The semantic tradition from Kant to Carnap

To the Vienna Station

J. ALBERTO COFFA

Edited by Linda Wessels
Indiana University

CAMBRIDGE
UNIVERSITY PRESS

This major new publication is a history of the semantic tradition in philosophy from the early nineteenth century through its incarnation in the work of the Vienna circle, the group of logical positivists that emerged in the years 1925–35 in Vienna and was characterized by a strong commitment to empiricism, a high regard for science, and a conviction that modern logic is the primary tool of analytic philosophy.

In the first part of the book, Alberto Coffa shows how the semantic tradition originated in opposition to Kant's theory that a priori knowledge is based on pure intuition and the constitutive powers of the mind, and was developed in the course of various foundational studies in mathematics, geometry, and logic.

In the second part, Coffa chronicles the development of this tradition by members and associates of the Vienna circle. He examines the attempts by Schlick and Reichenbach to interpret Einstein's theory of relativity, the search by Carnap and Tarski for a definition of truth, the new semantic theory of the a priori offered by Wittgenstein and Carnap, the logical positivists' transcendental approach to epistemology, and their failure to deal adequately with realism and the empirical foundations of knowledge. Much of Coffa's analysis draws on the unpublished notes and correspondence of these philosophers.

However, Coffa's book is not merely a history of the semantic tradition from Kant "to the Vienna Station." The author critically reassesses the role of semantic notions in understanding the ground of a priori knowledge and its relation to empirical knowledge; he also questions the turn the tradition has taken since Vienna.

"This is an extraordinary and long-awaited book. Its history is meticulous, but it treats its subjects as philosophers to be responded to and even fought with, rather than as museum pieces. It is worth the wait."
Richard Creath, Arizona State University

"Alberto Coffa is an intellectual pioneer. His book is the first comprehensive treatment of the development of logical positivism that is rigorous and sophisticated from both a historical and a technical point of view. It will constitute an indispensable basis for all future research in the area."
Michael Friedman, The University of Illinois at Chicago

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211, USA
10 Stamford Road, Oakleigh, Victoria 3166, Australia

First published 1991
First paperback edition 1993

Library of Congress Cataloging-in-Publication Data

Coffa, J. Alberto.
The semantic tradition from Kant to Carnap : to the Vienna Station / J. Alberto Coffa ; edited by Linda Wessels.
p. cm.
ISBN 0-521-37429-4
1. Semantics – History. 2. Kant, Immanuel, 1724–1804 – Contributions in semantics. 3. Bolzano, Bernard, 1781–1848 – Contributions in semantics. 4. Frege, Gottlob, 1848–1925 – Contributions in semantics. 5. Vienna circle. I. Wessels, Linda. II. Title.
B840.C58 1991
149'.94 – dc20 90-40588
CIP

British Library Cataloguing in Publication Data

Coffa, J. Alberto
The semantic tradition from Kant to Carnap : to the Vienna Station.
1. Meaning. Philosophical perspective, history
I. Title II. Wessels, Linda
121.6809

ISBN 0-521-37429-4 hardback
ISBN 0-521-44707-0 paperback

Transferred to digital printing 2003

For
Theresa and Justa
and
Diana and Julia

Contents

Editor's preface

On Christmas Eve, 1984, Alberto declared that a "good penultimate draft" of his book would be finished by the end of the year. The day after Christmas he became ill, and in the early morning hours of December 30, he died. The typescript Alberto left behind was, indeed, almost completed: The research had been finished and the arguments and theses were already in place; all but the introduction and the last chapter had been written in full, and extensive notes for those had been drafted. Some portions of the typescript had already been carefully crafted, and the intended shape of much of the rest was clear. In many places Alberto's dry wit showed through – you could almost see behind the prose the tilt of his head, the lopsided smile, the brief twinkle in his eye. With the help of a number of people, that typescript has been readied for publication. Using the notes Alberto left, I have completed the introduction and final chapter. Repetitions, digressions, and minor errors have been removed, theses and arguments delineated more clearly, grammar corrected, prose smoothed – all, I hope, without disturbing the twinkle. The result is not what Alberto would have produced, but perhaps it is something he would have found acceptable. He had intended to write a conclusion, discussing some of the implications of his study for contemporary philosophy. I have not attempted to write that conclusion.

Alberto began writing this book in the spring and summer of 1981 while a fellow of the Center for Philosophy of Science at the University of Pittsburgh. He was extremely grateful to the center for the time that allowed him to begin this project, and to his colleagues in the Department of History and Philosophy of Science, Indiana University, for both the time and the supportive and congenial environment that allowed him to continue his work. During the writing of this book, Alberto had fruitful discussions with many people; many also supplied him with information about useful sources and materials, and provided both intellectual and spiritual support. I cannot begin to compile a complete list of people he would have thanked, but certainly on that list would have been those with whom he had long and regular conversations on the philosophical matters that are central to this book: Thomas M. Simpson, Eduardo Garcia

Belsunce, Hector Castañeda, Simon Blackburn, and Alberto's students Frank Pecchioni and Tom Oberdan. There is no doubt that he would also have gratefully acknowledged the invaluable and constant support and encouragement of Adolf Grünbaum, his teacher and friend. To the many others whose assistance and influence should be acknowledged, I apologize for not including your names and thank you for the help you gave Alberto.

My own thanks go first to Gordon Steinhoff, who took on the heroic task of tracking down the sources of the quotations in the typescript, checking the quotations and translations for accuracy, completing the citations and references, and compiling the bibliography. I also gratefully acknowledge the assistance of Michael Friedman, who read the typescript at two different stages of my work on it and gave me numerous valuable suggestions for correcting and editing it. In addition, I thank Nicholas Griffin and an anonymous referee for their editorial suggestions, and Eduardo Garcia Belsunce, Edward Grant, and Tom Oberdan for lending their expertise where needed. Finally, I thank John Winnie for his suggestions and help on the typescript, and for the encouragement and support that sustained me through the long process of bringing it to publication.

Acknowledgments

Quotations from unpublished material are drawn from four sources: the Archives of Scientific Philosophy in the Twentieth Century, University of Pittsburgh Libraries; the Bertrand Russell Archives, MacMaster University; the Vienna Circle Archive, University of Amsterdam; the Wittgenstein Papers, microfilm distributed by Cornell University Library. My special thanks to Gerald Heverly, curator, and Stephen Wagner, research assistant, of the Archives of Scientific Philosophy, University of Pittsburgh Libraries, and to A. J. Kox, executive secretary of the Vienna Circle Foundation, University of Amsterdam, for their help in locating items in the Archives of Scientific Philosophy and the Vienna Circle Archive, respectively, and in checking the accuracy of quotations from material in their archives; thanks also to Nicholas Griffin for tracking down items in the Bertrand Russell Archives for me.

Passages from items in the Archives of Scientific Philosophy in the Twentieth Century that were authored by Rudolf Carnap, Frank P. Ramsey, and Hans Reichenbach are quoted by permission of the University of Pittsburgh; all rights reserved. The quotation from a letter by Kurt Gödel is included by permission of the Institute for Advanced Study, Princeton, N.J. Quotations from letters by Moritz Schlick are included by permission of the Vienna Circle Foundation.

Quotations from items in the Bertrand Russell Archives are included by permission of the Bertrand Russell Archives Copyright Permissions Committee.

Passages from items in the Vienna Circle Archive that were authored by Moritz Schlick are quoted by permission of the Vienna Circle Foundation. The unpublished letter by Rudolf Carnap from which a quotation is drawn has been copyrighted by Hanna Carnap Thost; all rights reserved. The quotation from that letter is included with the kind permission of Mrs. Thost. The quotation from a letter by Albert Einstein is included by permission of The Hebrew University of Jerusalem, Israel. Passages from a letter by Hans Reichenbach are quoted by permission of Maria Reichenbach.

Quotations from unpublished and posthumously published works by Wittgenstein are included with the kind permission of G. E. M. Anscombe and G. H. von Wright.

Introduction

The primary topic of this book is a decade in the philosophical life of what might loosely be called Vienna. Between 1925 and 1935 in the neighborhood of Vienna, the usually sluggish step of the Spirit suddenly quickened as some of his most enlightened voices started talking to one another. Wittgenstein, Tarski, Carnap, Schlick, Popper, and Reichenbach were, perhaps, no wiser than some of their contemporaries, but circumstances led them to interact during that decade, and the result of that dialogue deserves our attention.

When I started writing this book, I intended to explain in the preface that this was the history of epistemology since Kant, the way Carnap would have written it had he been Hegel. Since then I have come to think that while the Spirit may not be malicious, he is certainly forgetful. In Vienna he took a decisive step forward on the subject of the a priori; but he also moved sideways and backward on other crucial matters. Most of his erratic behavior could have been avoided had he been aware of some of his achievements in the preceding century. He may, perhaps, be excused since his best insights were due to the least noticeable of his voices.

Within the field of epistemology one may discern three major currents of thought in the nineteenth century: positivism, Kantianism, and what I propose to call the semantic tradition. What distinguished their proponents primarily was their attitude toward the a priori. Positivists denied it, and Kantians explained it through the Copernican turn. The semantic tradition consisted of those who believed in the a priori but not in the constitutive powers of mind. They also suspected that the root of all idealist confusion lay in misunderstandings concerning matters of meaning. Semanticists are easily detected: They devote an uncommon amount of attention to concepts, propositions, senses – to the content and structure of what we say, as opposed to the psychic acts in which we say it. The others cannot see the point of wasting so much time on semantic trivia.

It would be hard to find a more crucial epistemological problem than that of the character of a priori knowledge. One of the basic intuitions behind almost every epistemology since Plato's is that there are two radically different types of claims: the a priori and the rest. In pre-Kantian philosophy, many had tacitly assumed that the notion of analyticity provided the key to that of apriority. Kant saw that a different account was needed since not every a priori judgment is analytic, and offered a new theory based on one of the most remarkable philosophical ideas ever produced: his Copernican turn. In addition to this, Kant placed the idea of pure intuition at the center of his account of the scientific a priori. Positivists could not accept the consequences of this vision and saw no way out of the dilemma other than to deny the existence of the a priori, even in the case of logic.

Faced with the Scylla of asserting that 2 + 2 = 4 is empirical and the Charybdis of explaining it through the operations of pure intuition, semanticists chose to turn the boat around and try to find a better route. That there is a priori knowledge – even of the synthetic type – was indubitable to all of them; but most semanticists regarded the appeal to pure intuition as a hindrance to the development of science. Part I describes the stages through which it came to be recognized that pure intuition must be excluded from the a priori sciences and that consequently the Kantian picture of mathematics and geometry must be replaced by some other.

Our story begins with Kant's views on analysis and some of his reasons for concluding that one must appeal to pure intuition in connection with the a priori (Chapter 1). We then turn to the leading episodes that undermined that conviction. Bolzano's reductionist project (Chapter 2), complemented by Frege's and Russell's logicist projects (Chapters 4 and 6), challenged Kant's convictions in the field of arithmetic. Helmholtz (Chapter 3), Poincaré, and Hilbert (Chapter 8) provided the decisive contributions to analogous developments in the field of geometry. By the end of the nineteenth century, it had become clear that a priori knowledge could not possibly be what Kant thought it was. Early in the twentieth century, the special and general theories of relativity appeared to pose yet another challenge to the Kantian view, now from the domain of physics (Chapter 10).

Semanticists were primarily interested not in showing that Kant had not solved the problem, but in solving it themselves. The basic assumption common to all members of that movement was that epistemology was in a state of disarray due primarily to semantic neglect. Semantics, not metaphysics, was their *prima philosophia*. In particular, they thought, the key to the a priori lay in an appreciation of the nature and role of concepts, propositions, and senses. Though no defensible doctrine of the a priori emerged from their writings (Chapter 7), the ground was laid by a patient sharpening of semantic insights in the writings of

Bolzano, Frege, Husserl, Russell, and the early Wittgenstein (Chapters 2, 4, 5, 6, and 8). Against this background, Wittgenstein and Carnap offered in the early 1930s the first genuine alternative to Kant's conception of the a priori (Chapters 13, 14, 15, and 17). Their view that meaning is responsible for the a priori was that period's most decisive contribution to philosophy.

Logical positivism started as a branch of neo-Kantianism that differed from other branches in taking science as an epistemological model (Chapters 9, 10, and 11). During the 1920s, the early members and associates of the group slowly broke away from their Kantian beginnings, Schlick and Reichenbach as they struggled to interpret the lessons of the new theory of relativity (Chapter 10), Carnap as he tried to develop his epistemological ideas as a theory of constitution (Chapters 11 and 12). Out of their high regard for science emerged a second major contribution of the Vienna group – a transcendental approach to epistemology, a new "philosophy of science" (Chapters 10, 17, and 18).

The Copernican turn that had inspired Kant's analysis of the a priori had also led to a theory of experience and an understanding of the link between knowledge and reality that naturally led to idealism. In the nineteenth century, many wanted to avoid idealism, but few knew how to do it short of refusing to think through the consequences of their convictions. Semanticists suspected that if Kant's tacit semantic premises were granted, then certain Kantian insights on the role of constitution in knowledge could only be interpreted as leading to idealism. Once again, they thought, the key to a reasonable attitude was a clear-headed semantics.

Empiricists have traditionally flirted with meaning but, in the end, have remained hostile to it. When meaning becomes anything more than a topic for oblique allusion, when it becomes an explicit subject of research, it appears as an alternative to empirical considerations. It begins to look like a factual domain that is impervious to scientific research. Those associated with the Viennese movement were above all empiricists and shared the traditional empiricists' horror of meaning. Unequipped with meaning, they found it difficult to avoid idealism (Chapters 9 and 10). Carnap came closer than the others to making sense of realism, but his distaste for all things metaphysical also prevented him from completing the incorporation of meaning into empiricism (Chapters 12 and 17). In the end logical positivism remained without meaning. The natural consequence was the debate in the early 1930s over "foundations of knowledge," which was not really about foundations at all, but about the link between what we know and the world (Chapter 19).

Our picture of the Viennese developments of Part II will not be balanced unless we bear in mind the truths and falsehoods the participants learned from the three great nineteenth-century traditions that together shaped their standpoints. In all fairness, this book should have included, in addition to Part I, two other introductory sections devoted to Kantianism and positivism. The finitude of my life, my mind, and my reader's patience were factors to consider. There was also the fact that nineteenth-century Kantianism and positivism are far better known than their less celebrated rival. Finally – why not admit it? – the proportion of insight to confusion is far, far greater among semanticists than among their more prestigious and respected colleagues.

PART I

The semantic tradition

1

Kant, analysis, and pure intuition

> It was disastrous that Kant . . . held the domain of the purely logical in the narrowest sense to be adequately taken care of by the remark that it falls under the principle of contradiction. Not only did he never see how little the laws of logic are all analytic propositions in the sense laid down by his own definition, but he failed to see how little his dragging in of an evident principle for analytic propositions really helped to clear up the achievements of analytic thinking.
>
> Husserl, *Logische Untersuchungen*, vol. 2, pt. 2

For better and worse, almost every philosophical development of significance since 1800 has been a response to Kant. This is especially true on the subject of a priori knowledge. The central problem of the *Critique* had been the a priori, and Kant had dealt with it from the complementary perspectives of judgment and experience. His "Copernican revolution" gave him a theory of experience and a non-Platonist account of the a priori. But when the *Critique* was well on its way, Kant discovered the notion of a synthetic a priori judgment, and he saw in this a particularly appealing way of formulating his project as that of explaining how such judgments are possible.

The constitutive dimension of Kant's theories of experience and the a priori will figure prominently in later developments. As we shall see, one of the turning points in our story will involve a Copernican turn, though the issue it concerns will be different from Kant's. Moreover, the early stages of logical positivism may be viewed as a development to the point of exhaustion of this aspect of Kant's original idea. In this chapter, however, we shall focus exclusively on the more superficial aspect of Kant's treatment of the a priori, involving synthetic a priori judgments; for it was the shallowness of Kant's treatment of this matter that led to doctrines which, in turn, elicited the semantic tradition.

One of the central points of agreement among the members of the semantic tradition is the idea that the major source of error behind Kant's theory of knowledge – especially of the a priori – is his confused doctrine of meaning and that the key to a correct doctrine of the a priori is the understanding of semantics. Our purpose in this first chapter is to examine the relevant features of Kant's epistemology and its semantic background. Our first problem will be to uncover Kant's semantic views.

In a sense, of course, he hadn't any; for part of the story we aim to tell is that of how semantics was born. In another sense he did, of course, for he was bound to have opinions, however tacit and unacknowledged, on what it is to convey information, on when we may succeed in doing so, and when we are bound to fail. Philosophers have often thought these topics unworthy of much attention. The analytic tradition that extends from Bolzano to Carnap places meaning at the heart of philosophy – or, rather, it finds that it has been there, unrecognized, all along and that the failure to think seriously about it is the root of the reductio ad absurdum of rationalism displayed in Kant's philosophy and its idealist offspring. The question is, Where does one look for the tacit semantics of those who did not address that topic explicitly?

In one of the many aphorisms Quine aimed at the semantic tradition, he noted that "meanings are what essences became . . . when wedded to the word." If this were right, those who would like to know what Kant thought about meanings would have to consult what he wrote about essences; since he wrote next to nothing on that subject, this would be the end of the search. Actually, meanings had a more honorable ancestor within the field of traditional logic, in the category of concepts or, more generally, representations. To find out what a post-Cartesian philosopher thought about meanings, we must look at the logic books he wrote or quoted from, for it is there that the notions of concept and judgment are treated. Meanings are what concepts became when wedded to the word.

Conceptual analysis

Kant was very proud of his distinction between analytic and synthetic judgments. He recognized that philosophers before him had understood the significance of the division between a priori and a posteriori judgments. But when Eberhard challenged his originality on the matter of analyticity, Kant replied, in an effort at irony, that everything new in science is eventually "discovered to have been known for ages" (Allison, *The Kant–Eberhard Controversy*, p. 154). Had he read Borges, he would have paraphrased him: "Great ideas create their ancestry" (see Borges, "Nathaniel Hawthorne," *Obras completas*, p. 678).

In fact, Kant had little reason to be proud. His treatment of the analytic–synthetic distinction is original in certain ways, as we shall see; but in the end, it is one of the least distinguished parts of his philosophy. In it, some long-standing confusions converge and others emerge, original with Kant; the latter were destined to have a long and damaging influence throughout the nineteenth century.

The picture of meaning dominant since the emergence of rationalism and empiricism took meanings to be inextricably associated with experience. It is plausible to think that in order to know the meaning of pain, love, rivalry, heroism, and so on, one must undergo certain experiences and that the more carefully one analyzes these experiences, the better one understands pain, love, and so on. It is but a small step to conclude that the meanings of 'pain', 'love', and so on consist precisely of those psychic phenomena that are the targets of our analysis. The same may be thought to hold for all expressions; they will mean something only insofar as, and to the extent that, they relate to human mental processes. Number expressions, for example, may be thought to derive their meaning from the mental processes in which they are involved – the natural numbers through processes of counting, geometric objects through acts of measurement, and so on.

On this way of looking at things, the basic semantic notion is that of "representations" (*Vorstellungen*) construed as "modifications of the mind" that "belong to inner sense" (Kant, *Critique*, A98–9), as mental states designed to represent something. A long tradition, canonized in the *Logique de Port Royal*, had declared ideas or representations the most important subject of logic, since "we can have knowledge of what is outside us only through the mediation of ideas in us" (Arnauld and Nicole, *La logique ou l'art de penser*, p. 63). In Leibniz's words, human souls "perceive what passes without them by what passes within them" (Clarke, *The Leibniz–Clarke Correspondence*, p. 83); indeed, "the nature of the monad" is "to represent" (Leibniz, "The Monadology" [1714], *Philosophical Papers and Letters*, pp. 648–9).

The word '*Vorstellung*' first became a technical term in Wolff's philosophy; it corresponded approximately to the earlier 'idea' and was intended to cover both intellectual and psychic processes. For Meier, author of the logic text that Kant followed in many of his courses on that subject, representations were "pictures or images (*Gemälde oder Bilder*) of those things that we represent to ourselves (*wir uns vorstellen*)" (Meier, *Auszug aus der Vernunftlehre*, sec. 24). In its pre-Kantian use, in Wolff, Lambert, and Meier, for example, 'representation' and 'concept' (*Begriff*) functioned as synonyms, and the pre-critical Kant largely followed this usage.[1]

One of the many ways philosophers have tried to understand meaning might be called the "chemical theory of representation," using an analogy occasionally found in the writings of Locke, Lambert, and even Kant. According to this theory, representations, like chemical compounds, are usually complexes of elements or "constituents," which may themselves be complex. Generally, when a representation is given to us, we are not consciously aware of this. Analysis is the process through which we identify the constituents of a complex representation. It is a process that must come to an end, after (perhaps) finitely many stages, in the identification of simple constituents. Moreover, the best way to know what a representation is, is to identify its constituents – preferably its ultimate simple constituents – and the form in which they are put together or linked to constitute the given representation. To know a concept fully, for example, is to define it; and definition (*Erklaerung*) is no more and no less than exhaustive and complete analysis.

Descartes's doctrine of ideas had promoted the notions of clarity and distinctness to the status of philosophical celebrities. Under the influence of the new rationalism, these two heterological notions soon came to be regarded as the highest virtue in the ethics of concepts and to figure prominently in the chapters of most logic textbooks. They took a more definite shape in the German philosophical tradition.

Even though representations are basically intended to represent other things, we can upon occasion turn the arrow of reference on them (Kant, *Critique,* A108). When we do so, when we become conscious of a representation, then, Kant said, it is "clear" (*klar;* e.g., *Logik,* p. 33). The much more crucial virtue of "distinctness" (*Deutlichkeit*) depends entirely on our mental relation to what Kant called the "manifold" given in representation. Consider intuitive representations first. If we intuitively represent (e.g., see) a house in the distance, we may not be consciously aware of the windows, doors, and other parts of it. But, claimed Kant, we surely see them, in some sense; for we know "that the intuited object is a house"; hence, "we must necessarily have a representation of the different parts of this house. . . . For if we did not see the parts, we would not see the house either. But we are not conscious of this presentation of the manifold of its parts" (*Logik,* p. 34; also *Logik Pölitz,* pp. 510–11; *Reflexionen zur Logik,* refl. 1676, p. 78; *Wiener Logik,* p. 841; Borges, "Argumentum Ornithologicum," *Obras completas,* p. 787). The venerable Wolff had praised "the great use of magnifying glasses towards gaining distinct notions" (*Logic,* pp. 27–8). Following his lead, Kant noted that when we look at the Milky Way with bare eyes, we have a clear but indistinct representation of it, since we do not see a discontinuous cluster of stars but, rather, a continuous streak of light. When we look at it through the telescope, however, our representation is (more) distinct (*Logik,* p. 35; also *Logik Pölitz,* p. 511). Echoing one of Leibniz's examples in his *Nouveaux essais* (bk. 2, chap. 2, sec. 1 and bk. 4, chap. 6, sec. 7), Kant illustrated the nature of a clear but indistinct representation: "Blue and yellow make green, but we are not always aware of the presence of these parts in green" (*Wiener Logik,* p. 841).[2]

Mutatis mutandis, the same is supposed to be true of conceptual representations. We may, for example, have a clear concept of virtue, and we may recognize some of its constituent marks without being entirely clear on what all or even most of them are. The process through which we achieve distinctness on this matter is precisely what Kant called analysis: "To analyze a concept [is] to become self-conscious of the manifold that I always think in it" (*Kritik,* B11/A7).

It is an essential element of Kant's doctrine of analysis that our understanding of the analyzed concept changes (for the better) during that process, whereas the concept does not:

> When we resolve the concept of virtue into its constituents, we make it distinct through analysis. In so making it distinct, however, we add nothing to the concept, we merely clarify it. (*Logik,* p. 35)

> When I make a concept distinct, then my cognition does not in the least increase in its content by this mere analysis . . . [through analysis] I learn to distinguish better or with greater clarity of consciousness what already was lying in the given concept. Just as by the mere illumination of a map nothing is added to it, so by the mere elucidation of a given concept by means of analysis of its marks no augmentation is made to this concept itself in the least. (*Logik,* p. 64)

As to definition, he had said in the *Wiener Logik* that it is "the most important chapter of logic" (p. 912) and went on to elaborate:

> All our concepts, insofar as they are given, whether *a priori* or *a posteriori,* can be defined only by means of analysis through dissection (*Zergliederung*). For, when it is given, I can make the concept distinct only by making the marks (*Merkmale*) it contains clear. That is what analysis does. If this analysis is complete . . . [and] in addition there are not so many marks, then it is precise and so constitutes a definition. (p. 914; see also *Logik Philippi,* p. 455)[3]

On pain of indistinctness, analysis must terminate after finitely many steps; one must therefore find simple, indefinable, unanalyzable concepts at the end of this process.[4] About these obviously crucial indefinables, Kant had suspiciously little to say, though he did explicitly note that besides being indefinable and unanalyzable, these marks are also indistinct (*Logik,* pp. 34–5; *Logik Philippi,* p. 342). Total distinctness is therefore achieved when a complex concept is reduced to its indistinct constituents. This peculiarity may arise from a mere terminological oddity. It is harder to understand Kant's insistence that clarity, the only logical virtue of simple concepts, is not a topic for logicians, or indeed philosophers, but concerns only the psychologist (see, e.g., *Kritik,* B414 note; *Anthropologie,* p. 137; *Untersuchung über natürlichen Theologie,* pp. 284, 286, 290).[5] This would not be the last time a philosopher would consign to psychology those portions of epistemology that threatened to run his philosophy aground.

The preceding quotations provide ample evidence of Kant's commitment to the chemical doctrine of the concept. But they also display another important feature of his understanding of analysis, a feature that his idealistic successors would obliterate but that would be emphasized by a different tradition in nineteenth-century philosophy. Unless we are prepared to regard Kant's explanations of conceptual analysis as thoroughly blundering attempts to convey his meaning, there is no way to avoid the conclusion that he was tacitly endorsing a distinction between the mental acts in which concepts are involved and those concepts themselves. If our understanding of the concept *virtue* can be bad at one time and good at another, these two different acts or states of understanding must somehow concern or involve the very same concept. Hence, in some sense of being, there must *be* a concept of virtue that is both the target of mental episodes and distinct from them. This concept need not be extrasubjective; but it must at least be intersubjective, since the very same conceptual representation is involved in different representings, or psychic acts of representation, in the same or different persons. To be sure, like his teachers and his followers, Kant did not observe that distinction consistently; but without it, it would be very hard to make sense of the things he said about analysis of concepts and about analytic knowledge. About the former, for example, he typically claimed, "By means of analytic distinctness we recognize in something no more than we had originally thought in it; instead we recognize better, i.e., more distinctly and clearly and with greater awareness, what we already knew" (*Logik Blomberg*, p. 131; see also *Critique*, A5–6/B9). Nor would it be possible to make sense of the countless references to the uncovering of tacit knowledge through analysis. Indeed, one of the few persistent themes throughout Kant's philosophy, from the early *Untersuchung über natürlichen Theologie* (1764) to his critical writings, was that philosophy is distinguished from other sciences in that its proper method is analysis of concepts, bringing forth or to the surface hidden knowledge, rather than constructing new knowledge. As in the Socratic model, the philosopher's task is to help people become aware of what they have known all along: "If we only knew what we know . . . we would be astonished by the treasures contained in our knowledge" (*Wiener Logik*, p. 843).

On one occasion, Kant went so far as to raise explicitly what we now call the "problem of analysis," the issue of the identity of *analysandum* and *analysans*, and his revealing answer provides evidence of his insecure grasp of the situation:

> Is the concept in the definition totally identical to that defined [through the process of analysis]? . . . we must bear in mind: *materialiter*, i.e., *quoad objectum*, these concepts are always completely identical; only concerning the form they are not, indeed, they should be not entirely identical; concerning the matter, I always think the same object, only not in the same way, but in a different way; what before the definition I represented confusedly, I now represent clearly. (*Logik Blomberg*, p. 265)

The distinction between matter and form – that root of so much philosophical confusion – is "explained" in the *Logik Philippi* as follows: "When I look at a worm under the microscope, the form of the worm changes but the object remains the same. . . . All philosophy concerns only form, since we consider an object piecemeal and become more clearly aware of the matter contained in it" (p. 341). This would appear to imply that in analysis the concepts with which we are concerned both before and after the analysis are the same but that our mode of knowledge of them is different (although these and similar passages may also be taken to display a confusion between a concept and its objects).

Even though Kant's views on the nature of hidden meaning and tacit knowledge called for a distinction between act and content in representations, it is also true that he often appeared to disregard that distinction and that whenever specificity was demanded, he chose to fall on the purely subjective side of the dichotomy. Had Kant been more sensitive to the distinction and to its overwhelming importance, he would have noticed that the link between concepts and analysis was much weaker than he thought. We shall see that as Kant went on to extend the ideas of analysis and synthesis from concepts to judgments, his emphasis on the subjective element in representation, to the neglect of its objective counterpart, combined with the chemical picture of concepts to produce a peculiarly Kantian confusion.

Analytic judgments

For Kant, the link between the analyticity of judgments and conceptual analysis was immediate. To begin with, "in a judgment, two concepts are in a relation" (*Philosophische Enzyklopädie*, p. 19), and in categorical judgments, the link is the subject–predicate relation. Categoric judgments are "the matter of all others" (*Reflexionen zur Logik*, refl. 3046, p. 631). Thus, all judgments have as their "matter" either concepts or other judgments (refl. 3046). A categorical judgment is analytic, Kant claimed, when the predicate concept is thought implicitly in, or is contained in, the subject concept; all other categorical judgments are synthetic (e.g., *Logic*, p. 117; *Prolegomena to Any Future Metaphysics*, p. 14).[6] Thus, an analytic judgment is the expression of the outcome of conceptual analysis.

Given its sources in the idea of conceptual analysis, it is hardly surprising that Kant's definition applies to judgments in the subject–predicate form. (Here 'subject' is construed in the traditional pre-Fregean manner, so that the subject of 'All As are Bs' is 'All As' or 'A'.) The familiar problems posed by the narrow scope of this definition do not concern us here, since they fail to reveal any serious flaws in Kant's picture of things. What does concern us is the much more crucial matter of the link between conceptual analysis and analytic judgments. These judgments are derived, according to Kant, "by dissection of the [subject] concept" (*Prolegomena to Any Future Metaphysics*, p. 17); they "merely break up the [subject concept] into those constituent concepts that have all along been thought in it" (*Kritik*, B11). A paradigm example of an analytic judgment is "Every x that conforms to the concept $(a + b)$ of body also conforms to (b) extension" (*Logik*, p. 111). "In the analytic judgment we keep to the given concept, and seek to extract something from it" (*Critique*, A154/B193).

Superficially, it might seem that Kant was saying little more than the many others before him who had also been concerned with the issue of conceptual analysis. Eberhard, for example, thought the notion of an analytic judgment was clearly present in Leibniz's writings. This evaluation entirely misses the element of novelty that Kant had incorporated into the chemical doctrine of the concept. The difference between Kant's position on this issue and those of his predecessors becomes clear when we examine their answers to the question, How do we determine the constituency of a concept; what criteria determine whether a concept B is "in" the concept A?

When Kant started to think about this question, there were two standard answers, one emerging from a long and venerable tradition, the other first put forth by Leibniz. The Leibniz–Arnauld correspondence clearly displays the conflict between these two standpoints. With his characteristic blend of genius and insanity, Leibniz had conceived a project in which the simple constituents of concepts would be represented by prime numbers and their composition by multiplication. From the Chinese number theorem (and certain assumptions about the nature of truth) he inferred that – given this "perfect language" – all matters of truth could be resolved by appeal to the algorithm of division. "For example," he explained,

> if the symbolic number of man is assumed to be 6 and that of ape to be 10, it is evident that neither does the concept of ape contain the concept of man, nor does the converse hold. . . . If, therefore, it is asked whether the concept of the wise man is contained in the concept of the just man . . . we have only to examine whether the symbolic number of the just man can be exactly divided by the symbolic number of the wise man. (*Logical Papers*, p. 22)

This procedure allows us to solve all questions concerning the truth value of universal affirmative propositions if we assume, with Leibniz, that in the true instances "the concept of the subject, taken absolutely and indefinitely, and in general regarded in itself, always contains the concept of the predicate" (*Logical Papers*, p. 22).

In response to Leibniz's astonishing claim that in every true proposition, whether necessary or contingent, the predicate is contained in the subject, Arnauld defended the historical view of the matter: In order for the predicate B to be in A, what is required is not just the truth but the necessity of 'All As are Bs'.

Sometime in the 1770s, Kant came to the conclusion that analyticity is neither truth (as for Leibniz) nor necessity (as for Arnauld) but something stronger than both: What is contained in a concept is less than what is true of it and even than what is necessarily true of its objects; to put it differently, analyticity is one thing and apriority another. It was then that he saw that there are a priori truths not grounded on conceptual analysis, that there are, as he chose to call them, synthetic a priori judgments. With this insight, his conception of philosophy changed radically. Earlier he had thought that the method of philosophy was analysis and that analysis could only ground analytic claims. Now he decided that philosophy was also, perhaps even predominantly, aimed at examining the foundations of very different sorts of judgments, those that are a priori but not analytic.

It would be hard to exaggerate the significance Kant attached to this discovery. The introduction of the *Critique* calls for a new science (sec. 3) designed to answer this previously unnoticed question: How can we have a priori knowledge of propositions in which the predicate concept is not part of the subject concept? "A certain mystery lies here concealed; if it had occurred to any of the ancients even to raise this question, this by itself would . . . have been a powerful influence against all systems of pure reason" (*Critique*, A10). But the existence of such remarkable judgments had been entirely overlooked, Kant thought. The fact that mathematical judgments were not analytic, for example, "has hitherto escaped the notice of those who are engaged in the analysis of human reason, and is, indeed, directly opposed to all their conjectures" (*Critique*, B14).

Whereas much of the *Critique* must have been written, or at least conceived, by the time he came to this new vision of conceptual analysis, Kant chose to place the consequences of this vision at the very beginning of the *Critique* and the *Prolegomena*. When Eberhard challenged Kant's originality on the subject of analysis and synthesis, Kant was furious, exhibiting his feelings in a celebrated polemical blast. And when in 1791 he drafted a response to a question posed by the Academy on what progress had been made in German philosophy since Leibniz and Wolff, Kant observed that "the first step" of the new critical philosophy had been to draw the analytic–synthetic distinction. He added, "Had this been clearly known in the time of Leibniz or Wolff, we would have found it not only reported but also emphasized as important by the treatises of logic and metaphysics" (*Preisschrift über die Fortschritte der Metaphysik* [1804]; see also *Critique*, B19).

Even so, an important question remains. Why didn't Kant think that his distinction was an utterly trivial consequence of the notion of conceptual analysis? The answer I would propose is this: When that distinction was conjoined with Kant's casual understanding of semantic matters, it appeared to entail nothing less than the Copernican turn. Once we realize that we know a priori some claims that cannot be grounded on a mere understanding of their content, it becomes clear that the things about which we have such knowledge cannot be as mind-independent as they were thought to be.

The heart of the problem is Kant's seemingly harmless assumption that the analytic–synthetic distinction is a correct explication of another one, between clarificatory judgments (*Erlaeuterungsurteile*) and ampliative judgments (*Erweiterungsurteile*). In all likelihood, Kant never realized that he was dealing with two different distinctions. Thus, some of his "definitions" of analytic and synthetic judgments tell us that the latter "extends my knowledge beyond what is contained in the [subject] concept" (Allison, *The Kant–Eberhard Controversy*, p. 141; see also *Logik*, p. 111; *Critique*, A8; *Prolegomena to Any Future Metaphysics*, sec. 2a). But it is essential to realize that we are dealing here with a second partition of the class of all (true, subject–predicate) judgments into those that we may ground, or identify as true, merely on the basis of the fact that we are clear about the concepts involved in the judgment, and those (other judgments) that call for an appeal to extraconceptual sources of knowledge. Roughly speaking, whereas Kant's first, *nominal* definition characterizes 'analytic' as true in virtue of definitions (analysis) and logic, the second one defines it as true in virtue of meaning.

The idea that his nominal definition coincides with the second version of 'analytic' is based on an assumption that apparently seemed so obvious to Kant that it did not deserve the slightest argument: *Concepts can provide a basis for knowledge only through a process of analysis.* Thus, he claimed that in a synthetic judgment,

> I must advance beyond the given [subject] concept, viewing as related to it something entirely different from what was thought in it. This relation [between the subject and the predicate concepts] is consequently never a relation of identity or contradiction: and *from the judgment taken in and by itself*, the truth or falsity of the relation can never be discovered. (*Kritik*, A154–5/B193–4; my italics)

Elsewhere he put it more concisely: "It is evident that from mere concepts only analytic knowledge . . . is to be obtained" (*Critique*, A47/B64–5).[7] We shall return to examine the immense damage caused by the preceding confusion. But for now we shall proceed with the course of Kant's reasoning. Once the true synthetic judgments of Kant's nominal definition are confused with judgments not grounded on purely conceptual knowledge, the obvious question to ask is, What *are* they grounded on?

Kant had explained that all analytic judgments are grounded on a single principle, what he sometimes called the "principle of analytic judgments" (e.g., *Critique*, A149–50/B189), the principle of identity or contradiction. What he had in mind, presumably, was a principle allowing one to predicate of a given concept those other concepts that we "think" in it as constituents. However this may work, the interesting point is that Kant assumed that his analytic–synthetic distinction characterized, as it were, epistemic natural kinds, so he felt justified in concluding that there had to be another principle involved in the grounding of all synthetic judgments. He called it, of course, the "highest principle of all synthetic judgments" (see, e.g., *Critique*, A154/B193; A158/B197).

In synthetic subject–predicate judgments, we put together two concepts that are not related as part and whole. Having casually and unknowingly consigned all semantic grounds of knowledge to the category of conceptual analysis and thus to nominal analyticity, Kant did not even consider the possibility that synthetic judgments, nominally construed, might also have a semantic ground. "In synthetic judgments," Kant thought, "I must have besides the concept of the subject [and that of the predicate] something else (X), upon which the understanding may rely, if it is to know that a predicate, not contained in this concept, nevertheless belongs to it" (*Critique*, A8). Thus, he concluded that the synthesis of disjoint concepts could never be due to a link provided by the conceptual constituents of the judgment, but should always be mediated by a third element, an X, as he sometimes called it (e.g., *Critique*, A9/B13), not directly present in the judgment. This X could not possibly be a concept, Kant thought, since we would then have in addition to the subject and the predicate concepts, a third concept, and "from mere concepts only analytic knowledge . . . is to be obtained." Since Kant recognized no semantic building blocks other than concepts and intuitions, it followed that the ground of all synthetic knowledge, the glue that links the concepts in a synthetic judgment, must always involve intuition. This is the content of the principle of synthetic judgments: Synthetic judgments "are only possible under the condition that an intuition underlies the concept of their subject" (Allison, *The Kant–Eberhard Controversy*, p. 152; also see the letter to Reinhold, ibid., p. 164). For example, after arguing that 7 + 5 = 12 is not analytic, Kant added that in order to ground this judgment, "we have to go outside these concepts, and call in the aid of the intuition which corresponds to one of them, our five fingers, for instance" (*Critique,* B15). In all mathematical judgments, "while the predicate is indeed attached necessarily to the concept, it is so in virtue of an intuition which must be added to the concept" (*Critique,* B17).

Kant was not particularly modest about his discovery of this "highest principle of all synthetic judgments," designed to solve "the most important to all questions" of transcendental logic – indeed, perhaps "the only question with which it is concerned": How are synthetic a priori judgments possible? (*Critique,* A154/B193). As he wrote to Eberhard in bitter defiance:

> It was therefore not merely a verbal quibble, but a step in the advance of knowledge, when the *Critique* first made known the distinction between judgments which rest entirely on the principle of identity or contradiction, from those which require another principle through the label 'analytic' in contradistinction to 'synthetic' judgments. For the notion of synthesis clearly indicates that something outside of the given concept must be added as a substrate which makes it possible to go beyond the concept with my predicate. Thus, the investigation is directed to the possibility of a synthesis of representations with regard to knowledge in general, which must soon lead to the recognition of intuition as an indispensable condition for knowledge, and pure intuition for a priori knowledge. (Allison, *The Kant–Eberhard Controversy,* p. 155)[8]

Kant's doctrine of pure intuition had multiple origins. We have identified two: the principle of synthetic judgments and the thesis of synthetic a priori knowledge.[9] The reasoning proceeds as follows: It is clear, to begin with, that in the nominal sense of 'analysis' there are a great many synthetic judgments that very few people would seriously regard as a posteriori. One could cite, with Kant, the examples of arithmetic and geometry, but there are more pedestrian examples, such as 'If this is red, then it is not blue', 'If this is one meter long, then it is not two meters long', and 'If *a* is taller than *b* and *b* than *c*, then *a* is taller than *c*'. In none of these judgments does the subject contain the predicate(s). Yet they are all surely necessary and hence, according to Kant, a priori. Moreover, by Kant's criterion, every judgment with a simple concept must be synthetic, and surely some such judgments are necessary.[10] Thus, using his nominal definition, Kant had no difficulty identifying synthetic a priori judgments. Indeed, the only considerations to be found in Kant's writings that resemble arguments for the existence of synthetic a priori judgments invariably appeal to the nominal version of Kant's distinction – arguing, quite plausibly, that this or that predicate concept is "obviously" not a constituent or not "thought in" this or that subject concept.

One can readily grant to Kant everything he wanted up to this point. But the next step in his reasoning is, in effect, to confuse synthetic in the nominal sense with ampliative, by assuming that synthetic judgments in the nominal sense cannot be given a purely conceptual ground. The reasons for this error will be examined later, but now we can see the critical roller coaster well on its way: *Something* must ground synthetic judgments, and it cannot be concepts; hence, it has to be intuitions, such as empirical intuitions of the sort Hume liked. But we have now discovered that some synthetic judgments are a priori, so they cannot be grounded by an empirical intuition. Hence, there must be a very special, nonempirical sort of intuition – let us simply call it 'pure intuition'.[11]

We have said that bad semantics was at the root of Kant's appeal to pure intutition. Having examined the motives that may have led him to confuse his two senses of analytic, it is possible to diagnose the problem somewhat more precisely, as emerging from a psychologistic conception of semantics. Consider, for example, the question of whether we have to understand the concept *bachelor* in order to ground the judgment that all bachelors are unmarried. One may imagine Kant's train of thought running along the following lines: Understanding the concept *bachelor* is surely necessary in order to engage in its analysis; but analytic judgments are no more and no less than the outcome of analysis. Therefore, an understanding of the concept – and not just of its structural features or its logical constituency – must be relevant to the grounding of analytic judgments. Or he may (also) have argued: To know or understand a concept is to know its definition; hence, conceptual knowledge is knowledge in virtue of definitions. Since analytic knowledge is exactly knowledge in virtue of (partial) definitions, all purely conceptual knowledge must be analytic.

These appealing but fallacious kinds of reasoning lose all of their attraction once one observes the act–content distinction.[12] Consider again

(*) All bachelors are unmarried.

According to the distinction drawn in the preceding section, two radically different acts of judgment correspond to this sentence; whether we associate with (*) one or the other will depend on how distinct our representation of *bachelor* is. If our representation is entirely distinct (and if *unmarried and male and adult* is the complete analysis of *bachelor*), then the subjective judgment we express by (*) may also be conveyed – even more explicitly – by means of

(**) All unmarried male adults are unmarried.

But if our representing of *bachelor* is indistinct, then the judgment we express by means of (*) will, in general, be quite different from the one we express by means of (**).[13] We may therefore say that if our understanding of *bachelor* is indistinct, the mental episodes expressed by (*) and (**) are quite different and that one purpose of analysis is to lead us from the sorts of mental states associated with (*) to those associated with (**). But it should be equally obvious that, psychology aside, the *contents* of (*) and (**) are, by Kant's own standards, identical, since it would otherwise be nonsense to insist – as Kant did – that only our understanding of the concept has changed through the process of analysis, not what we were saying when we used it.

We are now ready to examine the extent to which our conceptual knowledge of the constituents of the subject concept in (*) is involved in the grounding of the judgment as analytic. When (*) is regarded as an expression of our judgment as produced in a state of indistinct understanding of *bachelor,* it is indeed reasonable to say that our improved understanding of that concept, our understanding of the meaning of 'bachelor', is essential to the grounding of (*), that is, for reaching the conviction that (*) is true. But consider now (*) as expressing our judgment in a state of distinctness toward *bachelor* or, again, consider the content of (*). To what extent does our understanding of the concepts involved play any role in *its* grounding? Or, to put this question differently, could we ground the content of (*) – that is, of (**) – without an understanding of all of the concepts involved? The answer is obvious when we look at (**): of course! All we really have to understand in order to provide a ground for (**) is the meaning of the concepts *conjunction* and *predication.* Thus, whereas the grounding of (*) *qua* indistinct subjective judgment through conceptual analysis calls for our understanding of all concepts involved, the grounding of what (*) says does not.[14]

We can now see how Kant's neglect of the nonpsychological dimension of semantics may have led him to confuse the analytic and the purely conceptual; for he would have been correct to think that an understanding of all concepts is essential for analysis (in the psychological sense) and that analytic judgments are the outcome of analysis. The shift from psychology to semantics is fatal to Kant's reasoning. The analysis of a concept does, indeed, require an understanding of it; but the grounding of an analytic judgment *qua* content calls only for an understanding of what we have earlier called its structure.

Another way of stating Kant's problem is to say that he confused conceptual knowledge with definitional knowledge; that is, he confused what can be grounded on concepts with the much smaller subclass of what can be grounded on definitions. As Kant saw it, analytic knowledge is possible only in the presence of conceptual complexity, but it should

have been clear that simple concepts, unaided by intuition, are as apt as their complex counterparts to act as grounds of a priori knowledge.

We have detected two tacit assumptions behind Kant's dealings with the analytic and the synthetic: According to the first, analytic coincides with true in virtue of concepts – or, as some would say much later, in virtue of meanings. On this assumption, considerations of a semantic sort are relevant to the establishment of only those judgments whose predicate is part of their subject. This implies that the ground of synthetic judgments does not lie in semantics. The second assumption tells us where it does lie. Given Kant's views on the nature of representation, it can only be assumed that the ground of synthetic knowledge is intuition – in the interesting cases, pure intuition.

Only through a complex and laborious process that took most of the nineteenth century did these Kantian confusions come to be recognized and neutralized. In the rest of Part I we shall look at the central stages of this process. It can be characterized most simply as the decline and fall of pure intuition. However much they may have disagreed on specific issues, the leading members of the anti-Kantian tradition we shall examine shared the conviction that Kant's system was built on a semantic swamp. They also agreed that the only way to avoid a similar fate was to place the theory of meanings, that is, the theory of concepts, judgments, and propositions, at the very top of their list of philosophical concerns. Semantics was born in the effort to avoid Kant's theory of the a priori. It was born in the writings of Bolzano.

2

Bolzano and the birth of semantics

All mathematical truths can and must be proven from mere concepts.

Bolzano, *Grossenlehre*

Kant was incorrect when he took logic to be complete.

Bolzano, *Gesamtausgabe,* ser. 2B, vol. 2, pt. 2

Modern Continental philosophy had always maintained close ties with scientific developments. In Kant the link became so close that the whole doctrine of the a priori had been motivated largely by a datum that had emerged from the sciences – an allegedly transparent feature of geometry, arithmetic, and the calculus that demanded philosophical explanation. Kant's successors in the nineteenth century were of two types: those who wanted to check whether what he said about the a priori sciences was true and those who didn't really care. The latter embraced his Copernican turn for "metaphysical" reasons. The former, by and large, devoted a great deal of time to an analysis of mathematical knowledge. As a result, their more gullible colleagues tended to look on them as low-level mathematicians trying to make a reputation in philosophy. "Mathematica sunt, non leguntur" is what Frege once guessed most philosophers would say about his writings. He was right. The same could have been said of the major writings of the semantic tradition.

The semantic tradition may be defined by its problem, its enemy, its goal, and its strategy. Its problem was the a priori; its enemy, Kant's pure intuition; its purpose, to develop a conception of the a priori in which pure intuition played no role; its strategy, to base that theory on a development of semantics.

If a theory is as sound as the problems it solves, it was reasonable to start a critical examination of the critical philosophy at the place where *it* had started, with an analysis of the character of the a priori knowledge from which Kant had derived his basic datum. The semantic tradition was not developed by people with a narrow-minded interest in the foundations of mathematics, but by those who suspected that Kant's understanding of arithmetic, the calculus, and geometry was based on irreparable misunderstandings and that these misunderstandings vitiated his general picture of the a priori. The following chapters sketch the story of the semantic tradition, a philosophical movement that, unlike positivism, took the a priori seriously and, unlike idealism, chose to look even more closely than Kant at his paradigm examples of the a priori.

While the idealists were removing every trace of objectivity from Kant's semantics, there was in a corner of the Austro-Hungarian empire, ignored by the leaders of German philosophy, a Czech priest by the name of Bernard Bolzano, who was engaged in the most far-reaching and successful effort to date to take semantics out of the swamp into which it had been sinking since the days of Descartes. Bolzano was the first to recognize that transcendental philosophy and its idealistic sequel were a reductio ad absurdum of the semantics of modern philosophy. He was also the first to see that the proper prolegomena to any future metaphysics was a study not of transcendental considerations but of what we say and its laws and that consequently the *prima philosophia* was not metaphysics or ontology but semantics. The development of these ideas in his monumental *Wissenschaftslehre* and in a variety of other writings established Bolzano as the founder of the semantic tradition.

Bolzano's philosophy was the kind that takes from and then gives life to science. His approach to semantics was developed in dialectical interplay with his decision to solve certain problems concerning the nature of mathematical knowledge. Kant had not even seen these problems; Bolzano solved them. And his solutions were made possible by, and were the source of, a new approach to the content and character of a priori knowledge. We shall illustrate the point by focusing on one of Bolzano's favorite mathematical topics, the calculus.

Intuition and the calculus

It would be grossly unfair to Kant to say that the main reasons he had for thinking that mathematics involves pure intuition were the semantic considerations examined in the preceding chapter. In fact, most mathematicians and philosophers at the time would have agreed that given the current state of mathematics, one could hardly draw any other conclusion. Geometry provided the most glaring example of the need to appeal to constructions in intuitions; but even the calculus, the most powerful branch of eighteenth-century mathematics, seemed to conform to that pattern.

In the British mathematical tradition, from which Kant appears to have learned most of what he knew about the calculus, Leibnizian infinitesimals were shunned; their role was played by rates of change. Motion – therefore, space and time – was placed at the very heart of the calculus. A variable was called a "flowing quantity" and its velocity a "fluxion." "I consider mathematical quantities," Newton wrote,

> not as consisting of very small parts, but as described by a continual motion. Lines are described, and thereby generated, not by the apposition of parts, but by the continued motion of points . . . angles by the rotation of the sides; portions of time by continued flux. . . . Fluxions are, as near as we please, as the increments of fluents generated in times, equal and as small as possible, and to speak accurately [*sic*], they are in the prime ratio of nascent increments. (*Quadrature of curves* [1704], as quoted in Kline, *Mathematical Thought from Ancient to Modern Times*, p. 363)

Thus, the derivative, the limit of an infinite sequence of ratios, was conceived as the value of the ratio at that instant of time right before the increment vanished (whatever that might mean). Limits were also characterized as dependent on temporal notions. Newton explained in *Principia:*

> Those ultimate ratios with which quantities vanish are not truly the ratios of ultimate quantities, but limits towards which the ratios of quantities decreasing without limit do always converge; and to which they approach nearer than by any given difference, but never go beyond, nor in effect attain to, till the quantities are diminished *in infinitum*. (p. 39)

The idea is to appeal to an infinite kinematic process with an end term, rather than to an infinite sequence with a limit.

Dubious theoretical considerations of this type were accompanied by no less dubious arguments in support of "claims" whose only merit was that they worked. For example, derivatives were calculated on the basis of indefensible sleights of hand, combining algebraic operations (such as division by an increment) with assumptions incompatible with algebra (division by 0). Even so, the calculus worked. It was not broken; why fix it? Those who thought of mathematics not as knowledge but as a scientifically useful technique, and those who did not think at all about such things, did not worry much about the "foundational" question. The first strident complaint came, in fact, from a philosopher.

Berkeley was the first to rise against the chaos at the foundation of the calculus. In 1734 he published a work designed to show that the boldest speculations of theologians compared favorably with the most sober statements of mathematicians on the foundations of the calculus. He noted with some glee that in England, following Newton, functions (fluents) were quantities that vary with time, and their derivatives (fluxions) "are said to be nearly as the increments of the flowing quantities, generated in the least equal particles of time; and to be accurately in the first proportion of the nascent, or in the last of the evanescent increments" (*The Analyst*, p. 66). As to the "foreign mathematicians" (i.e., Leibniz and his followers):

> Instead of flowing quantities and their fluxions, they consider the variable finite quantities as increasing or diminishing by the continual addition or subduction of infinitely small quantities. Instead of the velocities wherewith increments are generated, they consider the increments or decrements themselves . . . which are supposed to be infinitely small. (p. 67)

Leibniz's "ghosts of departed quantities" clearly made no more sense than Newton's "nascent increments." Berkeley drew his well-known conclusion:

> Nothing is easier than to devise expressions or notations, for fluxions and infinitesimals. . . . These expressions indeed are clear and distinct, and the mind finds no difficulty in conceiving them to be continued beyond any assignable bounds. But if we remove the veil and look underneath, if, laying aside the expressions, we set ourselves attentively to consider the things themselves which are supposed to be expressed or marked thereby, we shall discover much emptiness, darkness, and confusion; nay, if I mistake not, direct impossibilities and contradictions. (*The Analyst*, p. 69)

At first few mathematicians took Berkeley's complaints very seriously, endorsing the pragmatist maxim attributed to d'Alembert: Allez en avant et la fois vous viendra! But by the end of the century – by the time Kant was arguing for the inevitability of the spatiotemporal character of the calculus – the very best mathematicians had begun to worry.

In 1784, three years after the publication of Kant's first *Critique*, the Berlin Academy proposed a question on the foundations of the calculus as its mathematical prize problem. The problem was, in effect, to explain the role of the infinitely small and the infinitely large in the calculus. The proposal explained:

> It is well known that higher mathematics continually uses infinitely large and infinitely small quantities. Nevertheless, geometers, and even the ancient analysts, have carefully avoided everything which approaches the infinite; and some great modern analysts hold that the terms of the expression "infinite magnitude" contradict one another. The Academy hopes, therefore, that it can be explained how so many true theorems have been deduced from a contradictory supposition, and that a principle can be delineated which is sure, clear – in a word, truly mathematical – which can appropriately be substituted for "the infinite." (Grabiner, *The Origins of Cauchy's Rigorous Calculus*, pp. 41–2)

The first major contribution to this topic was Lagrange's *Théorie des fonctions analytiques* (1797). (The full title is almost a manifesto: "Theory of analytic functions, detached from any consideration of infinitely small or evanescent quantities, of limits or of fluxions, and reduced to the algebraic analysis of finite quantities.") There he explained that his two main purposes were to unify the calculus with algebra and, above all, to disengage the calculus from the "metaphysical" considerations involving infinitesimals and fluxions. The main problem with fluxions was their involvement with the "foreign" concept of motion and the obscurity of the associated notion of a limit.

Lagrange's main qualm about the notion of limit was that it was too vague and too geometric; as usually presented, it considered quantities "in the state in which they cease, so to speak, to be quantities"; and the ratio of two finite quantities "no longer offers a clear and precise idea to the mind, when the terms of the ratio become zero simultaneously" (see Grabiner, *The Origins of Cauchy's Rigorous Calculus*, p. 44). Lagrange's criticisms of standard foundations were decisive, and his project to eliminate infinity through a reduction to the theory of numbers (algebra) rather than through constructive processes remained a central feature of later nineteenth-century developments. But his own foundational work on the calculus aimed to avoid rather than to explicate the basic notions of limit, continuity, and the like. His proposed foundation was soon shown to be untenable.[1]

The next major step in this development was taken by Bolzano. Bolzano's mathematical work embraced an astonishing range of topics, including geometry, topology, the theory of functions, the theory of the infinite, and even the notion of the actual infinitesimal.[2] Here we shall illustrate only its philosophical thrust with a few references to Bolzano's role in initiating the research program that came to be known as the rigorization of the calculus.

It is worth pausing briefly to introduce a matter that will occupy us heavily in later chapters: the sense and purpose of foundationalist or reductionist projects such as the reduction of mathematics to arithmetic, or of arithmetic to logic. It is widely thought that the principle inspiring such reconstructive efforts was epistemological, that they were basically a search for certainty. This is a serious error. It is true, of course, that most of those engaging in these projects believed in the possibility of achieving something in the neighborhood of Cartesian certainty for those principles of logic or arithmetic on which a priori knowledge was to be based. But it would be a gross misunderstanding to see in this belief the basic aim of the enterprise. A no less important purpose was the clarification of what was being said.

The word 'rigor', normally used by mathematicians and historians to describe the purpose and the achievement of the major nineteenth-century foundational projects, is ambiguous; it is both a semantic and an epistemological notion. The search for rigor might be, and often was, a search for certainty, for an unshakable "*Grund*." But it also was a search for a clear account of the basic notions of a discipline (an "ideological" reduction; see Chapter 11). Ignorant people may think it childish to worry about the difference between 'For every epsilon, there is a delta that works for all x' and 'For any epsilon and for all x, there is a delta . . .', since anyone can see it. Such people would be well advised to study the history of the calculus and consider the difficulties that emerged from a failure to distinguish between convergence and uniform convergence. Modern critics of foundationalist projects have been blind to their clarificatory dimension, sometimes confusing a search for meaning with a search for essences ("essentialism"). The epistemological perspective on foundationalist projects makes it especially easy to miss their basic achievements, since in few cases, if any, was an actual reduction achieved to what the reductionists regarded as a fully satisfactory basis of certainty. In most of the major cases, however, there was a clear advance in the direction of reducing the more obscure to the clearer. Bolzano's work is a good example of this.

It is now widely agreed that Bolzano's first decisive contribution to the rigorization of the calculus was his "Rein analytischer Beweis des Lehrsatzes" (1817). The problem examined in this paper is one the Kantians would have declared childish: How do we know that a continuous function taking values both above and below zero must take a zero value in between? What is essential is not, of course, the specific content of the theorem, but the particular perspective from which Bolzano approached it. The question was not, What argument are we to give in order to convince ourselves that this claim is true? It was, rather, What exactly does this claim *say?* As we shall see, Bolzano's lasting contribution to this problem was his insight into the structure of the required proof sequence; but this insight was, in turn, dependent on a clear and novel picture of the content of the intermediate-value theorem. In order to reach his specific conception of that content, Bolzano first had to explicate what was meant by the central notions in the theorem, especially the notion of continuity.

Bolzano began his paper by criticizing a variety of standard proofs of the theorem and, by implication, a variety of interpretations of its content. Some proofs, he explained, depend on a truth borrowed from geometry, to wit, that every continuous line with a positive and a negative coordinate must intersect the x-axis. But this geometric proposition is, first of all, a particular case of the theorem under consideration and, more importantly, is itself in need of proof – a proof that must no doubt derive from the more general theorem. Another equally objectionable form of proof introduces the notion of continuity in terms of the notions of time and motion. The latter, however, "are just as foreign to general mathematics as the concept of space." A correct proof must begin by giving a proper definition of the basic notions involved in the theorem and must prove the claim "analytically," that is, avoiding intuition and appealing only to basic assumptions concerning numbers and functions.

The first task, therefore, was to cleanse the notion of continuity of any spatiotemporal "dynamic" character and to turn it into an "analytic," arithmetical notion. Thus, we are told that

> the expression *that a function fx varies according to the law of continuity for all values of x inside or outside certain limits* means only this: *if x is some such value, the difference f(x + w) – fx can be made smaller than any given quantity provided that w can be taken as small as we please.* (pp. 427–8)

This is, in effect, the first clear presentation of the epsilon–delta definition of continuity. Bolzano then stated the (so-called) Cauchy criterion for convergence, proving the necessity of that criterion and arguing for its sufficiency in a way marred by the unavailability of a definition of real number (which would emerge fifty years later). He then proved that a bounded set of real numbers has a least upper bound and from this finally derived the intermediate-value theorem.[3]

From the philosophical standpoint, leaving aside its specifically mathematical significance, the most interesting feature of Bolzano's proof is the careful elimination of anything that might have to do with geometry or spatiotemporal considerations. All "dynamic" notions of the calculus (continuity, limit, etc.) have been made static. The notion that a function "approaches" a value had become a misleading metaphor, which really said something about certain arithmetical inequalities that have no connection whatever with time. As a result of Bolzano's proof, the central notions of the calculus were on their way to being "arithmetized." The arithmetization – or "rigorization" – of the calculus would be completed in later years by Cauchy, Weierstrass, Cantor, and Dedekind.[4]

Bolzano saw a clear philosophical purpose behind this project. In a sketch of an autobiography, he once wrote (talking of himself in the third person):

> From very early on he dared to contradict him [Kant] directly on the theory of time and space, for he did not comprehend or grant that our synthetic a priori judgments must be mediated by intuition and, in particular, he did not believe that the intuition of time lies at the ground of the synthetic judgments of arithmetic, or that in the theorems of geometry it is allowable to rest so much on the mere claim of the visual appearance, as in the Euclidean fashion. He was all the more reluctant to grant this, since very early on he found a way to derive from concepts many geometric truths that were known before only on the basis of mere visual appearance. ("Zur Lebensbeschreibung," *Gesamtausgabe*, ser. 2A, vol. 12, pt. 1, p. 68)

There is, indeed, no theme throughout Bolzano's mathematical and philosophical work more consistent than the commitment to take pure intuition out of a priori knowledge. In the field of mathematics, this took the form of persistently excluding spatiotemporal ideas from subjects other than geometry and continually questioning the value of any kind of intuition in mathematics.[5] Already in *Beyträge zu einer begründeteren Darstellung der Mathematik* (1810), Bolzano had raised the question of the nature of mathematics and its relation to philosophy. The critical philosophy, he explained, offers one answer:

> It claims to have discovered a distinct and characteristic difference between the two fundamental types of human *a priori* knowledge, the philosophical and the mathematical, to wit, that *mathematical knowledge must be able* to represent – i.e., *construct* – adequately all of its concepts in a *pure intuition,* and thereby to *demonstrate* all of its theorems. *Philosophical knowledge,* on the other hand, lacking all intuition, must be satisfied with mere *discursive concepts.* The essence of mathematics would therefore be most properly expressed through this explanation: *It is a science of reason through the construction of concepts.* . . . As for me, I will frankly acknowledge that until now – as with in fact so many other doctrines of the critical philosophy – I have been unable to accept the correctness of Kantian assertions concerning *pure intuition* and the *construction of concepts through it.* I also still believe that surely there lies an internal contradiction in the *concept of a pure* (i.e., a priori) *intuition;* and even less can I convince myself that it is necessary to construct the concept of number in time, and that consequently the intuition of time is an essential part of arithmetic. (pp. 106–7)

Bolzano's "Rein analytischer Beweis des Lehrsatzes" was only one of a variety of contributions to this project of excluding intuition from mathematics. Since Kant had argued that all synthetic knowledge requires intuition, Bolzano's mathematical project involved an implicit challenge to the principle of synthetic judgments and, therefore, to the heart of Kantian semantics. Bolzano chose to make that challenge quite explicit.

The root of the problem

Bolzano agreed with Kant's teachers that all knowledge consists of representations; and he also agreed with Kant that representations are ultimately reducible to concepts or intuitions. But the chaos one finds in the literature both before and after Kant on the nature of these representations is characteristically replaced in Bolzano by a clear, careful statement of what they are.

To begin with, Kant's occasional ambiguities between subjective and intersubjective elements in representation are entirely eliminated. Bolzano started by distinguishing two senses of the word 'representation'. First, there are the representations that psychologists (and idealists) consider to the exclusion of all others, the mental states or determinations of the soul, as Kant had called them – such as my state of mind as I perceive a physical object. These are called "subjective representations" or "representations in us" (Bolzano, *Gesamtausgabe*, ser. 2B, vol. 18, pt. 2, p. 64). Second, there is the much more important intersubjective content of the psychological representation or, as Bolzano called it, the "representation in itself" or "objective representations."

The key to what objective representations are emerges from the fact that each meaningful grammatical unit is associated with a host of subjective representations but with only *one* objective representation, which has being even when the object of the representation does not. For example, "the subjective representations that occur in the minds of my readers when they see the word 'nothing' should be almost equal to one another, but they are nevertheless many. On the other hand, there is only one objective representation designated by this word" (Bolzano, *Theory of Science*, p. 62).

Whereas subjective representations are real, that is, "they have real existence at the time when they are present in a subject" (p. 61), objective representations are not. They are

> not to be found in the realm of the real. An objective representation does not require a subject but subsists, not indeed as something *existing*, but as a certain *something* even though no thinking being may have it; also, it is not multiplied when it is thought by one, two, three, or more beings. . . . For this reason, any word, unless it is ambiguous, designates only one objective representation. (p. 62)

Objective representations are the substance (*Stoff*) or content of subjective representations. Their being in no way depends on the existence of subjective acts, just as the meaningfulness of expressions in no way depends on anybody's bearing the appropriate meanings in mind; and like meanings, there is only one for each linguistic unit unless the given expression is ambiguous. Clearly, Bolzano's objective representations are the "meanings" or "senses" of his successors in the semantic tradition.[6] The distinction between objective and subjective representations amounts to a separation of meaning from psychological processes.

Bolzano further distinguished between an objective representation and the object of that representation. For example, the objective representation associated with the word 'table' (i.e., the meaning of 'table') should not be confused with tables, the objects of that representation. Even though both are objective, only one of them is real and only one of them is the topic of discourse when the word 'table' is involved. One must consequently distinguish three semantically relevant elements associated with a grammatical unit: (a) the objective representation or meaning, (b) the object of the representation (e.g., the entity referred to by a

proper name, *if any*), and (c) the psychological process that takes place when we perceive or think about the object of the representation (*Theory of Science*, p. 62). These distinctions had been acknowledged in one way or another by most major philosophers before Bolzano. What makes his contribution to this matter so remarkable is that he was the first one to recognize fully the enormous destructive implications of even a merely halfhearted recognition of these distinctions. As he once put it, the "proton pseudon" of the new idealistic philosophy is that "the *concept in itself* is not clearly understood, and is confused sometimes with the *thought* and sometimes with the *thing* that is its object" ("Über der Begriff des Schönen" [1843], p. 6).

One of the most important consequences Bolzano drew from these distinctions was a radical reformulation of an implicit semantic assumption that had enjoyed widespread acceptance since the days of Leibniz, the doctrine that an appropriate analysis of a subjective representation should identify in it as many parts as there are in the object represented. Leibniz had argued, for example, that the representation (*idée*) of green is indistinct because even though it appears to us to be simple, it is, in fact, complex; for physics establishes that "green emerges from the combination of blue and yellow. Thus one is justified in thinking that the idea of green is composed of those two other ideas . . . hence there are perceptions of which we are not aware" (*Nouveaux essais*, p. 100). Kant was so impressed with Leibniz's point that he used this very example in his logic lectures to explain how representations could be clear but indistinct and how distinctness could be achieved through the identification of constituents (*Wiener Logik*, p. 841). And in one of his reflections, Kant noted that a representation must be isomorphic to what it represents: "[Representation] is that determination of the spirit (*Bestimmung der Seele*) that refers to other things. What I call referring (*Beziehen*) is when its features conform to those of the external things" ("Die Vernunftlehre," Reflexionen 1676, *Kants gesammelte Schriften*, vol. 16, pp. 76–7). But, he added, a representation does not stand to what it represents as the painting to its subject. The representation

> is composed out of its component concepts in the same way in which the entire represented thing is composed out of its parts. Just as, for example, one can say that the notes of a musical piece are a representation of the harmonic connection of the tones, not because each note is similar to each tone but because the notes are connected to each other just as the tones themselves. (p. 78)

Bolzano thought that there was an important kernel of truth in all this, but that the intended point was blunted by the idealist "proton pseudon," the confusion between the objective representation and its object. Kant was assuming that "the parts of a representation are the same as the representations of the parts of its object" (Bolzano, *WL*, sec. 63). This is clearly false since, for example, the representation of a simple object may be complex (as in *the center of mass of the solar system*).[7] There is a tacit isomorphism in representation, Bolzano thought, but it is between the mental representation and its objective counterpart. And it is because this isomorphism is, for the most part, tacit or unconscious that semantic analysis is essential:

> We think a certain representation in itself, i.e. we have a corresponding mental representation, only if we think all the parts of which it consists, i.e. if we also have mental representations of these parts. But it is not necessarily the case that we are always clearly conscious of, and able to disclose, what we think. Thus it may occur that we think a complex representation in itself, and are conscious that we think it, without being conscious of the thinking of its individual parts or being able to indicate them. (*Theory of Science*, p. 69)

Semantic analysis can remedy this difficulty by bringing subjective and objective representations to isomorphic match. Bolzano's doctrine redirects conceptual analysis onto the path that will eventually lead to Frege's *Begriffsschrift*.

Finally, although the most obvious purpose of objective representations is to represent their objects, Bolzano thought that their most important task was to join together into propositions the objective content of subjective judgments. Once again, one must observe here the distinction between the subjective and the objective realms. Subjective propositions – the judgments of standard logic treatises – are mental states constituted by mental representations. Their content, the propositions in themselves (as Bolzano called them), have objective representations as their constituents. "It seems indisputable to me," he wrote, "that every, even the simplest, proposition is composed of certain parts, and it seems equally clear that these parts do not merely occur in the verbal expression as subject and predicate . . . but that they are already contained in the proposition in itself" (*Theory of Science*, p. 65). The constituents of the objective proposition expressed by the sentence *S* are, in fact, the objective representations associated with the grammatical units of *S*. Moreover, a proposition is not about its constituents but about the objects of its constituent representations (see *WL*, secs. 48–52).

The preceding sketch provides an illustration of the sense in which Bolzano is responsible for the kind of picture-theoretic semantics that would develop decades later in the writings of Frege, Russell, and Wittgenstein. Philosophical semantics was not invented for its own sake, however, but for the sake of epistemology. It was invented so that the character of knowledge, in particular a priori knowledge, could be better understood. Let us now see how Bolzano put it to that use.

Modality, analyticity, and the a priori

Few issues divide philosophers more revealingly than their attitude toward the gap between what is and what must be, toward fact and modality. The basic issue might be illustrated as follows: Even though both

(1) This man is a featherless biped

and

(2) If all men are mortal and all Greeks are men, then all Greeks are mortal

are true, they appear to differ in an elusive modal trait. As some would put it, (1) merely *is* true, whereas (2) *must* be true. Alternatively, (1) is known only through observation, whereas (2) is known a priori. Before undergoing philosophical treatment, most people would agree that a modal feeling is associated with (2) but not with (1) and that this seems to be related to the difference between the modes of access to the truth of these statements. One of the perennial philosophical issues is whether the modal feeling associated with (2) is a reliable indicator of some important trait of the associated claim or whether it is no more than the product of confusion, worthy only of being purged from our thought. Most philosophers have taken the first position; the ludicrous character of the theories of modality and the a priori that they proceeded to offer may have been the powerful fuel that moved many sane philosophers to consider the second. Bolzano was one of the most prominent nineteenth-century proponents of the "positivist" standpoint that the modal feeling is deceptive and should be explained away. And yet no one in the nineteenth century came closer than he did to an appreciation of the facts that would lead, around 1930, to a new doctrine of necessity and the a priori. In order to see this, we must first examine Bolzano's contribution to the problem of analyticity.

There are two basic ways of construing logical necessity and related modal attributes of propositions. According to the first, what we might call the Leibnizian way, to determine whether a proposition *P* is logically true, we fix *P* and change the world, watching what happens to the truth value of *P*. To see whether (1) is logically true, for example, we examine different possible worlds to see whether this man is featherless in all of them. Finding that he is not, we conclude that (1) is not logically true.

According to the second procedure, when we want to determine whether *P* is a logical truth, we do not change the world; we change *P* instead and look at whether the truth values of the resulting propositions – evaluated in this fixed world of ours – change as well. Instead of envisaging new circumstances, we envisage, in effect, new claims about the given circumstances. This idea, loosely related to Aristotle's introduction of the variable into logical considerations, was first developed in Bolzano's writings.[8]

Russell once said that it makes no sense to say of a true proposition that *it* could have been false (*Principles*, p. 12). Perhaps unable to make sense of talk concerning different worlds, he was also unable to make sense of different truth values for *this* proposition. He would have concurred with Bolzano's judgment that "every given proposition is either true or false and never changes; either it is true forever, or false forever, unless we change some *part* of it, and hence consider no longer the same but some other proposition" (*Theory of Science*, p. 194).

According to Bolzano, this tacit change of the proposition is, in fact, what is involved in most modal claims:

> We often take certain representations in a given proposition to be variable and, without being clearly aware of it, replace these variable parts by certain other representations and observe the truth values which these propositions take on. . . . Given a proposition, we could merely inquire whether it is true or false. But some very remarkable properties of propositions can be discovered if, in addition, we consider the truth values of all those propositions which can be generated from it, if we take some of its constituent representations as variable and replace them by any other representations whatever. (p. 194)

Whereas the treatment of modality in relation to possible worlds seems to leave no room for human choice, the Bolzano approach is clearly relative to a specification of the constituents that are to be regarded as variable. Which propositions will be associated with the one under consideration (as its "Bolzano companions") depends entirely on which of its constituents are considered variable. Thus, Bolzano introduced the notion of *general validity* (*Allgemeingültigkeit*): A proposition *P* is generally valid relative to the set $x_1, \ldots, x_n$ of constituents when all of its Bolzano companions (all propositions obtained by replacing $x_1, \ldots, x_n$ by arbitrary but grammatically admissible representations)[9] are true.

Bolzano briefly considered the possibility of leaving only logical concepts fixed, but observed that "the whole domain of concepts belonging to logic is not circumscribed to the extent that controversies could not arise at times" (pp. 198–9), and he therefore let the matter drop. To the extent that the distinction can be drawn, he proposed to call *logically analytic* all those propositions that are universally valid relative to all their nonlogical concepts. He reserved the term 'analytic' for the much less promising notion of a proposition that is universally valid relative to some constituent or other (pp. 197–8).

Whereas Bolzano's notion of analyticity does not appear to capture an interesting concept, the same could hardly be said about his general validity.[10] Much more important than singling out those of his notions destined to have a brilliant future, however, is recognizing the basic

insight underlying Bolzano's approach to the subject of analyticity. This insight emerges most clearly when, after putting forth his own proposals, Bolzano (as usual) turns to examine the major alternatives available in the philosophical literature. After a detailed examination of the flaws in the Kantian notion of analyticity he concludes:

> Generally, it seems to me that none of these explications sufficiently emphasizes what makes these [analytic] propositions important. I believe that this importance lies in the fact that their truth or falsity does not depend upon the individual representations of which they are composed. . . . This is the reason why I gave the above definition. (p. 201)

It would be hard to exaggerate the significance of this insight or the extent to which it undermines the basis of Kant's philosophy; for Bolzano is saying, in effect, that Kant's contention that truths based only on conceptual knowledge must be analytic is very nearly the opposite of the truth, since the basic feature of analytic knowledge, in Kant's nominal sense, is that it *ignores* most concepts or representations. Bolzano's point is precisely the one spelled out at the end of Chapter 1. That he was the first to make it (and the only one to see it clearly for quite a number of decades) was a consequence of the fact that he was the first to distinguish meticulously between the content of a conceptual representation and its psychological trappings.

Bolzano's insight led to an improved understanding of the notion of analyticity; but more significantly, it led also to the following question: If the conceptual resources that must be mobilized in order to justify analytic judgments are only a modest fragment of those available to us, what job do the remaining concepts do? Surely they must do some work! It seems absurd to suppose that the modest stock of concepts that pertain to logic justify some claims (the analytic ones), but that all remaining concepts have no comparable talents. If conjunction and implication suffice to establish the truth of 'If this is A and B, then it is A', then color concepts, for example, should be capable of contributing to the justification of some other claims. Which ones? Whatever the answer, we can see that Bolzano's careful semantic investigations show Kant's position on this matter to be an utterly untenable middle ground; for Kant insisted on the capacity of concepts to establish the validity of certain claims, but at the same time ignored the remaining vast continent of conceptual resources. Kant's semantic confusions led him to ignore the grounding force of descriptive concepts. That in turn led him to postulate pure intuition. When the confusions were exposed, it opened once again the question of whether arithmetical and geometric knowledge require something beyond the realm of concepts for their justification.

Considerations such as these likely played a role in Bolzano's later views on the nature of synthetic a priori judgments. As we have seen, he readily granted the existence of such judgments (in the nominal sense of synthetic):

Not everything that can be predicated of an object, even with necessity, lies already in the concept of that object. For example, one can predicate of every rectilinear triangle that the sum of its three angles = 180 . . . nevertheless, no one will believe that these properties of the triangle are contained as constituents of this concept. ("Logische Vorbegriffe," *Gesamtausgabe,* ser. 2A, vol. 5, p. 178)[11]

Kant's solution, involving the principle of synthetic judgments, was flatly rejected:

Kant poses the question, "What justifies our understanding in assigning to a subject a predicate that is by no means contained in the concept (or explanation) of the former?" – And he thought he had discovered that this justification could only be an intuition that we link with the concept of the subject and that also contains the predicate. Thus, for all concepts from which we can construct synthetic judgments, there must be corresponding intuitions. If these intuitions were always merely empirical, the judgments which they mediate should also always be empirical. Since, nonetheless, there are synthetic *a priori* judgments – (as such things are undeniably contained in mathematics and pure natural science); there must also be *a priori* intuitions – however odd this might sound. And once one has decided that there can be such, one will also convince oneself easily that for the purposes of mathematics and pure natural science, time and space are these intuitions. (*Beyträge zu einer begründeteren Darstellung der Mathematik,* pp. 234–5)[12]

What, then, is it that justifies the belief in these a priori judgments? Bolzano's most common explanations were of the empiricist type, thus inconsistent with the alleged a priori status of the relevant judgments. But in the *Wissenschaftslehre,* he finally came to recognize that his semantic insights could be put to good use at this very point:

What justifies the understanding to attribute to a subject *A* a predicate *B* which does not lie in the concept *A*? Nothing I say but that the understanding *has* and *knows* the two concepts *A* and *B.* I think that we must be in a position to judge about certain concepts merely because we have them. . . . Since this holds generally, it also holds in the case when these concepts are simple. But in this case, the judgments which we make about them are certainly synthetic [in Kant's nominal sense]. (*Theory of Science,* p. 347)

This says, in effect, that not only the *content* but also the justification of synthetic a priori judgments is purely conceptual.

This was no more than a flash of insight, however, and was not destined to play a major role in Bolzano's system. Bolzano's official account of how a priori knowledge is grounded was very different from the one we just saw in *Wissenschaftslehre.*

The basis of logical truth

Those willing to make sense of the modal difference between (1) and (2) will typically say that the truth of (1) both consists of a certain correspondence with facts and can be determined by an appeal to that correspondence, but will be willing to grant only that the truth of (2) may consist in that correspondence. It would be just plain muddleheaded to appeal to that correspondence in order to justify (2) – as muddleheaded as an attempt to determine empirically whether all bachelors are unmarried. The modalist will readily grant that the truth of

(3) All men are featherless bipeds

is entirely reducible to that of all of its instances, in the sense that there is nothing more to (3) than the conjunction of its instances such as (1), and these are merely factual. Hence, in cases like (3), the modalist will agree with Bolzano that an examination of a host of other propositions is essential to determine the truth value of (3), since (3) is, in the end, no more than the conjunction of those propositions. But for the modalist, the ground of (2) – and of necessary claims in general – does not emerge from below, from the facts, but from above. The modalist may grant that there is a fact that makes (2) true and that innumerable other facts make the Bolzano companions of (2) true as well. But these facts are, for the modalist, irrelevant to the justification of (2); something prior to and independent of these facts determines and explains the truth of (2) and, at the same time, its peculiar modal character. Traditionally, the "form" of the proposition was cited as the reason for its truth: (2) is true not in virtue of facts – which are, indeed, as (2) says they are – but in virtue of its form.

In his characteristic cool-headed manner, Bolzano noted the overwhelming weight of the tradition in favor of this approach and then examined the various attempts to explain the form–matter distinction. He concluded, quite rightly, that there is little but confusion behind the traditional ways of drawing the distinction. For those who would insist on using some notion of form, he offered an honest definition: The form of the proposition *P* relative to its constituents $x_1, \ldots, x_n$ is, in effect, the class of propositions that differ from *P* at most in the constituents in question (see *WL,* sec. 186). But Bolzano was clearly not very fond of those logicians and philosophers "blinded by the erudite twilight of the words 'form' and 'matter'" (*Theory of Science*, p. 164).

Bolzano concluded that the idea of form as traditionally construed was worthless, and he could see no other candidate for the role of a suprafactual ground of logical truth. Hence, he saw no reason to place (2) in a different category than (1); one might as well call them both analytic. In particular, the truth grounds of (2) are essentially of the same sort as those of (1):

> The only reason why we are so certain that the rules *barbara*, *celarent*, etc. are valid, is because they have been confirmed in thousands of arguments in which we have applied them. This is also the true reason why we are so confident, in mathematics, that factors in different order give the same product, or that the sum of the angles in a triangle is equal to two right angles. (*Theory of Science*, p. 354)

The ground of (2), like that of (1), derives from below, from the facts.

The same attitude is clearly displayed in Bolzano's interpretation of Aristotle's celebrated definition of a syllogism as "a discourse in which, certain things being stated, something other than what is stated follows of necessity from their being so" (*Prior Analytics*, 24b19).

Here is Bolzano's comment:

> The 'follows of necessity' [in Aristotle's characterization] can hardly be interpreted in any other way than this: that the conclusion becomes true *whenever* the premises are true. Now it is obvious that we cannot say of one and the same class of propositions that one of them becomes true *whenever* the others are true, unless we envisage some of their parts as variable. . . . The desired formulation was this: as soon as the exchange of certain representations makes the premisses true, the conclusion must [*sic*] also become true. (*Theory of Science*, p. 220)[13]

In general, Bolzano's interpretation of statements of valid inference such as (2) is this: "Everybody feels that the sense of the assertion can only be to say that in each case where a substitution of representations makes the antecedents true, the consequent will also express a truth" (p. 253). For him the only way we can make sense of the idea of a necessary link between the premise and the conclusion of a valid inference is by assuming that some of the constituents of (2) are tacitly taken as variable and that we are required to examine the truth values of all of the appropriate instances. The basis of the necessity of (2) is the just plain truth of the appropriate instances. This leaves us with a strange category of analytic statements that includes not only (1) and (2), but also

(4) If this is red, then it is not-blue.

The modalist would want to regard (4) as necessary, but certainly not in virtue of its form. For him, the inference from 'This is red' to 'This is not-blue' expresses a syllogism in the exact sense of Aristotle's characterization, and form surely plays no role in the necessity involved in this inference. The idea that our acceptance of (4) should conform to the inductive strategy of checking on the truth values of antecedents and consequents for its instances is too ludicrous to even be taken seriously. Bolzano might have appealed at this point to his doctrine of the conceptual ground of synthetic a priori knowledge. But if so, why not be equally generous in the case of (logically) analytic knowledge? Having divided his natural kinds in the wrong places, Bolzano was not properly disposed to ask the right questions. It would take almost a century to reach once again the level that Bolzano was approaching at this point, to rearrange those categories, and to formulate the right questions.

Decades later others could see in Bolzano's work a nearly complete version of a successful defense of a necessitarian standpoint. If concepts can provide a justification for synthetic knowledge and if logical concepts are the only ones that are kept fixed in the case of logically analytic knowledge, why not say that such knowledge is grounded on logical concepts? Or, to put it in modern terms, why not say that logical truth is truth in virtue of the meanings of the logical terms? Why not say that the analyticity of (2) is based not on the fact that it and other sentences have certain truth values, but rather on the fact that some of the constituent words have certain meanings? By placing (1) and (2) in the same category of analytic judgments, Bolzano made it harder to see that in some cases – such as (2) – the understanding of what is being said is not only a necessary, but also a sufficient justification of logical knowledge.

Bolzano's indecisiveness on this epistemological matter is one of a variety of indicators of the fact that, great as his contribution was, the semantic tradition still had a long way to go. Bolzano excelled when his subject was the content of a priori statements, where he argued with unsurpassed technical and philosophical authority that claims widely thought to involve intuition in their content, in fact, did not. But he was far less successful when he turned to the justification of these avowedly a priori claims. There his views were more conservative, and he tended to infer from his justified rejection of standard aprioristic accounts that nothing was left but a form of positivism. From *Beyträge zu einer begründeteren Darstellung der Mathematik* to the *Wissenschaftslehre*, he kept repeating that the reason we are so confident of mathematical laws such as that of the commutativity of multiplication is that "they have been confirmed in thousands of arguments in which we have applied them" (*Theory of Science*, p. 354).

Bolzano's sketch of a picture-theoretic semantics was also no more than a sketch. The central ideal of logical analysis, the realization that language is an extraordinarily misleading guide to content, was still in the future.[14] For Bolzano language was a rather reliable picture of the form of objective propositions. He wrote as if, by and large, German sentences were isomorphic maps of the corresponding objective propositions. Thus, the objective proposition expressed by 'This triangle is large' consists of the representations *this, triangle, has,* and *largeness.* The isomorphism obtains even in the case of names: '35' and '53' were said to express representations whose constituents are identical (presumably, the representations *3* and *5,* whatever they may be) and to differ only in "the way in which these parts are connected" (*Theory of Science,* p. 69).[15] Finally, even the notion of content was not given its full role. Bolzano was still attached to the long-standing tradition that thinks of deductive relations as somehow analogous to causal connections and seeks to draw within the class of valid logical links a further distinction aimed at identifying the "proper ground" of certain claims. These and other matters would finally be settled a few decades later, in Frege's writings. Before we turn to them, however, we must consider what happened to Kant's views on geometry in the meantime.

3

Geometry, pure intuition, and the a priori

In philosophy, an intuition can only be an example; in mathematics, on the other hand, an intuition is the essential thing.

Kant, *Logik Busolt*

For Helmholtz, however, there existed the option: either "necessity of thought" or "empirical origin." But it is appropriate to add to these: necessity of intuition, and this as pure.

Cohen, *Kants Theorie der Erfahrung*

So it is entirely implausible that outside the range of pure mathematics we will ever make use of these hypotheses of non-Euclidean spaces.

Riehl, *Phil. Krit.,* vol. 2

From the beginning of the nineteenth century, Kant's pure intuition had a rough time in analysis. The rigorization of the calculus banished intuition from the notions of function, continuity, limit, infinitesimal, and all else that had elicited Berkeley's justified complaint. The arithmetization of analysis cornered the pure intuition of time into arithmetic, where Frege would soon deal it a death blow (Chapter 4). Mathematics was not just the theory of abstract magnitudes, numbers, functions, and infinitesimals, however. It was also the science of space, geometry, and here Kantians could rest assured that intuition would never be dethroned. Or so it seemed for a while.

During the nineteenth century, geometry was the battleground of two major epistemological wars. The first, the subject of this chapter, concerned the role of pure intuition in knowledge; the second, surveyed in Chapter 7, took it for granted that that role was nil and questioned the nature of geometric concepts. It is interesting that in both cases the semantic tradition had nothing to contribute to these developments. As we shall see, a proper picture of geometry demanded a synthesis of Kantian and semantic insights that neither of these two conflicting traditions was in a position to undertake. For the time being, however, our topic is the nature and role of pure intuition in geometric knowledge.

The presence of pure intuition in geometry, Kant thought, is manifested by a peculiar type of necessity that attaches to geometric judgments. We begin by examining this idiosyncratic modality.

One of the central distinctions in Kant's theory of modality was between a kind of necessity derived from intuition (*Anschauungsnothwendigkeit*) and another derived from thought (*Denknothwendigkeit*). The former has its source in features of the human sensibility, the latter in features of the understanding.[1] Perhaps the best short explanation of this distinction occurs in a passage in Frege's *Grundlagen*, in which he is trying to explain why he thinks that arithmetic is a part of logic. "Empirical propositions," he writes,

> hold good of what is physically or psychologically actual, the truths of geometry govern all that is spatially intuitable, whether actual or product of our fancy. The wildest visions of delirium, the boldest inventions of legend and poetry, where animals speak and stars stand still, where men are turned to stone and trees turn into men, where the drowning haul themselves up out of swamps by their own topknots – all these remain, so long as they remain intuitable, still subject to the axioms of geometry. Conceptual thought alone can after a fashion shake off this yoke, when it assumes, say, a space of four dimensions or positive curvature. To study such conceptions is not useless by any means; but it is to leave the ground of intuition entirely behind. . . . For purposes of conceptual thought we can always assume the contrary of some one or other of the geometrical axioms, without involving ourselves in any self-contradictions. . . . The fact that this is possible shows that the axioms of geometry are independent of one another and of the primitive laws of logic, and consequently are synthetic. Can the same be said of the fundamental propositions of the science of number? Here, we have only to try denying any one of them, and complete confusion ensues. Even to think at all seems no longer possible. (*The Foundations of Arithmetic*, pp. 20–1)

Frege is saying that the laws of geometry and arithmetic, unlike those of physics, are necessary. But whereas geometric propositions are necessities of intuition, the laws of arithmetic are necessary in a far deeper sense: Thought itself becomes impossible if we deny them. To the extent that logic is the pure theory of concepts, arithmetic must be a part of logic. The doctrine is, of course, not Kantian, but the ideological framework certainly is.

What does it mean to say that geometric laws are necessities of intuition? Kant's writings contain no more than a few confusing hints. Of course, he had no reason to be overly concerned about working out the details: Who would seriously doubt circa 1800 that geometry was necessary or that its necessity had something to do with geometric constructions?[2] But the situation changed shortly after Kant's death, when non-Euclidean geometry made its first public appearance. By the second half of the nineteenth century, people began to wonder first about the exclusive necessity of Euclidean geometry and then about the role of intuition in *any* geometry, Euclidean or otherwise. As neo-Kantians were forced to address this issue more extensively, it slowly emerged that the master's silence was not a sign of unspoken wisdom.

Revealingly, the cleverest among neo-Kantians silently pushed Kant's pure intuition to a corner of their doctrine of geometry; what they offered as the *truly* Kantian theory of geometry suspiciously resembled one of Helmholtz's central contributions to that field. By the end of the century, the neo-Kantians' writings on this topic had become an unwitting testimony to the fact that geometry called for a ground entirely different from the one Kant had envisaged. In conjunction with the rigorization of the calculus and with what Frege would soon be doing to arithmetic, these episodes converged to establish what Bolzano had claimed in 1810: that Kant's pure intuition plays no role whatever in mathematics.

Kant's mixed message

Kant thought that geometry was a good example of how little you could do in science with mere concepts. If you try to prove a geometric theorem from pure concepts, "All your labor is in vain; and you find that you are constrained to have recourse to intuition, as is always done in geometry. You therefore give yourself an object in intuition . . . [indeed] an object *a priori* in intuition, and ground upon this your synthetic proposition" (*Critique*, A47–8/B65). Well, exactly how do you give yourself an object a priori in intuition, and does the geometer really *need* to do that? There are three ways to read Kant's views on pure intuition, what we might call Platonist, constructivist, and structuralist. Let us consider them in sequence.

Every now and then one finds in Kant statements that suggest that pure intuition differs from empirical intuition in that the objects it represents are "pure" rather than empirical (see "Vorarbeiten zu Ausgleichung eines auf Missverstand beruhenden mathematischen Streits," *Kants gesammelte Schriften*, vol. 23, p. 201). In the *Critique* Kant gave a transcendental twist to the primary–secondary quality distinction when he argued that "qualities cannot be presented in any intuition that is not empirical," but quantities can: "The shape of a cone we can form for ourselves in intuition, unassisted by any experience, according to its concept alone, but the color of this cone must be previously given in some experience or other" (*Critique*, A715/B743). This seems to present the geometric cone as an object of a *different kind* than the objects given in empirical intuition, rather than as the form of objects given in empirical intuition. Here the difference between form and content (or matter) appears to correspond to the difference between innate and acquired, as if the colorless image of a cone could be formed by someone who had no previous experience, and this image would be a pure intuition. Whewell seems to have interpreted Kant along these lines when he thought of the Kantian intuition as an "imaginary looking" (*History of Scientific Ideas*, vol. I, p. 140; cited by Mill, *Logic*, bk. 2, chap. 5, sec. 5); and so did Riehl, albeit disapprovingly, when he saw in pure intuition an echo of Platonic forms (*Phil. Krit.*, vol. 2, p. 104). Consider also a typical reference to the "construction" of a geometric concept involved in the process of a proof: We construct it, Kant said, "by representing the object that corresponds to this concept either by imagination alone, in pure intuition, or in accordance therewith also on paper, in empirical intuition" (*Critique*, A713/B741). The explicitly drawn duality encourages us to think of pure intuition as given in a domain involving imagination and of empirical intuition as belonging to an entirely different domain.

Though these and other passages invite a Platonist construal, this interpretation is almost certainly at odds with Kant's intentions. There is a second interpretation that, like the first one, associates pure intuition with the givenness of objects (as one would expect of any sort of Kantian intuition) but takes these objects to be empirical. This interpretation focuses on Kant's remarks about the construction of mathematical concepts.

Few elements of the critical philosophy are better known than Kant's attempt to bring together what he had cast asunder in his distinction between sensibility and understanding. Detached as those two faculties may be, there can be no human knowledge unless they join forces: Concepts without intuition are empty, and intuition without concepts is blind:[3]

> If our knowledge claims are to have objective reality, i.e., if they are to refer to an object and thereby have meaning and sense, it must be possible that the object be given in some manner. Otherwise the concepts are empty, and even though one indeed has thought with them, through this thinking nothing has really been known; one has merely played with representations. (*Kritik*, A155/B194–5)

Mathematical concepts are linked to intuition by the celebrated construction of the concept. Kant and his followers used this phrase repeatedly in allegedly explanatory contexts in which one can almost see them frowning down on their readers, daring them to exhibit their stupidity by asking what it means. The truth is that neither Kant nor his followers had any very definite idea of what that "construction" was. The plausibility of any Kantian thesis involving this notion is inversely proportional to the clarity with which it is explained. It is interesting that whenever Kant did make an effort to illustrate what he meant by

"construction in intuition," an empirical intuition entered in. For example, Kant explained that the construction of a figure makes it "present to the senses" (*Critique*, A240/B299). When we prove a proposition about triangles we may construct the concept "on paper, in empirical intuition. . . . The single figure which we draw is empirical, and yet it serves to express the concept, without impairing its universality" (*Critique*, A713–14/B741–2). Moreover, "we cannot think a line without *drawing* it in thought, or a circle without *describing* it. We cannot represent the three dimensions of space save by *setting* three lines at right angles to one another from the same point"[4] (*Critique*, B154). The same is true of arithmetic: In order to produce the synthesis required for the proof that 7 + 5 = 12, "we call in the aid of the intuition that corresponds to one of [those concepts], our five fingers, for instance, or, as Segner does in his Arithmetic, five points" (*Kritik*, B15;[5] Bolzano comments on these remarks in *WL*, sec. 305).

The structuralist interpretation differs from the first two in treating pure intuition as something quite unlike intuition, which is a singular representation. According to the structuralist Kant, what is pure and a priori is not a kind of object but a mode of knowledge of empirical objects. All objects of intuition are empirical, and pure intuition is the "mere form" of empirical intuition (*Critique*, A239/B298). It follows that pure intuition is not a sort of singular representation but a formal feature of such representations, a *lex menti insita*, as Kant once put it. When in this structuralist gear, Kant would explain that when we construct the concept of a triangle, for example, we do not really construct an instance of that concept or even give any particular object to intuition, but what we construct is only the form of an object. Indeed, given this construction the possibility of its object might still be doubtful (*Critique*, A223/B271; see also A239/B298).

However one interprets the nature of pure intuition, there are two related, but distinguishable problems faced by Kant's account of geometry: How does pure intuition support the necessity of Euclidean geometry, and why must a geometric argument be a chain of inference guided throughout by intuition (*Critique*, A717)?[6]

As we saw earlier, Kant explained that when we prove a proposition about triangles, we may construct that concept "on paper, in empirical intuition"; and he added that "the single figure that we draw is empirical, and yet it serves to express the concept without impairing its universality" (*Critique*, B741). One might have thought that what threatens the universality of the procedure is not the *empirical* character of the figure involved but the fact that it is a *singular*, specific object and that *it* is all that has been considered. Be that as it may, Kant added, "For in this empirical intuition we consider only the act whereby we construct the concept, and abstract from the many determinations (for instance, the magnitude of the sides and of the angles), which are quite indifferent, as not altering the concept 'triangle'" (*Critique,* A714/B742). Note that Kant wanted to abstract not only from those determinations that fix parameters that the concept leaves undetermined (such as those Kant enumerated in the preceding passage in parentheses) but also those in which the empirical object of our empirical intuition fails to qualify as an instance of the constructed concept (e.g., the three-dimensionality of the constructed triangle, the wiggly nature of its lines, etc.). Determinations of the first sort yield instances of the concept, while the second result in objects that are, at best (and in a sense badly in need of elucidation), mere approximations of instances of the given concepts. If, *per impossibile,* we were somehow given all instances of a concept in intuition, we could abstract from their idiosyncrasies by simply considering only what is true of all of them and thus achieve the intended result. But no object ever given to us in *any* kind of intuition is, for example, an instance of the concept of a triangle.

How do we decide which determinations are to be abstracted, which features of the constructed figure are relevant to a proof? Kant was, understandably, not terribly exercised by this question. His complete answer, such as it was, is encapsulated in that celebrated aphorism to the effect that the geometer "must not ascribe to the figure anything save what necessarily follows from what he has himself set into it in accordance with his concept [of it]" (*Critique,* Bxii). When pursued to its logical conclusion, however, this remark leads to an uncomfortable dilemma; for what necessarily follows from what the geometer has set in the figure either (a) follows from his concept of that figure, quite independently of *any* features of the figure ("formal" or otherwise), or (b) follows only when we examine in addition to the concept itself some *relevant* features of the figure. In the first case we have the position Russell came to endorse around 1900: The synthesis in mathematical and logical knowledge can be produced from concepts alone, without appeal to any kind of intuition.[7] This clearly conflicts with the principle of synthetic judgments and the associated link between mathematics and intuition. In the second case – probably Kant's own choice – we are left with the original question: Which of the various features exhibited by the empirically constructed figure (whether in the mind or on paper) are allowable grounds of inference? It would appear that by Kant's own standards, the only guides in this decision are the axioms and theorems of geometry. But *before* we can use the intuitional X to provide a ground for the synthesis expressed in the axioms, we must have those very axioms in order to determine what X is. Thus, Kant's prescription for identifying the features to be abstracted leads beyond Kantianism to the

view that we cannot synthesize the axioms until we have them. In the Kantian idiolect, geometric axioms would have a regulative role and would therefore not pertain either to the domain of sensibility (intuition) or to the understanding (construed so as to exclude Reason). We shall soon see how some neo-Kantians chose to pursue this un-Kantian course in order to construe Kant as having anticipated Helmholtz's insights.

Helmholtz's challenge

The claim that Euclidean geometry is a necessity of intuition had been disputed by empiricists on the familiar grounds that what we cannot imagine may well exist. Mill, for example, had distinguished between the sense in which antipodes are inconceivable and the sense in which the enclosure of space by two straight lines is inconceivable. In the former case everyone can surely represent the circumstances under consideration even if they appear incredible. In the latter, however, "we cannot represent to ourselves" the alleged circumstances:

> We cannot represent to ourselves two and two as making five; nor two straight lines as enclosing space. We cannot represent to ourselves a round square; nor a body all black, and at the same time all white. These things are literally inconceivable to us, our minds and our experience being what they are. (*Hamilton's Philosophy,* pp. 69–70)

But even this strong sense of inconceivability is consistent with the possibility and even the truth of inconceivable claims; for even though we cannot represent round squares, things black and white all over, and so on, we can represent circumstances in which we could represent them.[8] The inconceivability arises only because our experience has taught us to associate or dissociate two representations:

> We should probably be as well able to conceive a round square as a hard square, or a heavy square, if it were not that, in our uniform experience, at the instant when a thing begins to be round it ceases to be square. . . . Thus our inability to form a conception always arises from our being compelled to form another contradictory to it. . . . We cannot conceive two and two as five, because an inseparable association compels us to conceive it as four; and it cannot be conceived as both, because four and five, like round and square, are so related in our experience, that each is associated with the cessation, or removal, of the other. . . . And we should probably have no difficulty in putting together the two ideas supposed to be incompatible, if our experience had not first inseparably associated one of them with the contradictory of the other. (*Hamilton's Philosophy,* pp. 70–1)

Mill illustrated our ability to represent the inconceivable with examples from several "a priori" disciplines. In arithmetic, for example, our commitment to the law that 2 + 2 = 4 would vanish if whenever two pairs of things "are either placed in proximity or are contemplated together, a fifth thing is immediately created and brought within the contemplation of the mind engaged in putting two and two together" (p. 71). The production of this fifth thing must be "instantaneous in the very act of seeing, [so] that we never should see the four things by themselves as four: the fifth thing would be inseparably involved in the act of perception by which we should ascertain the sum of the two pairs" (p. 73). Clearly, Mill was thinking about adding up things like rabbits or cows, not things like solutions of third-degree equations or Roman consuls. As Frege would point out in *Grundlagen* (1884, secs. 7 and 8), the latter are not easily "placed in proximity" or involved in "acts of perception." A world in which when someone adds the first two Roman consuls to the next two a fifth one appears, presumably with his distinct proper name, his own political record, and so on, is not a world at all but the product of a confused mind; for in that world the decision to add would alter the past, and on pain of contradiction there could not be one person adding a group of objects and another not.

Mill's arguments against the a priori character of geometry were no better. For example, he quoted approvingly James Fitzjames Stephen's remark that a "world in which every object was round, with the single exception of a straight inaccessible railway, would be a world in which everyone would believe that two straight lines enclosed a space" and then observed: "In such a world, therefore, the impossibility of conceiving that two straight lines can enclose a space would not exist" (*Hamilton's Philosophy*, p. 72).[9] If Mill was the wisest positivist, as he probably was, the Kantians had little to fear from the positivist challenge to their doctrine.

The first decisive step in the overthrow of the notion of a necessity of intuition did not come from positivism. Widespread opinion notwithstanding, it did not come from the discovery of hyperbolic geometry either, or even from the recognition of its consistency. Ironically, it emerged from an attempt to show that the new geometries were not a challenge to Euclid.

In 1868 Beltrami published a paper entitled "An Attempt to Interpret Non-Euclidean Geometry," in which he introduced his celebrated pseudospherical model. Had the interpretation offered in that paper been successful, it would have established the consistency of hyperbolic geometry.[10] Despite appearances, the doctrines put forth could have brought nothing but comfort to Kantian souls. For Beltrami's ultimate goal was not so much to interpret as to *reduce* hyperbolic geometry to Euclidean geometry and to argue that there was no more geometric sense in the former than it could derive from the latter.

Beltrami's stated purpose was "to find a real substratum" for Lobatchewski's geometry, but only for its two-dimensional fragment (*Opere matematiche*, p. 375). Beltrami concluded that the hyperbolic plane is, in fact, the Euclidean pseudosphere in disguise, since the Euclidean metric of the Euclidean pseudosphere coincides (locally anyway) with that of Lobachevski's plane. He argued that no analogous interpretation could be given for three-dimensional hyperbolic space, however, since he thought that the part of space in which the interpretation is constructed must have a metric not reducible to the standard Euclidean form $dx^2 + dy^2 + dz^2$:

> Since until now the notion of a space different from [Euclid's] appears to be lacking or transcends, at least, the domain of ordinary geometry, it is reasonable to suppose that, even though the analytical considerations on which the preceding constructions rest can be extended from the field of two variables to that of three, the results obtained in this latter case can not yet be constructed with ordinary geometry. (p. 397)

And in his next study on the topic he would insist that his two-dimensional model gives

> a true and proper interpretation, since one can *construct* [the appropriate concepts] on a *real* surface; on the other hand, those which embrace three dimensions can be represented only analytically, since the space in which such a representation would materialize is different from the one that we generally call by that name. ("Teoria fondamentale degli spazi di curvatura costante" [1868–9], *Opere matematiche*, p. 427)

Far from posing any threats to Kant's philosophy, Beltrami's work was consistent with and possibly even grounded upon it. Kant had never doubted the logical consistency of non-Euclidean geometries. He would surely have said of hyperbolic geometry that it is impossible but not *logically* impossible (since its "negation," Euclidean geometry, is not logically necessary but only intuitionally necessary). So the fact that there is an interpretation of hyperbolic geometry is hardly surprising, nor is it surprising that this interpretation has to be given in terms of Euclidean intuitable notions. Nor is it surprising that wherever that reduction to Euclidean intuitions fails, we must abandon the project of giving an interpretation to Lobachevski's theory. One could hardly find a more appealing package of good news for Kantians in a geometric monograph. Yet less than three years later, Helmholtz would see in Beltrami's study a refutation of the Kantian notion of the intuitional necessity of Euclidean geometry. With characteristic boldness, Helmholtz recognized the potential of Beltrami's analytic representation. He was, in a sense, the first to realize that what is now called the Beltrami–Klein model of hyperbolic geometry is, indeed, a model of hyperbolic geometry. Let us first look briefly at this model.

To facilitate his metric analysis of the pseudosphere, Beltrami introduced an auxiliary surface, the interior of a Euclidean circle. An isomorphic mapping for the pseudosphere will induce a hyperbolic metric on this circle. The intrinsic metric of the pseudosphere is determined by associating with any two points P and Q on it the Euclidean length $d(P, Q)$ of the geodesic that links them along the pseudospherical surface. This metric function can be expressed as a function $f(X, Y)$ of the points X and Y that are the projections of P and Q, respectively, on the auxiliary circle. One may now choose to abandon Beltrami and look at $d(X, Y)$ not as a device for calculating the intrinsic Euclidean distance between P and Q (along the pseudosphere), but as giving the "distance" between X and Y. Thus construed, the function f defines a metric on the open surface inside the auxiliary circle that is hyperbolic, since it is the image of a hyperbolic metric under an isomorphism. By the new metric standard, the chords of the circle are infinitely long straight lines. Angles are correspondingly remetrized. Even though the open surface is a model of hyperbolic geometry, Beltrami did not think for a moment that this open surface could qualify as an "interpretation" (in his sense) of hyperbolic geometry. If it could, then by his own arguments three-dimensional hyperbolic geometry would also be interpretable. No doubt, the "arbitrary" (i.e., non-Euclidean) character of the metric defined by f was the decisive reason for ruling out the "auxiliary" space as a possible interpretation.

It was Helmholtz who observed that the "straight lines" in the above open surface are far closer relatives of standard straight lines than those found in Beltrami's preferred model. This was the basis of his well-known proof that we can intuitively represent non-Euclidean spaces, thus showing that Euclidean geometry is not a necessity of intuition.

The first step in Helmholtz's argument was to remove the ambiguity from Kant's notion of an intuitive representation:

> By the much misused expression 'to represent' or 'to be able to think of how something happens', I understand that one could depict the series of sense-impressions one would have if such a thing happened in a particular case. I do not see how one could understand anything else by it without abandoning the whole sense of the expression. ("On the Origin and Significance of the Axioms of Geometry" [1870], *Epistemological Writings*, p. 5)

Helmholtz used this analysis of representation to show that non-Euclidean geometry is representable. He prefaced his argument with two "warm-up" stories designed to dislodge our faith in the truthfulness of intuition. The first was a Flatland case, in which two-dimensional beings living on a curved surface develop a non-Euclidean geometry on the basis of their perceptions. Since the idea that such beings might have anything like our perceptions is well-nigh incoherent, the philosophical point of this popular example is virtually nil. The second story overcame this difficulty by presenting a three-dimensional world, what we shall call a 'mirror universe'.

Imagine a spherical mirror on whose surface S all events in our Euclidean space are reflected. Now imagine a three-dimensional world delimited by S and a plane through the focal point of the spherical mirror. In this world, physical objects behave exactly the way they "appear" to behave in the mirror. Thus, for every object in our Euclidean space, there will be a corresponding object in the mirror universe. When an object O in our space moves away from S to infinity, the corresponding mirror object, O^*, will move away from S toward its focal point; O does not change shape as it moves, but O^* does, shrinking (by our metric standards) as it moves away from S.

How do we determine that the geometry of our space is Euclidean? We might, for example, draw a right triangle and measure its three sides; we observe that the measurements are 3, 4, and 5 units, respectively, thus confirming the Pythagorean theorem that separates Euclidean geometry from its constant-curvature rivals. But as I perform these measurements, a little man, looking and moving just like I do, but changing his shape as he moves, measures the sides of a triangle that looks very unrectangular (to us). Yet since his "meterstick" also changes its length as it moves, he also finds that it fits three, four, and five times, respectively, on the sides of the triangle. Unaware of the "fact" that his meterstick is changing size as he moves, the poor little man infers – at the same time *we* do – that his space must be Euclidean. In general, whenever a geometric statement concerning an object in our Euclidean universe is true (relative to our metric standards), the same statement will be true of the *corresponding* mirror object (relative to the metric standards in the mirror world). It follows that from the viewpoint of the mirror universe, the surface is also convex – rather than concave, as one might at first think, for all objects on S (and therefore S itself) are their own corresponding objects. In spite of the striking difference between the two universes, precisely the same geometry is valid in both; indeed, both are Euclidean. And the symmetry goes even further. From the standpoint of the mirror metric, we are the inhabitants of a mirror universe in which objects change shape as they move, shrinking as they move closer to their focal point. Whether things look "funny" or not is entirely a matter of perspective. Helmholtz did to intuition in geometry what Montesquieu had done a century and a half earlier to intuition in political philosophy.

Having softened our faith in intuition, Helmholtz delivered the decisive blow. Once again he set out to describe a three-dimensional universe that we can represent intuitively; but this time the geometry of the universe would be non-Euclidean. It is here that Helmholtz appealed to Beltrami's auxiliary sphere. Using Beltrami's results, Helmholtz set out to "deduce how the objects of a pseudospherical world would appear to an observer, whose visual estimation and spatial experiences had, exactly like ours, been developed in flat space" (p. 21). Beltrami's pseudospherical model, like Helmholtz's Flatland, was of no use because it was two-dimensional. Helmholtz concentrated instead on what was for Beltrami a merely analytic representation, the auxiliary circle – or, for the three-dimensional case, the auxiliary sphere. In a bold philosophical step, Helmholtz took the function f to be a metric in the space enclosed by the auxiliary sphere. According to this metric, the axioms and theorems of hyperbolic geometry are true. Moreover, the Euclidean straight secants are also straight lines. Helmholtz fashioned for the imagination a spherical universe endowed with the f-metric, or, in the more concrete terms preferred by Helmholtz, a universe in which the solid objects preserve f-length under transport to the same extent that the solid objects of our universe preserve their Euclidean length (for some Euclidean metric) under transport.

Can we intuitively represent this space? We have in fact just done so in general terms and could be as specific as required, appealing to the details of Beltrami's construction. But the representation given so far is, as it were, external. We can imagine this remarkable world in which "solid" objects change size in remarkable ways and even note the relativity of this description: We are no more entitled to judge the behavior of their metric standards by ours than they are to judge ours by theirs. But can we represent this world from the inside, not as an "impartial Euclidean" observer might, but as an inhabitant of that universe would? Helmholtz answered with an interspace traveler story. The Euclidean observer is sent to the center of the mirror universe, and we are told what that universe looks like to him. Since the straight lines in that universe are as straight as his old Euclidean lines, he

> would continue to see the lines of light rays, or the lines of sight of his eyes, as straight lines like those existing in flat space, and like they actually are in [Beltrami's] spherical image of pseudospherical space. The visual image of the objects in pseudospherical space would therefore give him the same impression as if he were at the centre of Beltrami's spherical image. (p. 21)

In particular, since the "universe" is not infinite (by the Euclidean standards to which our space traveler is accustomed) but is bounded by the surface of a sphere of radius R, he would at first think (and "see") that all objects are roughly within a distance R. However, as soon as he started moving around (as he must, according to Helmholtz, if he were to be capable of geometry at all), he would encounter a number of surprises that would alter his way of thinking and therefore – according to Helmholtz – his way of seeing.

This challenge to the idea of *Anschauungsnotwendigkeit* is perhaps the most striking of Helmholtz's criticisms of Kant's philosophy of geometry. He also raised other questions about what Kantians called the "applicability" of Euclidean geometry. In particular, he wondered how Kantians could explain why the very same geometry that is allegedly grounded on pure intuition also happens to be so readily applicable to our empirical world. Helmholtz noted three difficulties. First, Kantians must assume that pure intuition gives them perfectly precise knowledge of the properties of, say, triangles or parallel lines if they are to be sure that Euclidean geometry is true – rather than some very small deviation from it. Second, even if we were endowed with such a supremely accurate mental eye, why should we think that the laws for the geometric triangles of pure intuition agree with the geometric laws governing the rather un-Platonic triangles we encounter in the world? Third, even if the laws of geometry in both the purely geometric and the empirical domains are the same, it does not follow that the metric behavior of ideal objects even remotely resembles that of their real counterparts (see, e.g., "Die Thatsachen in der Wahrnehmung" [1878], especially pp. 397–8). Helmholtz's mirror universe established that two geometric domains in which exactly the same geometric laws are fulfilled may disagree radically on congruence judgments.

It is unlikely that any neo-Kantian ever understood Helmholtz's third point.[11] In reply to the other two, they pointed out that Helmholtz's charges presupposed what we have called the Platonic interpretation of Kant's words and argued that this was quite wrong since for Kant there are no properly geometric objects. For support they could appeal to those passages in which Kant said that when constructing a concept, we construct not an object but "only the form of an object" (*Critique*, A223/B271; see also A239/B298). As we saw, even in light of that construction, "the possibility of its objects would still be doubtful" (*Critique*, A224/B271). The problem of the applicability of "pure geometry" to the world is solved as follows: The constructive synthesis through which the concept (of, say, a triangle) is constructed in imagination "is precisely the same as that which we exercise in the apprehension of an appearance, in making for ourselves an empirical concept of it" (*Critique*, A224/B271).

What distinguished the Helmholtzian neo-Kantians from their "philosophical" counterparts was their reaction to a passage like this: The former regarded it as a problem, the latter as a solution. The Helmholtzians noticed the obvious fact that such a passage perhaps suggests an interesting idea, but is nearly meaningless as it stands. They then did their best to associate some clear and definite sense with such words, relating them to what they or others had discovered in the fields of the psychology of perception or geometry. The philosophical neo-Kantian response to such efforts was invariably an echo of Cohen's professorial "The critics have not understood Kant" (*Die Gegner haben Kant nicht verstanden*).

Helmholtz's philosophies of geometry

Helmholtz's "On the Origin and Significance of the Axioms of Geometry" is a paradigm of a seminal study. It is an explosion of new, deep, and often conflicting ideas on the essence of geometry. In addition to refuting the intuitional necessity of geometry (as recounted in the preceding section), the paper presents (a) Helmholtz's official empiricist philosophy of geometry, which was destined to have a major influence in later decades; (b) an implicit but quite clear refutation of a crucial part of (a); (c) an aprioristic view of geometry, inconsistent with (a); and (d) the first statement of geometric conventionalism – formulated as a possibility, but not fully endorsed because of its obvious conflict with (a) and its apparent conflict with (c). Helmholtz's official empiricism joined with Mill's to inspire an increasingly influential but narrow geometric empiricism. His aprioristic doctrine was avidly grasped by neo-Kantians, who liked it so much that they attributed it to Kant. Poincaré was the first to see clearly beyond Helmholtz, recognizing not only the limitations of geometric empiricism but, more significantly, the consistency and indeed the adequacy of Helmholtz's conventionalist and apriorist doctrines.

Helmholtz's official empiricist doctrine rested on his claim that empirical facts lie at the foundation of geometry. The most basic of these facts is described by the axiom of free mobility, which says that geometric configurations can be moved up to each other without any change in their form or dimensions (*Epistemological Writings*, p. 4). Helmholtz had argued in "On the Facts Underlying Geometry" (1868, *Epistemological Writings*) that from this axiom, plus the "fact" that space is infinite, one could prove the central hypothesis of Riemannian geometry, that the metric must have the form[12]

$$ds^2 = g_{ij}\, dx_i\, dy_j.$$

The assumption that space is infinite seemed justified by physical theory, but what justifies the axiom of free mobility? Helmholtz took it to be an "observational fact" ("On the Origins and Meaning of the Axioms of Geometry," *Epistemological Writings*, p. 15), "something we have all experienced from earliest youth onwards" (p. 4). But it is quite clear that the inference from observation to free mobility is refuted by Helmholtz's own mirror-universe example; for the inhabitants of both universes would "see" from their earliest youth onward that their metersticks and other solid objects satisfied the axiom of free mobility and would also "see" that the metersticks in the other universe did *not* satisfy the axiom. They could not both be right in their inference from experience to free mobility, yet by Helmholtz's reasoning, if one of them is right, so is the other. Hence, neither would be right, and Helmholtz's inference to the axiom is ungrounded.[13]

Side by side with this untenable geometric empiricism one finds in Helmholtz's writings a different and much more promising theory of geometry, for the roots of conventionalism are clearly under the surface of much that Helmholtz has to say about the essence of geometry. One might say that the purpose of geometric conventionalism, as developed by Poincaré and others in the late nineteenth century, was to perform a balancing act widely regarded as quite impossible: to grant to Kantians the a priori character of many scientific principles (geometry prominently included) and at the same time insist on their replaceability and on the existence of equally necessary alternatives to them. In *An Examination of Sir W. Hamilton's Philosophy*, Mill had expressed with characteristic clarity what was, no doubt, a widely shared view among both empiricists and their Kantian opponents. One of the latter had complained because Mill failed to distinguish between the necessity of thinking something and the thinking of that thing as necessary. Mill replied by acknowledging the distinction but noting that the ground for the latter is always an argument for the former. He added, "If I disprove the necessity of thinking the thing at all, I disprove that it must be thought as necessary" (p. 270). Much of the most interesting philosophy of science developed in the past few decades has been inspired by the opposite idea: Many fundamental scientific principles are by no means necessarily thought – indeed, it takes great effort to develop the systems of knowledge that embody them; but their denial also seems oddly impossible – they need not be thought, but if they are thought at all, they must be thought as necessary. This doctrine, whatever its intrinsic merits, is neither empiricist nor Kantian. It emerges directly from ideas that flourished, as we shall see, in Vienna around 1930. But its roots lie in the conventionalism of the late nineteenth century (see Chapter 7) and, even farther back, in the seminal writings of Helmholtz.

No one before Helmholtz was so acutely aware of *both* the need to allow for a variety of systems of geometry *and* the peculiar preempirical role that such systems play in the organization of our knowledge. The opening paragraph of Helmholtz's "On the Facts Underlying Geometry" (1868) states a remarkable fact concerning geometric axioms: To test the axioms we must first know which objects are rigid, which surfaces flat and which edges straight, but "we only decide whether a body is rigid, its side flat and its edges straight, by means of the very propositions whose factual correctness the examination is supposed to show" (*Epistemological Writings*, p. 39).[14] Statements exhibiting this extraordinary feature are not found only in geometry. Another example is Helmholtz's first axiom of the theory of measurement: "If two magnitudes are both alike with a third, they are alike amongst themselves" ("Numbering and Measuring from an Epistemological Viewpoint," *Epistemological Writings*, p. 94). According to Helmholtz, this axiom "is not a law having objective significance; it only determines which physical relations we are allowed to recognize as alikeness" (p. 94). The principle of causality also "has an exceptional status" because "it is the presupposition for the validity of all other [laws]; . . . it is the basis of all thought and conduct. Until we have it we cannot even test it: thus we can only believe it, conduct ourselves according to it" (Königsberger, *Hermann von Helmholtz*, vol. 1, p. 248).

How are we to interpret such statements? In his more lucid moments, Helmholtz suggested that to answer this question we must look closely at how to distinguish within our knowledge between what has "objectively valid sense" and what is merely "definition or the consequence of definitions, or depends on the form of description" ("On the Facts Underlying Geometry," *Epistemological Writings*, p. 39). Thus, at times he was inclined to think of geometric axioms as something like "definitions" and asserted that the "first axiom of arithmetic," the law that magnitudes equal to a third one must be equal to each other, "can be properly regarded as the definition of equality. The axiom must be satisfied in those cases in which two pairs of magnitudes must be recognized as mutually identical" (*Einleitung*, p. 27; see also "Numbering and Measuring . . .," *Epistemological Writings*, p. 78). Perhaps the most intriguing and striking statement of this position appears in "On the Origin and Significance of the Axioms of Geometry." After suggesting that geometric axioms deal with the mechanical behavior of rigid bodies under motion, he added:

> Of course, one could also understand the concept of rigid geometric spatial configurations as a transcendental concept, formed independently of actual experiences and to which these need not necessarily correspond, as in fact our natural bodies do not correspond in an entirely pure and undistorted manner to the concepts that we have abstracted from them inductively. If we adopted this concept of rigidity understood as an ideal, a strict Kantian surely could then regard geometric axioms as *a priori* propositions given through transcendental intuition, and these propositions could not be confirmed or refuted by any experience because one should first have to decide in agreement with them whether given natural bodies should be considered rigid. But we should then add that under this interpretation, geometric axioms would certainly not be synthetic statements in Kant's sense; for they would then only assert an analytic consequence of the concept of rigid geometric configuration necessary for measurement, since one could accept as rigid only those configurations which satisfied the axioms. (*Schriften zur Erkenntnistheorie*, pp. 23–4)

The interpretation Helmholtz offered here as a possible retreat for a Kantian is exactly the one he had espoused in the remark quoted earlier from "On the Facts Underlying Geometry" (p. 39) concerning how we can decide whether bodies are rigid. We shall examine its implications shortly, when we look at the neo-Kantians' reading of this pregnant remark.

Plugging the leaks

When the Germans began to recover from idealism, the first thing that occurred to them was to go back to Kant and start over again, trying to get it right this time. 'Neo-Kantianism' is the blanket name for a variety of movements that had little more in common than a distrust of the post-Kantians who preceded them and the belief that what Kant meant (but didn't quite manage to say) was profound and true. In this general sense of the term, Helmholtz initiated one of the earliest neo-Kantian movements. In "Uber das Sehen des Menschen" (1855), he called for a re-evaluation and reinterpretation of transcendental philosophy in light of the new research in psychology of perception (pp. 76–7). The great historian of Greek philosophy Eduard Zeller would eventually join Helmholtz in his attempt to offer an image of Kantianism consistent with current science and philosophy. As we shall see, this movement would continue into the twentieth century, manifested in the work of Planck, Schlick, and many others inclined to add a scientific realist twist to transcendental philosophy.

What is generally known as neo-Kantianism, however, is a fragment of this larger movement that had a much weaker interest in science that did Helmholtz or, indeed, Kant. The most important exponent of this "philosophical" neo-Kantianism was Hermann Cohen, founder of the celebrated Marburg school; from this school would eventually emerge Natorp, Heimsoeth, Ortega, and Cassirer. Rickert and Windelband led a different branch of the movement that was more concerned with an extension of Kant's thought to the cultural sciences. Outside the Marburg school, Alois Riehl attempted to show that Kant's picture of knowledge was consistent with the rather unruly behavior since 1800 of the non-cultural sciences. One point on which the "strict" neo-Kantians agreed was that Helmholtz's criticisms of Kant had missed the mark. Wherever there was a genuine discrepancy, Helmholtz was wrong, and wherever

Helmholtz had made an interesting point, the point could be found in Kant, if you knew how to read him. In defending their hero, the neo-Kantians were greatly aided by the protean, dialectical nature of Kant's remarks on pure intuition (see the first section of this chapter, "Kant's mixed messages").

Helmholtz's position can be further clarified if we consider briefly the responses of neo-Kantians to Helmholtz's challenge. There were those who thought Helmholtz was simply not a very good philosopher and those who thought his philosophy was excellent but his Kantian scholarship bad. We shall examine one example from each group.

Cohen quite properly challenged Helmholtz's characterization of rigidity as a physical property that we can recognize in objects as a matter of empirical fact. But he observed, also correctly, that Helmholtz's writings contain a different account of the matter:

> He thought that one could conceive the notion of a rigid geometric spatial configuration as a transcendental concept, and thereby consider the geometric axioms as sentences given through transcendental intuition. But in that case the geometric axioms would turn into analytic sentences. "For they would then assert merely what followed analytically from the concept of the rigid geometric configuration required for measurement." Here Helmholtz is relying on the usual nominal definition of analytic and synthetic, which we have long left behind. The concept of a geometric configuration in general, not to speak of one appropriate for measurement, has no connection with the concept of analytic truth but is, from its origin and character, a synthetic notion; for it presupposes intuition. (*Kants Theorie der Erfahrung*, p. 232)

To his credit, Cohen appears to have been the first to recognize clearly that Kant's use of 'analytic' is ambiguous. Cohen argued that there are two senses of 'analytic' and of 'synthetic' in Kant, in effect, the first and third senses identified in Chapter 1 (see *Kants Theorie der Erfahrung*, chap. 11). Kant sometimes meant by synthetic "predicate not thought in the subject," and at other times he meant "having intuition as the ground of the synthesis." Instead of regarding this as the outcome and source of several confusions, however, Cohen took the ambiguity to be another proof of Kant's subtlety. According to Cohen, the first definition is nominal, whereas the second is real. The distinction between these two sorts of definitions can be illustrated by an example due to the venerable Wolff, who had explained in his *Logic* that a nominal definition of a clock would be "a machine that shows the hours," whereas "if I point out its structure, I give a real definition" (pp. 41–2). Apparently, a real definition gives an account of the causes or sources of the features ascribed in the nominal. The conclusion is that Kant's second definition of 'analytic' is not equivalent to the first, but goes far deeper; it identifies the essence of analyticity.

Clearly, Cohen only succeeded in baptizing the difficulty, for he did not even notice that the extensions of the two definitions differ. Nor did he realize the difference between his "nominal" sense of 'analytic' and our crucial *second* sense – true in virtue of concepts. Like all other Kantians, he uncritically assumed that "knowledge claims that must be derived from given concepts . . . are analytic" in the nominal sense (*Kants Theorie der Erfahrung*, p. 115).[15]

Having missed this crucial distinction, it is only natural that Cohen would confuse the sense of analyticity suggested by Helmholtz's remark above with Kant's nominal sense, and that he should take its inadequacy to Helmholtz's intentions as sufficient reason for concluding that intuition is required for the purposes at hand. But this is clearly an indefensible construal of Helmholtz's words. According to Kant's "nominal" sense, we identify the analytic consequences of a concept *C* by looking at the constituents of *C*. Nothing of the sort is involved in the relation invoked by Helmholtz, however, as Cohen observed. According to Helmholtz, the geometric axioms involving a particular geometric concept *C* follow analytically from *C*, even though these axioms are not grounded on an analysis of *C*. Rather, we do not have access to the concept *C* except through the endorsement of those axioms. As Sellars once put it, certain concepts presuppose laws and are inconceivable without them; a geometric axiom may not tell us anything about points, lines, and so on, but instead tell us something about the concepts of point, line, and so on. On this view, our knowledge of geometric axioms would be very much like what Kant regarded as transcendental knowledge, for it would deal not with objects of any sort, but with our knowledge of objects and, in particular, with that part of our knowledge that seems a priori.

Thus, however obscurely and confusedly, Helmholtz seemed to be appealing to a notion of analyticity that did not involve going into the concept to look for its constituents, but going outside it, to look for "analytic" links with other concepts. We might call this view "holistic," since it recognizes an intimate relation between a concept and a larger context, a propositional context, and takes this context to be in some sense prior to the concept. The propositional context is prior in the sense that the context defines the concept, or better, acceptance of the propositions that form the context is part of what is involved in recognizing the concept for what it is. The analytic consequences of the concept are the Fregean consequences of those claims that, as others would put it (Chapter 14), constitute the concept.

Unlike Cohen, Riehl tried to read Helmholtz's insights into Kant's writings. In *Der philosophische Kriticismus*, Riehl presented a version of Kant's conception of geometry that displayed the rather severe beating Kant's opinions had taken as a result of certain facts about geometry that had emerged over the previous few decades. Riehl recognized that some of the central elements in Kant's theory of mathematics were untenable and endeavored to adjust the doctrine to accommodate these recent, embarrassingly un-Kantian developments.

For example, Riehl frankly acknowledged that he could not make sense of the Kantian notion of a pure intuition in geometry (pt. 1, chap. 2, sec. 2). He saw no way to interpret that notion except as a revival of the idea that forms can subsist independently of the empirical entities in which they are embodied. As a true Kantian, Riehl regarded such objects as metaphysical fantasies of pre-Kantian philosophy. Consequently, there is no a priori representation, either conceptual or intuitional, that can be construed as an entity; instead, there are a priori functions of consciousness that are imposed as conditions of experience (p. 86). Form is no more than an abstract endpoint for an order of sensations (p. 104).

Nonetheless, Riehl thought that Kant was right about everything that mattered and his current critics wrong. If geometric axioms cannot be grounded on analysis or intuition, it does not follow that they must be grounded on fact. There is another possible ground of knowledge that Kant had in fact acknowledged but failed to explore with sufficient depth:

> It is a prejudice to believe that what cannot be derived from pure mathematics must for that very reason be derived from pure experience. Above mathematics and experience there are the dominating principles of logic – and when one proves that certain knowledge claims are neither mathematical nor empirical, one has thereby proved that they have a logical origin. (p. 175)

This is, in fact, the case with Euclid's axioms: "Concerning the fundamental properties of space, what can never be decided either by intuition or by analysis is already decided logically" (p. 178). Geometric concepts are not derived from experience, according to Riehl, nor "proved by means of facts"; they are a priori concepts because they are created through the faculty of thought. Rather than having the facts verifying geometric concepts, we have that "conversely, the facts are to be verified through them" (p. 177). According to Riehl, this doctrine of geometric concepts as "logical" can be found (albeit tacitly and obliquely) in Kant's writings and, quite explicitly, in Helmholtz's:

> No one has expressed more clearly than Helmholtz this independence of the ideal configurations of geometry from their corporeal representations in reality, and the dependence of our knowledge and judgment concerning the latter upon the former, as he does when he says that whether a body is rigid, its surface flat, and its edges straight is to be decided by means of the very (geometric) propositions whose factual (empirical) correctness was to be exhibited by the test. (p. 177)[16]

From his Helmholtzian premises, Riehl tried to derive the Kantian conclusion concerning subjectivity of geometric knowledge: "The subject is certainly not, as Kant taught, the sole carrier of the spatiotemporal relations of appearances; he is indeed the author of their determined thought form" (p. 116). And as in Kant, even though these forms are, in some sense, up to us, we do not really have a choice on what these forms will be: "This form of knowledge is necessarily valid for the conscious grasping of the relations of intuition" (p. 116). In other words, Euclidean space is really necessary after all – not a necessity of intuition, but a "logical" necessity, that is, a necessity grounded on concepts alone. But notice how far this is from Kant's ideas. Now conceptual necessity is not grounded on anything like an analysis of concepts; there is, indeed, no talk about analysis in the nominal sense. Rather, conceptual necessity emerges by some unspecified route, involving somehow the endorsement of certain axioms; these are accepted or recognized as true (in some sense of that expression, perhaps only misleadingly associated with its constituent phrases), somehow in virtue of the concepts involved – or, as some would put it later, in virtue of the meanings involved. When carried to its natural conclusion, this line of thought would lead to the view that even though each set of geometric axioms is, in the appropriate sense, "logically" true, it is also the case that each set of axioms is as good as any other. Helmholtz's "transcendental" version of geometry led inevitably to the principle of tolerance in geometry.

4

Frege's semantics and the a priori in arithmetic

Wouldn't Locke's sensualism, Berkeley's idealism, and so much more that is tied up with these philosophies have been impossible if they had distinguished adequately between thinking in the narrow [objective] sense and representing; between the constituents (concepts, objects, relations) and the representations? Even if human thinking does not take place without representations, the content of a judgment is something objective, the same for all. . . . What we are saying for the whole content is true also of its constituents that we can distinguish within it.

Frege, Draft of a reply to Kerry, *Nachlass*

The erroneous belief that a thought (a judgment, as it is usually called) is something psychological like a representation . . . leads necessarily to epistemological idealism.

Frege, "Logik," *Nachlass*

Through the present example . . . we see how pure thought, irrespective of the content given by the senses or even by an a priori intuition, can bring forth judgments deriving solely from the content that springs from its own constitution, which at first sight appear to be possible only on the basis of some intuition. One can compare this with condensation, through which it is possible to transform the air that to a child's consciousness appears as nothing into an invisible fluid in the shape of drops.

Frege, *Begriffschrift*

Before Frege, the best logic texts might have started with a paragraph such as this:

Every categorical proposition has a subject, a predicate, a copula, a quality, and a quantity. Subject and predicate are called 'terms'. For example, in 'the pious man is happy', 'the pious man' and 'happy' are terms of which 'the pious man' is the subject, 'happy' is the predicate and 'is' is the copula. The 'quality' of the proposition is affirmation or negation . . . the 'quantity' of a proposition is its universality or particularity. (Leibniz, *Opuscules et fragments inédits de Leibniz*, pp. 77–8)

Frege radically altered the character of logic. He rejected the traditional doctrine of all five of these categories and offered a new account that guided the development of logic for the following century. Frege re-

placed the partition between subject and predicate with one between object and function. He argued that the copula is not a separate element linking subject and predicate but merely a part or function of the concept displayed in its unsaturatedness; that the category of quality derives from a confusion between unasserted propositional content and its assertion; and that the proper interpretation of quantity requires a theory of quantification that recognizes the functional character of the quantified concept and the existence of higher-level concepts.

It is widely agreed that these discoveries signaled the birth of modern logic. They are, however, no more than the by-products of the much more fundamental enterprise that inspired Frege from his earliest writings: an investigation of the character of what we say when we convey information by means of judgments – not just of what we do say, but of what we could say or judge. From his earliest writings Frege's main concern was with meaning or content, with what he called "the logical" – that is, with semantics.

Begriffsschrift

According to Kant's early conception of mathematical knowledge, the distinguishing virtue of mathematical symbolism is that it represents isomorphically the features of its topic. In arithmetic, for example, Kant had argued that symbols, with their capacity to grow and diminish and their mutual relations, offer a model of the corresponding features of numbers. He thought the case of geometry was even more striking, because there the symbols actually resemble the symbolized. "Mathematical symbols," he argued,

are sensible vehicles of knowledge, so that one can be as confident with them that no concept has been neglected and that each single comparison has taken place through simple rules, etc., as one is of what one sees with one's eyes. The task is greatly facilitated by the fact that one must not think things in their general representation, but only about the signs known individually and with sensible knowledge. In the case of philosophy, on the contrary, words, the symbols of philosophical knowledge, serve only to remind us of the general concept being designated. One must always keep their meaning before one's eyes and the pure understanding maintained in constant effort; and how imperceptibly does a characteristic of an abstracted concept escape us, for there is nothing sensible to reveal to us its omission. (Kant, *Untersuchung über natürliche Theologie*, pp. 291–2; see also *Critique*, A715–18/B743–6)

And elsewhere he added:

The signs employed in philosophical considerations are nothing more than words that fail to represent through their own composition the partial concepts that constitute the whole idea signified by the word; nor can their connection designate the relations of philosophical thoughts. That is why in each act of thinking for this mode of knowledge one must have the thing itself before one's eyes, and it becomes necessary to represent the general in abstracto, without being able to avail oneself of that important and helpful device of handling only single signs rather than general concepts of the thing itself. (Kant, *Untersuchung über natürliche Theologie*, pp. 278–9)[1]

Thus, in Kant's youthful opinion, the symbolism of mathematics was what he might have called an *Anschauungsschrift*, a symbolic system designed to display in sensible intuition a reliable model of the domain of mathematical discourse. Because of the constructive nature of its topic, mathematics lends itself perfectly to isomorphic representation. Philosophy, in contrast, deals with given, nonconstructed concepts and therefore is not capable of this sort of treatment. In other words, there is an *Anschauungsschrift*, but there is no *Begriffsschrift*; and even if there were one, it would interest the mathematician and not the philosopher.

Frege's first book, his *Begriffsschrift* of 1879, put forth a program that directly opposed Kant. Its aim was to design a symbolism that would do for philosophy what Kant thought could be done only for mathematics – a symbolism that portrays not the things it is about but what we may say about them, that gives a picture not of things thought but of thought itself, objectively considered. "Right from the start," he explained in a retrospective account, "I had in mind the *expression of a content*. . . . But the content must be given more precisely than in a natural language" ("Booles rechnende Logik," p. 13).

Unlike Bolzano, Frege recognized from the very beginning that for most sentences of natural languages "the connection of words corresponds only partially to the structure of the concepts" ("Booles rechnende Logik," p. 13). But instead of drawing Kant's defeatist conclusion, Frege attempted to identify what others would call a "perfect language," a fragment of German that expressed perspicuously the content of what we say. "The business of the logician," he explained, "is to conduct an ongoing struggle . . . in part against language and grammar insofar as they fail to give clear expression to the logical" ("Logik" [1879–91], *Nachlass*, p. 7).

"The logical" – it would be a serious error to misunderstand what Frege meant by this recurring expression in his early writings. What Frege and Russell called "logical," what Husserl called a "logical" investigation, what Meinong called "*Gegenstandstheorie*,"[2] and what Wittgenstein termed a "logico-philosophical" observation are close relatives; they should not be confused with what is now called logic, after formalism and set theory have come to dominate the field. Their "logic" was our semantics, a doctrine of content, its nature and structure, and not merely of its "formal" fragment.

For example, Frege explained that an understanding of several languages reveals the fact that natural languages contain a large number of nonrepresentational features, elements that do not stand for anything "logical." He concluded that a familiarity with several languages is quite useful, because "differences between languages can reduce the difficulty in grasping the logical" ("Logik" [1879–91], *Nachlass*, p. 6; also in a later draft, "Logik" [1897], *Nachlass*, p. 154). When Frege defined as his goal "to isolate what is logical" ("Logik" [1879–91], *Nachlass*, p. 6) and "to separate sharply the psychological from the logical, the subjective from the objective" (*The Foundations of Arithmetic*, p. xxii), he was clearly implying that his target, the logical or objective element in thought, is not what remains in judgment when content is excluded but what remains when we discard the specifically *psychological* element.[3]

Frege devoted considerable effort to separating his own conceptions of "logic" from that of the mere computational logicians such as Jevons, Boole, and Schroeder. Whereas these people, he explained, were engaged in the Leibnizian project of developing a calculus ratiocinator, his own goal was the much more ambitious one of designing a *lingua characteristica*. Traditional logicians were concerned basically with the problem of identifying mathematical algorithms aimed at solving traditional logical problems – what follows from what, what is valid, and so on. Frege's goal went far beyond what we now call formal logic and into semantics, meanings, and contents, where he found the ultimate foundation of inference, validity, and much more.

Frege's criticisms of Boole are particularly revealing. In Boole's work, he complained, "content has been entirely ignored" ("Booles rechnende Logik," p. 13). Boole's aim was to produce algorithms for solving logical problems, but his strategy could not satisfy "anyone interested in keeping the closest link in the relations between signs and the relations between the things themselves" (p. 13). Unlike Boole, "I did not want to represent an abstract logic in formulas, but to express a content through written signs in a more precise and surveyable (*übersichtlicherer*) fashion" ("Ueber den Zweck der Begriffsschrifts" [1882–3], *Begriffsschrift*, p. 97). "Boole's formula-language symbolic logic represents only the formal part of language, and that only incompletely" ("Booles rechnende Logik," p. 14). "Boole's formula language presents only a part of our thought; the whole of it cannot be taken care of by a machine or replaced by a purely mechanical activity" (p. 39). It is the whole of our thought that concerns the *lingua characteristica*. "We can derive a real usefulness from [a formula language] only when the content is not merely indicated but constructed out of its constituents by means of the same logical signs that are used in the computation" (p. 39).

Frege's project involved identifying a fragment of the German language that fulfills two conditions: (a) Every German sentence has a translation into this fragment, and (b) the grammatical form of every sentence in this fragment mirrors isomorphically the constituents of the content it expresses, as well as their arrangement in that content. The fact that *Begriffsschrift* introduced symbols not available in pre-Fregean German, essential as it was for the practical feasibility of the project, was a semantically insignificant factor; for such symbols could be entirely eliminated in principle, in favor of standard German language expressions – precisely those in terms of which the meanings of Frege's symbols were expressed. Given this "perfect language," derivability relations and conditions for validity would follow without any need to appeal to algebraic tricks extrinsic to the propositions under consideration, but simply by an analysis of the constituents of the claims involved and their structural relations, as manifested perspicuously (i.e., syntactically) in their reformulation in the perfect language. In effect, the idea was to produce a language in which, even though inference was based on meaning, one need no longer think about meanings (just as Kant had said that the nature of mathematical symbolism makes it unnecessary to think about *its* meaning), since one could now restrict oneself to the signs "present to the senses" and their symbolic correlations. Little wonder that fifty years later a heretic disciple was tempted to cut the remaining link with meanings and take the perspicuous language as the whole concern of logic and scientific philosophy.

How is one to identify the details of this perfect language? Frege's strategy, and the epoch-making results that emerged from it, seem to have been inspired by a semantic conception that, oddly enough, he never made quite explicit. Indeed, the central elements of that semantics were *essentially* tacit; for as soon as he recognized their presence in his system (apparently in the late 1880s), he hastened to eliminate them.

The basic semantic categories

The similarity between Frege's early semantics and Bolzano's is quite striking. As we saw, Frege emphasized in the beginning of his *Grundlagen* the importance of separating "sharply the psychological from the logical, the subjective from the objective"[4] (p. xxii). One must be especially careful, he urged, to distinguish between objective and subjective representations; the former are "the same for all," but the latter are not. A word is often accompanied by a subjective representation that nevertheless "is not its meaning"; "the word . . . means an objective representation" (p. 37). Extending an undeserved olive branch to the past, Frege added, "It is because Kant associated both meanings with the word ['representation'] that his doctrine assumed such a subjective, idealist complexion, and his true view [!] was made so difficult to discover" (*The Foundations of Arithmetic*, p. 37).

There can be no doubt, however, that Frege's treatment of these matters is even farther removed from Kant's than Bolzano's. Unlike Kant, and in agreement with his general goal of desubjectivizing semantics, Bolzano had distinguished three elements associated with every representation: (a) the subjective representation, (b) its objective counterpart, and (c) its object. But having agreed with Kant that representations are either concepts or intuitions, he had a hard time producing an objective counterpart between the subjective intuition (e.g., the seeing of a rose) and its object (the rose). Bolzano was still far too dependent on that tradition in which all we see are "ideas" and "phenomena." "What I see when someone holds a rose before me," he explained, "is a representation" (*WL*, vol. 1, p. 217) – hence, presumably, not a rose, for roses, unlike objective representations, are in space and time and, unlike subjective representations, persist when the human mind vanishes.[5] The object of intuition appears to be subjective and the objective counterpart remains a mystery. We are in one of the darkest corners of Bolzano's philosophy.

From the very beginning Frege cast these Kantian hesitations aside: "Objective representations," he explained in the *Grundlagen*, "can be divided into concepts and objects," not into concepts and intuitions[6] (p. 37). Objective representations matter not for their own sake but for what we can do with them by linking them together, for when the link is appropriate, the outcome is something akin to a Kantian judgment minus its subjective, psychological component; it is the content of a judgment minus its subjective dimension. This is what Frege called a *content of possible judgment* (*beurteilbarer Inhalt*, henceforth, cpj). A cpj is the target of what Russell later would call propositional attitudes: understanding, assuming, asserting, wondering, and so on. Most importantly for Frege, these things are what we claim to know. Thus, an understanding of human knowledge depends on a proper understanding of cpj's: A theory of knowledge presupposes a semantics, and until we understand the latter, we ought not try to deal with the former.

The whole of Frege's early semantics centered around these three basic notions: concept, object, and cpj. The distance between Frege and Kant is underscored by the paucity of remarks on that most disturbing of Kantian problems, the character of the objects of knowledge and their constitution through the categories. Objects are no problem for Frege – they are the tables and chairs of everyday experience, the numbers and classes of mathematical knowledge, the truth values of his logic, and so on. His semantic interest is aimed almost exclusively at the other two topics, concepts and cpj's. Moreover, what he has to say about them has an oddly complementary nature, since his account of each depends on an account of the other, so that one is bound to understand them jointly or not at all. Whether dialectical or circular, Frege's reasoning on this topic does not lend itself graciously to didactic exposition. We begin by looking at the way he came to think about concepts, contrasting his views with the more standard picture of the matter.

The concept: roots of holism and unsaturation

According to the abstractionist theory of the concept that was still quite popular in Frege's time, the best way to understand what concepts *are* is to look at their genesis. It is important to emphasize that the young Frege agreed with this point, even if he disagreed with the abstractionist account of how human beings define or construct concepts. Abstractionists claimed that concepts emerge through a process that takes us from certain data to a concept via a process of elimination. The given at the starting point of this process appears to consist of intuitions. A succinct statement of the abstractionist theory is found in Husserl's early writings. Husserl began his *Habilitationsschrift* devoted to the concept of number, by explaining that he would assume

> that concepts originate through a comparison of the specific representations that fall under them. Disregarding the characteristics (*Merkmale*) that differ, one holds firmly to the ones that are common; and these latter are the ones which then constitute the general concept. (*Begriff der Zahl*, p. 299)

The standard designations of concepts as "general representations" and "common nouns" are motivated by the widely held belief that the essential characteristic of a concept is its capacity to refer to more than one thing. It was therefore widely held that a theory of the concept should explain, above all, that power of multiple reference. The abstractionist theory appears inspired by the curious idea that one can explain the generality of *reference* in a concept by involving a multitude of things in the story of how the concept came to emerge. But as Frege pointed out, the theory has no way to distinguish between a case in which one decides to neglect features of an object because they differ from those of others and one in which a person is simply forgetful and lets the details of a single instance drop out of his memory. Frege's biting and accurate critique of this procedure (as displayed in Husserl's *Philosophie der Arithmetik*) is worth recalling:

> [Detaching our attention] is particularly effective. We attend less to a property, and it disappears. By making one characteristic after another disappear, we get more and more abstract concepts. . . . Inattention is a most efficacious logical faculty; presumably this accounts for the absentmindedness of professors. Suppose there are a black and a white cat sitting side by side before us. We stop attending to their color, and they become colorless, but are still sitting side by side. We stop attending to their posture, and they are no longer sitting (though they have not assumed another posture), but each one is still in its place. We stop attending to position; they cease to have place, but still remain different. In this way, perhaps, we obtain from each one of them a general concept of Cat. By continued application of this procedure, we obtain from each object a more and more bloodless phantom. (From Frege's review of Husserl's *Philosophie der Arithmetik* [1894], *Translations*, pp. 84–5)

If concepts cannot derive from abstraction, how *do* they emerge? In Frege's view, the process of concept formation is dependent on that of judgment. Frege noted that logicians, from Aristotle to Boole, had seen logic as a theory of inference in which the construction of concepts "is presupposed as something that has already been completed." He contrasted this with his own approach: "I start from judgments and their content, not from concepts. . . . I allow the formation of concepts to proceed only from judgments" ("Booles rechnende Logik," p. 17); the representations of properties and relations "come simultaneously with the first judgment in which they are ascribed to things" (p. 19).[7]

Frege's strategy for dealing with the concept was to assume that we are given the cpj's and their constituent objects; we then generate concepts by carving them out of cpj's, as we exclude this or that object from the given cpj.[8]

In its basic outline, Frege's doctrine of concept formation corresponds rather closely to holistic remarks found in Bolzano and other previous writers. But nothing in the work of Bolzano, or of anyone else, compares with the richness of detail and results that emerged when Frege adopted this strategy. For one thing, most pre-Fregean philosophers, including Bolzano, relied rather blindly on surface grammar and on the subject–copula–predicate form. For another, the holistic doctrine became a fruitful semantic tool only when conjoined with another idea entirely original with Frege: that the step from judgment to concept is analogous to a similar step taken in mathematics, linking a function and its values.

The instrument of generality in mathematics is the variable, and its most frequent context is the function name. Consider, for example, the function

$$2 \cdot x^3 + x;$$

the values it assigns to 1, 2, and so on are

$$2 \cdot 1^3 + 1, \qquad 2 \cdot 2^3 + 2, \qquad \text{and so on.}$$

It occurred to Frege that if we look at this process backward, we obtain a very enlightening picture of the nature of a function. Instead of moving from the function to its values, move from the values (or, rather, from those *particular* names of the values) to the function (or, rather, to the name of the function):

From this we may discern that the essence of the function lies in what is common to these expressions; i.e., in what is present in

'$2\cdot(x)^3 + x$'

over and above the letter 'x'. We could write this as follows:

'$2\cdot(\ )^3 + (\ )$'.

("Funktion und Begriff" [1891], *Kleine Schriften*, p. 128)

The (name of the) function is therefore seen as deriving from (certain names of) its values by removing from the latter (the names of) one or more objects. Frege drew an important conclusion: The function not only correlates arguments and values but is also "unsaturated," "in need of completion," and like concepts, "predicative."[9]

Frege saw that the backward process from argument (name) to function (name) may be applied not only to expressions that designate numbers but to all meaningful expressions, including sentences, and recognized in this the key to the nature of general representation. For example, starting with

(*) John is tall,

we can take John (or 'John') out and are left with

x is tall,

a function of a more general type than the mathematical variety, since it does not take numbers as values. According to Frege, this is the concept *tall* or, as he preferred to write it, *x is tall*. The holistic and the functional idea are now linked through the fact that the concept *tall* is the function we obtain when we carve out of a cpj such as (*) an object such as John.

In 1882 Frege explained to a correspondent (possibly Marty), "I do not believe that concept formation can precede judgment . . . but I think of a concept as having arisen by decomposition from a cpj" (Letter to Marty [?] [1882], *Wiss. Briefwechsel*, p. 164). He went on to explain how the construction takes place. Consider the cpj

(*) $3 > 2$.

Depending on what we choose to regard as the "subject" of (*), the claim will be regarded as the attribution of different concepts to different objects. If we regard 3 as its subject, for example, then (*) says that 3 falls under the concept *being greater than 2*. A similar concept results if 2 is chosen as the subject. Finally,

we can also regard '3 and 2' as a complex subject. As a predicate we then have the concept of the relation of the greater to the smaller. In general, I represent the falling of an individual under a concept by $F(x)$, where x is the subject (argument) and $F(\ )$ the predicate (function), and where the empty place in the parentheses after F represents nonsaturation. (p. 164)

Two years later he would explain in *Grundlagen*:

When from a cpj that deals with the objects *a* and *b* we subtract *a* and *b*, we obtain a remainder, a relation concept that is, accordingly, in need of completion in two ways. If from the sentence

"the earth is more massive than the moon"

we substract "the earth," we obtain the concept "more massive than the moon." If alternatively, we substract the object "the moon," we get the concept "more massive than the earth." But if we subtract them both at once, then we are left with a relation concept. (p. 82)

Frege's holist theory of the concept was revolutionary. Before Frege, singular and general representations had been considered two species of the same semantic natural kind. He was not simply rejecting the old idea of the concept as a "common name," as a name of more than one thing (although he was certainly doing that as well). Behind the term 'general representations' lay the idea that both sorts of representations emerge from basically the same process, such as that proposed by abstraction theories. According to these theories, there is a primordial sort of singular representation (the given), which is the most powerful, desirable, and complete form of representation; we obtain a less specific form of representation by weakening the features of the primordial ones. This character is mysteriously transmitted even to a priori concepts. On Frege's view, so-called general representations are so different from their singular counterparts that one might better regard them as falling into two radically distinct semantic categories. The difference is revealed by the difference in the procedures that lead to their emergence. Singular representations are proper names, and these are supposed to be given quite independently of judgment; general representations emerge only *after* judgment. Traditional logic texts treated concept, judgment, and reasoning in that order; Frege was proposing that the order of the first two be reversed. His new perspective not only showed how wrong it is to think of concepts as general names, ignoring their predicative (unsaturated) dimension; it also showed how wrong it is to think of concepts as *general* representations; for

x is identical with Sir Walter Scott

is a concept that, by its very essence, lacks the capacity to designate anything more than one object.

Quantification and arithmetic

Most of Frege's colleagues were suspicious of so much semantic subtlety. Why should one care about the nature of concepts? Here we can only scratch the surface of an answer by briefly recalling Frege's two major achievements of this early period: his theory of quantification and his analysis of arithmetic. As we shall see, they are entirely dependent on his semantic conception of things, and the picture they offer of mathematical knowledge is both an enormous improvement on earlier efforts and a further major step away from the Kantian position on the role of concepts and intuition in a priori knowledge.

Quantification

Traditionally, quantified sentences such as 'All As are Bs' and 'Some As are Bs' had been thought to make subject–predicate claims, their subjects being 'all As' and 'some As', respectively. As we shall see, Russell's *Principles,* written more than two decades after Frege's *Begriffsschrift,* was still inspired by this pre-Fregean view of quantification.[10] One could hardly provide more striking evidence of the revolutionary character of Frege's views.

Frege analyzed the content of

(1) All men are mortal

as follows: He first considered the apparently unrelated matter of how complex propositional contents emerge from simpler ones by means of logical operations such as negation and material implication. Consider, for example, how

(2) If John is a man, then John is mortal

is formed out of

(3) John is a man

and

(4) John is mortal.

According to Frege, the cpj expressed by (2) is uniquely characterized when we give what we would now call its truth conditions, that is, when we say that it is true in all and only those circumstances that make (4) true or (3) false.[11] A similar construction introduces all other connectives (negation suffices to define all the rest). At this point Frege had characterized the language of propositional logic. Next consider what happens when we remove the object John (the word 'John'?) from (2); Frege represented the result as

(5) If x is a man, then x is mortal

where the variable 'x' is no more than a convenient device to identify the blank left by the departed name. As we know, (5) stands for a Fregean concept, one that will be true of an object a precisely when a is not a man or else it is mortal. Now (1) can be interpreted as saying *about that concept* that every single object in the universe is an instance of it. A quantified sentence such as (1) is therefore to be interpreted as involving a concept of "higher order," a concept that applies not, like (5), to objects, but to (first-level) concepts, to (5) itself. The universal quantifier is no more and no less than a second-level concept that applies to the first-level concept (5) precisely when (5) is true of every single object – in other words, precisely when (1) is true. Appearances and tradition notwithstanding, (1) does not say anything about all men, or about *all men,* or about any particular man. Nor is quantification to be construed in the medieval–Russellian manner, as an operation that transforms the so-called subject concept (e.g., *man*) into a denoting expression (e.g., *all men*). It is a second-level concept whose topic is what we now call the scope of the quantifier. The Fregean picture of quantification readily lends itself to iteration. Once existential quantification is defined from its universal counterpart and negation in the standard manner, we can easily disclose ambiguities hidden in ordinary language – such as that between simple and uniform convergence.

Frege's *Begriffsschrift* was not really a new language but a fragment of German; everything that can be said in Frege's concept script can be said equally unambiguously in German. Yet the German language does not contain expressions that are manageable and unambiguous and that serve the same purpose as Frege's new symbolism. That is why his notational system became (pragmatically) essential. Given time and unlimited patience, one could explain in a natural language the ambiguities easily sorted out in Frege's notation; but nothing in unreconstructed natural languages can do nearly as well.

By combining his semantic insights with the new notational system, Frege displayed explicitly in *Begriffsschrift* the first clear formulation of a formal language with propositional connectives and quantification over both individuals and first-level functions. Beyond this, the monograph identified a set of "logical laws" and rules of inference for inferring other laws from them, even though no effort was made to determine what distinguishing feature of these formulas determines their membership in that class. The system was developed with a subtlety and rigor far exceeding the standards displayed in later decades by Peano, Russell, and even Hilbert in his early logical writings. The role of inference rules, clearly recognized in *Begriffsschrift*, would remain a mystery to Frege's most distinguished colleagues until well into the twentieth century.

Remarkable as these achievements were, they were only the beginning. Five years later Frege would publish another short monograph aimed at showing that when we get our semantic facts straight, it becomes clear that Kant's philosophy of arithmetic – indeed, of the whole science of number – is incorrect.

Arithmetic

In *Grundlagen* Frege was concerned, once again, with the proper semantic interpretation of certain notions. Here, however, his emphasis was not on concepts traditionally assigned to the domain of logic such as quantification, copulation, and sentential connection, but on notions widely regarded as extralogical. His topic was number.

Three years after *Grundlagen* was published, Husserl explained in his *Habilitationsschrift*:

> Today it is generally agreed that a rigorous and thorough development of higher analysis (the totality of *arithmetica universalis* in Newton's sense), excluding all auxiliary concepts borrowed from geometry, would have to emanate from elementary arithmetic alone, in which analysis is grounded. But this elementary arithmetic has, as a matter of fact, its sole foundation in the concept of number. (*Begriff der Zahl*, p. 294)

This was so because after half a century of work by Cauchy, Dedekind, Cantor, and Husserl's teacher, Weierstrass, a large portion of Bolzano's project to conceptualize analysis had been fulfilled. This achievement is sometimes known as the "arithmetization" of the calculus, because it reduced all of the mathematics of numbers to the science of natural numbers plus a vaguely logical discipline of classes. This project showed that whatever foundation there is to analysis is to be found in the theory of the natural numbers. One could conclude that any intuition present in analysis is to be found in arithmetic – hence the philosophical significance of the nature of arithmetic.

When Frege turned to examine this issue in *Grundlagen* (1884), he characteristically posed it at a pre-epistemological, semantic level. The questions he asked did not have the familiar Kantian ring: How does arithmetic acquire objective validity? How can it be applied? How do we come to know its objects and its justification? His basic questions were, What do number statements say and what do they say it about? The question of ground was not even raised.

Frege began by singling out for special attention what he called "number statements," statements that say there are *n* things of a given sort *T*; for example,

(*) Jupiter has four moons.

He reviewed the two basic doctrines on what these statements are about. According to the first one, (*) is about a certain object, perhaps the moons of Jupiter, or the class of those moons, or the "heap" or conglomerate that they somehow constitute. According to the second, (*) is not about any objective element but about some subjective counterpart such as our representation of moons, or certain mental processes of "addition," conjunction, or whatever. Frege described and criticized various versions of these two possible interpretations, refuting decisively every single one of them. This portion of the monograph, leading to Frege's own proposed solution, is one of the most dazzling examples of sound philosophical writing ever produced.

The key, according to Frege, is to recognize that the question "How many?" makes no sense if we identify an object as its target, but acquires sense if its target is a concept. If we point to the cards on the table and ask about them, "How many?" the right answer may be eighty (cards), two (packs of cards), or almost any other number we may choose. To allow for a definite, unique answer, we must make reference, explicit or implicit, to a concept (x is a card, x is a pack of cards, etc.). Since the number attribute is fixed only when the concept is determined, it is natural to regard the concept itself as the topic of the number statement. Thus, syntactic appearances notwithstanding, number statements, *like quantified statements*, are about concepts. In fact, Frege explained, they *are* quantified statements, even though ordinary language hides that fact.

For example, the number statement 'The earth has one moon' is about the concept *is a moon of the earth*, and it says that exactly one object (at least one and at most one) falls under it. To say that an object falls under a given concept is, according to *Begriffsschrift*, simply to apply an existential quantifier to it; and to say that at most one object falls under it is easy to do in terms of universal quantification, identity, and the propositional connectives. In a similar fashion one can render an interpretation of (*) and of all statements of the form 'There are *n Fs*', where '*n*' is a *standard* numeral.

This brilliant solution, later described by Frege as the "most fundamental of my results" in *Grundlagen* (*Grundgesetze*, p. ix), showed that a wide range of statements previously regarded as extralogical and involving an appeal to either empirical intuition (Mill) or to pure intuition (Kant) involved only reference to concepts. Bolzano's hope of conceptualizing mathematics had been taken a giant step forward.

We shall return in Chapter 7 to Frege's conception of logic and arithmetic, and to some of the problems it left unsolved. But whatever the problems, there can be no doubt that Frege's work shed a great deal of light on the character of arithmetical knowledge. It is worth emphasizing once again that the clearer mathematics became in his hands, the further it moved away from Kantian doctrine. It wasn't so much that Frege had argued that arithmetic is analytic. Since he had defined 'analytic' as derivable from logic and definitions, one might consider his conception of logic so different from Kant's as to make a conflict on this issue virtually impossible. Rather, the basic conflict was, as we might expect, on the subject of intuition. On one thing Kant and Frege agreed: Logic is grounded at the level of the understanding where sensibility and its forms play no role (recall the quotation from Frege at the beginning of this chapter). The reduction of arithmetic to logic conflicted with the Kantian postulation of an appeal to sensibility in the realm of arithmetic. As we shall have reason to observe at a later stage, there is no explicit theory of the ground of analytic knowledge in Frege's writings; but these brief considerations indicate that he intended to place both the content and the ground of arithmetical knowledge at the level of Kant's analytic doctrine, to the exclusion of his aesthetic doctrine.

Quite apart from their relevance to Kantianism, Frege's early treatments of quantification and number statements provided the model for a reconstructive conception of language that would inspire a variety of schools within the analytic tradition. No one before Frege had taken so seriously the task of casting a language in which ordinary things could be said in an extraordinarily clear manner. No one before him had applied his translation techniques so effectively to the solution or dissolution of philosophical problems. We shall soon see that this aspect of Frege's approach appeared independently in the works of Russell, and then of Wittgenstein and Carnap. But we must now turn to yet another major discovery of Frege's, one that his successors were far slower to appreciate.

The discovery of sense

Few things have proved more difficult to achieve in the development of semantics than recognition of the fact that between our subjective representations and the world of things we talk about, there is a third element: what we say. Perhaps Chapter 5, which deals mostly with developments that took place more than a decade after Frege came to grips with the situation, should be read before this account of Frege's discovery; for that chapter describes the difficulties that many of the best philosophical minds encountered near the turn of the century largely because they were unable to understand that what we say, sense, cannot be constituted either from psychological content or from the real-world correlates of our representations. Psychologistic logicians had pursued the first approach; most of Frege's successors pursued the second. They all attempted to understand sense by forcing it into a world to which it does not belong.

As we have seen, many of the things Frege said during his first decade of research suggest that he also had started on the assumption that cpj's should have as constituents both the objects they deal with and the concepts attributed to them. The ominous hesitations on use and mention implicit in some of the references in the section on roots of holism and unsaturation are not Homer's nods, but symptoms of the fact that Frege had not thought through the implications of what he was saying. Moreover, from the very beginning, Frege had acknowledged a major exception to the general account of content, an exception that would become the rule in the picture of propositional content that emerged in the 1890s. Although we do not know much about the course of Frege's thought as he moved toward the recognition of sense, the opening remarks of "On Sense and Reference" (1892) give us an appreciation of the role played by one of the most influential factors: the nature of identity. Identity had posed a difficulty to Frege's cpj semantics from the very beginning. Consider, for example, the following identity statements:

(*) The author of *Waverley* = the author of *Waverley*

and

(**) The author of *Waverley* = Scott.

What does cpj semantics say about the content conveyed by these sentences? If we interpret identity as relational, its content will be a cpj containing the identity relation with both "holes" saturated by the objects named by the relevant terms. Then, since the author of *Waverley is* Scott, the content of (*) is identical to the content of (**). Consequently, if knowing that a sentence is true is knowing that its cpj is true, anyone who knows that (*) is true must also know that (**) is true. As Frege put it in "On Sense and Reference," "Now if we were to regard equality as a relation between that which the names '*a*' and '*b*' designate, it would seem that $a = b$ could not differ from $a = a$ (i.e. provided $a = b$ is true)" (*Translations*, p. 56). Frege seemed to have thought at first that this difficulty did not point to a problem within his semantics but revealed instead the very idiosyncratic character of the identity relation. Thus, he came to think that whereas all other relations relate their objects, identity says something about a very different domain. Let us look closely at what that domain was.

Superficially, section 8 of *Begriffsschrift* (in which the topic is discussed) appears to say that identity must be construed as a relation between purely syntactic expressions, as if (*) and (**) were really about 'Scott' and 'the author of *Waverley*'. But a more careful reading shows that this is not so. Frege's point is that different names of the same object will, in general, be associated with different ways of determining which object they name. For example, we might give to ourselves a geometric point "directly in intuition" and baptize it with the proper name '*A*', and then we might give it as the point fulfilling certain geometric conditions:

> To each way of determining the point there corresponds a particular name. Hence the need for a sign for identity of content (*Inhalt*) rests upon the following consideration: the same content can be completely determined in different ways; but the content of a *judgment* is that in a particular case *two ways of determining it* (*Bestimmungsweisen*) really yield the *same result*. Before this judgment can be made, two distinct names, corresponding to the two ways of determining the content, must be assigned to what these ways determine. (van Heijenoort, *From Frege to Gödel*, p. 21; Frege's italics)[12]

Clearly the cpj's (*) and (**) do not talk about names, nor do they, of course, have names among their constituents. According to Frege, what (*) and (**) *say* is that two methods of determining a referent yield the same referent. If what they say is the cpj, then cpj's have as constituents not the referents but the methods for determining them, that is, what Frege will eventually call the "sense" of the corresponding singular terms.

Thus, circa 1880, whereas the contents of most relational judgments looked like the future Russellian propositions, the content of identity statements looked like the future Fregean *Gedanke*, for they involved the senses rather than the referents of the appropriate names. The main difference with the views of his later self was that the young Frege confused content and topic, thus concluding that the sense of the names in question was the topic of the identity claim.

What happened circa 1890 is that Frege detected and removed this major confusion from his system. First, he decided that identity does not call for a special treatment after all, but that '$a = b$' *is* a relation between what the names a and b signify. It does not follow, however, that the content of this sentence is the same as that of '$a = a$'. In the spirit of *Begriffsschrift*, he reiterated in "Uber Sinn und Bedeutung" (1892) that when 'a' and 'b' refer to the same object, 'a' may differ from 'b' merely in physical shape (as in 'England' and 'Inglaterra') – in which case '$a = a$' and '$a = b$' *do* say exactly the same thing (have the same content); or 'a' may differ from 'b' also in the *way* it designates or refers (*in der Weise, wie es etwas bezeichnet*; p. 41). In the latter case the contents are obviously different even if the objects related are not. What was wrong with

the old solution to the puzzle of identity was that it misidentified the source of the trouble, for the trouble arose only when we failed to distinguish between what we say and what we say it about (or, as Bolzano would put it, between objective representation and its object). On the new solution, a sentence of the form '$a = b$' would concern not the topic (*die Sache selbst*) but the mode of designation (*die Bezeichnungsweise*, p. 41) or, as Frege now called it, the sense (*Sinn*). It was clear that both topic and sense are essential, but they play radically different semantic roles.

Semantic dualism

Two styles of semantic analysis played a prominent role in the semantic tradition, what we will call semantic monism and semantic dualism. Monists and dualists agree on the broadly picture-theoretic assumption that for the purposes of semantic analysis we must partition a language into its basic grammatical units and associate with them appropriate semantic correlates. The conflict lies in the number and character of the semantic entities required: The monist thinks we need associate no more than one semantic entity with each grammatical unit, and all of these entities come basically from the same place, the world. The dualist thinks that we must associate two different elements with each piece of grammar: roughly its contribution to what the sentence says and its contribution to what it is about. The latter is in the world; but where the former is, and even whether this question makes any sense at all, are matters on which the dualist would normally hesitate.

We have seen that the semantic system underlying Frege's thought during the first and most creative decade of his intellectual life was a form of semantic monism. Then, around 1890, he recognized its presence and replaced it by a dualist system. According to Frege's official semantic doctrine after 1890, there are two radically different semantic categories associated with each independently meaningful expression: its sense and its significance. The words '*Sinn*' and '*Bedeutung*' had occurred in other semantic systems both before and after Frege. What distinguished Frege's usage was that he regarded these notions as semantic categories, rather than as names of specific semantic objects or relations. Wittgenstein would acknowledge in the *Tractatus* both a *Sinn* and a *Bedeutung*, for example; yet his system is a version of semantic monism, for on his view only names have a *Bedeutung* and only pictures have a *Sinn*. For Frege every meaningful grammatical unit does *two* semantically relevant things: It expresses its sense and it signifies its significance. By doing each of these things, the meaningful expression has the potential to contribute to two parallel processes whose ultimate goal is symbolically represented in the syntactic structure of the fully analyzed sentence.

Though Frege emphasized the distinction between what we understand and what we are talking about, he thought there must be a very intimate link between those two. Take a sentence such as

(1) The author of *Waverley* is tall.

The universe includes many objects and many properties, but only one of each affects in any way the truth value of (1). The only relevant object is the one who wrote *Waverley*, Scott; the only relevant property is *height*. Whether (1) is true or false depends entirely on whether a single object in the universe has a single property. And the interesting fact is that this information is entirely contained in sentence (1). Since that information consists of two different items, it is natural to think of (1) as being divided into two grammatical units corresponding to each of those items. We may therefore think of it as being partitioned into 'the author of *Waverley*' and '*x* is tall'. We then go on to recognize a very basic semantic feature of (1), that these grammatical units are associated with the truth-relevant parts of the world, that is, with those elements of the world that are the only ones relevant to the determination of whether what (1) says is true. Quite generally, Frege thought, the grammatical units of a claim are associated with the elements that constitute its topic and determine its truth value. For the case of direct discourse, the mode of language of science, it is natural to baptize the class of all such elements "the real world," for it includes all we talk about when we don't talk about talking. In direct discourse each grammatical unit is associated with a corresponding element in the real world, and Frege said of such elements that they are meant by the expression in question, or that they are its "significance" or "meaning" (*Bedeutung*).[13]

Grammatical units partition the world into what matters and what does not as far as truth is concerned; they give us what we have called the truth-relevant elements. Responsible use of language does not require that we have a definite idea of what that partition is, however (and herein lies a verificationist tale). All that is required for the purposes of communication or responsible discourse in general is that what we say be intelligible, and this usually has little to do with being able to tell whether what we say is true, or even with the possession of effective methods to identify either its referents or truth values.

That sense is nonetheless intimately related to the truth-value-relevant elements in the world is strongly suggested by a seldom-noticed fact: Frege took it as self-evident that the grammatical analysis appropriate to the study of sense coincides with the grammatical analysis appropriate to the study of significance. The understanding of the sense of a sentence

is not, for Frege, a holistic phenomenon but comes about through the understanding of the sense of its parts. This is the natural explanation of the fact that we can understand sentences we have never seen before (Letter to Jourdain [1914], *Wiss. Briefwechsel*, p. 127) and is also the reason we need to talk about the sense of parts of the sentence, as well as that of the whole they constitute. But notice that there is, in principle, no reason the grammatical units that provide the building blocks of propositional sense should be the same that provide the truth-relevant features of the world, *unless* sense (and therefore understanding) are essentially a matter of doing something concerning those worldly units. Thus, even if the sense of those expressions that select truth-relevant elements need not be an *effective* method of getting at those elements, it must be some sort of device establishing a correlation, however ineffective, with them.

The realm of sense, ignored by most philosophers before Frege and challenged by many after him, was Frege's response to what has remained to this day one of the central topics of semantics, the character of propositional understanding. This involved at least two major problems: What is it that we understand, and under what circumstances does understanding take place (or how do we distinguish apparent propositional understanding from *the real thing*)?

The solution to these problems that would prevail among members of the semantic tradition after Frege was that the targets of propositional attitudes are certain real entities (for Russell, particulars or universals; for Husserl, essences, etc.) and that the basic attitude of understanding is, in effect, intuition (acquaintance, *Wesensschau*, self-evidence). In drawing a sharp distinction between sense and the real world, Frege was denying that the things we understand – that is, what we say and its constituents – are elements of the real world, whether individuals or properties, universals, concepts, or essences. He also denied that understanding is a glorified form of "seeing" aimed at these entities. Frege's subtle point is that while understanding does involve "giving" the object in question, it need not necessarily be "given" in the mode of acquaintance. Understanding and acquaintance should not be equated; indeed, the former is usually concerned with what might well be called "knowledge by description." What we understand when we understand a definite description, for example, is its sense; and this, in turn, is a way of giving the referent, but a way that need not be effective or even lead to an end result (a referent).

Thus, in Fregean doctrine the sentence is the real center of semantic activity. To understand the character of a constituent of a sentence is to know the job it is supposed to perform in connection with its partners in the sentence, in order to associate that sentence with each of its two semantic dimensions. The basic task of the sentence is to say something. What the sentence says is its sense, and the sense of each part is simply its contribution to this message. On the matter of sense, the parts of the sentence are mere instruments, totally subservient to the overriding goal, which is to constitute the sense of the sentence. The sense of a part of a sentence is interesting only because it helps us understand the sense of the sentence. The second semantic task concerns significance. Since the significance of a sentence is its truth value, the reason we are interested in the significance of its grammatical units cannot be related to the "construction" of the significance of the whole to which they belong. Indeed, the truth value of the sentence is determined not by its sense alone but by its sense plus the way things stand with certain elements in the real world. The second job of the parts of the sentence is, then, to determine what those truth-relevant elements are. Both the sense and the significance of the whole sentence "emerge" from that of the parts.

Frege's earlier semantic framework had provided the structure for some of the deepest foundational work of the century – including his theory of quantification and his analysis of arithmetic. Frege's philosophy after 1890 is the record of his remarkable but only partial success at grounding those foundational achievements on this newfound dualism. To explore the difficulties he encountered in this effort would require a more detailed analysis than is appropriate here. "It isn't all gold, but there is gold in it," Frege had written to his son concerning his unpublished papers. His published writings may not be all gold either, but on semantic matters Frege had seen far deeper than any of his contemporaries and had reshaped the course of the semantic tradition.

5

Meaning and ontology

O doubtful names which are like the true names, what errors and anguish have you provoked among men!

From Book of Crates, in Bertholet, *La chimie au Moyen Age*, vol. 3

In the last two decades of the nineteenth century, the semantic tradition took a turn toward ontology that would alienate those empiricists who, of course, wanted to avoid idealism, but not at the price of Platonism. A variety of issues were involved, all having to do with whether knowledge is independent of what is known: Are the objects of knowledge mind-independent? Is what we say about them mind-independent? Are their properties and relations mind-independent? These are very different questions, but the growing bias toward semantic monism tended to conflate them.

These questions elicited two separate developments widely regarded as the landmarks of a certain type of realism. The first centered around the notions of intentionality and denoting; the second around the rights and wrongs of holism. The links between these doctrines and what came to be known as "logical atomism" is the topic of the following two chapters.

The purpose of most people involved in these developments was to oppose the growing tide of German idealism and neo-Kantianism. Worthy as the project was, it was marred by an excessive reliance on psychologistic semantic categories and by a damaging confusion concerning the subject–predicate form. The former affected the semanticists' theory of empirical representation, leading to untold confusion via the so-called problem of our knowledge of the external world. The latter, our main interest at this point, led to the conflation of semantics with Platonism.

The subject–predicate form has been blamed for inspiring a number of philosophical confusions. Perhaps the best known is the doctrine that there are no ultimately relational propositions (that all relations are "internal," as Russell put it). Russell identified the fallacy and its alleged links with idealism, and he talked more eloquently and permanently than anyone else about its dangers, which he certainly avoided. But there was a less celebrated confusion created by the tendency to think of subjects and predicates, one that Bolzano and Frege had quietly avoided but that inspired most of those who followed in Brentano's footsteps, Russell prominently included. It might be described as follows.

The notion of the "subject" of a proposition is ambiguous between the subject concept and the object(s) falling under it. The subject of *Adam is friendly,* for example, could be either the concept *Adam* or Adam himself, and Leibniz's famous predicate-in-the-subject doctrine trades on this ambiguity: Adam (but not *Adam*) must have the property of friendliness if the proposition in question is true. Taking 'subject' in its objectual sense, someone might think that all propositions must have a subject–predicate form, that there must be some thing that they are about and something they say about it. According to this picture, every basic statement can be analyzed as involving two properties, the subject concept and the predicate concept, respectively. It is essential to recognize that these concepts play very different semantic roles: The former is a mere instrument for identifying the topic or "subject" of the proposition; the latter is what is said about it. One can think of the subject concept as looking around until it finds something that exemplifies it; the latter is what the proposition talks about, its "logical subject" as Russell would call it. Once it has been found, the proposition is ready to convey its information by asserting of that object the predicate concept. In this picture of things, the penalty for vacuity of a concept depends on its location in the proposition. If the predicate concept has no instances, we have said something false; but if the subject concept has no instances, the price is far greater: We have said nothing. To put it differently, the presence of the subject concept in the proposition does not suffice to guarantee its meaningfulness; it must also have instances.

People tempted to think this way can be identified by a tendency to say things like the following: It is self-evident that every proposition must be about something; the claim that chimeras do not exist must be understood as saying about chimeras that they do not exist (and *not* about the concept of a chimera that it has no instances); quite generally, every representation and every belief must have an object other than itself; 'I am the subject of a proposition; therefore I am' is no less evident than Descartes's *Cogito.*[1] A primary character of thought, according to such persons, is its relation to objects, its aboutness. Their approach stands or falls with their understanding of the notions of *objects* and *aboutness.* In fact, it falls.

Neither Bolzano nor Frege felt the slightest temptation to reason along these lines. Frege, in particular, saw that there is no ambiguity in the notion of instantiation and analyzed existence as the having of instances. He went so far as to deny that 'All As are Bs' is about As *even* when there are As (which makes one wonder why we should examine their properties in order to see whether the claim is true). So very far was Frege from the temptation to believe in golden mountains. Others were not as lucky.

Misplaced objectivity

Five years before Frege's *Begriffsschrift* appeared, Brentano published a book destined to have a far greater influence in the nineteenth century, his *Psychology from an Empirical Standpoint.* Husserl, Stumpf, Meinong, and Twardowski drew lessons from his work that Brentano would later describe as absurd but that decisively influenced the course of twentieth-century Continental philosophy.

There are several intimate links between Brentano's work and the protoanalytic tradition that we have examined in earlier chapters. The most obvious are his hostility toward the Kantian-idealist movement and his interest in careful, patient, piecemeal intellectual work. There is also his emphasis on the metaphysical pre-Kantian, indeed, pre-Cartesian, classics, particularly evident in his attempt to link his idea of intentional in-existence with medieval themes and his attitude toward Aristotle. Finally, there is the fact that many of the main problems that concerned him were semantic.

There was a major difference, however: Brentano never quite came to grips with the fact that the problems in question were not psychological but semantic in nature, and therefore called for semantic solutions. The title of the book in which Brentano launched his new project is an accurate indication of the halfhearted, indecisive character of its break with Kant and his semantic ancestry. In fact, Brentano thought that logic itself derived its justification from psychology,[2] and he was greatly upset when his most distinguished disciple, Husserl, took Bolzano as his model.[3]

Brentano's most influential idea was stated in a celebrated and obscure paragraph of his *Psychology:*

> Every psychic phenomenon is characterized by what the Scholastics in the Middle Ages called the intentional (and sometimes the mental) in-existence of an object, and what we should like to call, although not quite unambiguously, reference (*Beziehung*) to a content (*Inhalt*), directedness (*Richtung*) toward an object (which in this context is not to be understood as something real), or immanent objectivity (*immanente Gegen-ständlichkeit*). Every psychical phenomenon contains something as its object, though not each in the same manner. In representation something is represented, in judgment something is acknowledged or rejected . . . in desiring it is desired, etc. (*Psychologie,* p. 115)

This intentional in-existence was for Brentano the characteristic trait of mental or psychic phenomena (perhaps roughly comparable to Kant's world of inner sense) that distinguished them from physical phenomena (those present to outer sense).

Brentano's principle of classification soon became so popular that two decades later Twardowski could start one of his books with "It is one of the best known positions of psychology, hardly contested by anyone, that every mental phenomenon intends an immanent object" (*On the Content and Object of Presentations*, p. 1). And Meinong, at about the same time, could assume, "The reader will concede without reservations that it is essential to everything psychic to have an object. . . . For no one doubts that we cannot have a representation without having something to represent, and likewise that we cannot judge without having something to judge" ("Über Gegenstände höherer Ordnung und deren Verhältnis zur inneren Wahrnehmung" [1899], *Gesamtausgabe*, vol. 2, p. 381).

But the widespread and ever-growing agreement on what Brentano had said was not matched by a corresponding agreement on what he meant. The best philosophers among his followers saw no way of accepting Brentano's premise without also accepting the view that the object that is the target of the intentional act is not a mere component of the act, but must enjoy a mind-independent form of being.[4]

The first step in the process leading to this remarkable conclusion was the elimination of Brentano's confusion between the content and the object of a representation. In a logic treatise written under Meinong's editorship, Höfler had noted that whereas the content of a representation is part of the subjective act of representation, the object, in general, is not (*Logik*, par. 6). Twardowski devoted a whole monograph to the topic, his influential *On the Content and Object of Presentations*.

A good way to introduce the content–object distinction, according to Twardowski, is to notice an ambiguity in the notion of the "represented" (*Vorgestelltes*). Following Brentano, he distinguished between determining and modifying uses of adjectives. 'Old friend' is an example of the former and 'false friend' of the latter: An old friend is a kind of friend; a false one is not. Twardowski's preferred example concerns a painting of a landscape. When we talk of the painted landscape, we might mean two different things, depending on whether 'painted' is intended in an attributive or in a modifying sense. In the former case, what is meant is the actual landscape that inspired the painter's artistic product; in the latter, what is meant is the painting itself. The same ambiguity obtains in the case of representations, for when we speak of "the represented" we might mean our subjective "painting," the content, or its target, the object. A person represents to himself an object, for example,

> a horse. In doing so, however, he represents to himself a mental content. The content is the copy of the horse in a sense similar to that in which the picture is the copy of the landscape. In representing to himself an object, a person represents to himself at the same time a content that is related to this object. (*On the Content and Object of Presentations*, p. 16)

Once content is distinguished from object and the former placed in the psychic act, the obvious next question is, What and where are the objects to which psychic acts refer?

Twardowski and Meinong, the best known of the first generation of Brentano's disciples, saw in his writings a revolutionary answer to that question. As they read him, Brentano had refuted Bolzano's contention that there can be representations without objects. Since all (and only) psychic phenomena have reference to an object, and since all representations are psychic phenomena, it follows that all representations must refer to objects. What these objects are depends on whether the representation is empirical. *This rose* (thought in the process of looking at a rose) is an example of an empirical representation; representations for which there is no empirical object in sight to attach to them, such as *the golden mountain* or *the round square*, are nonempirical. It is the latter kind that concerns us at this stage, but first it is worth saying a word about the former.

The limited character of the rebellion against psychologism in the Austrian realist movement is illustrated by its version of the object of perceptual representation. Instead of saying that the object of a judgment of perception such as *This is a rose* (said while looking at a rose) is the rose in question, these semanticists accepted the idealist's notion that judgments involve phenomenal elements such as colors and sounds, and then reified those categories. Semanticists and idealists agree that the judgment concerns not roses in any ordinary sense but phenomenal colors, tactile sensations, and the like. But idealists, reasonably enough, can make no sense of phenomenal qualities in a mind-independent world; semanticists say they can. For all their virtues of rigor and insight, Brentano's writings must be apportioned much of the blame for the sense-datum approach to epistemology.

If in the field of empirical representation Austrian realism reified psychologism, in the field of representations without empirical objects it ontologized semantics. Where Brentano's sense of proportion had tempered his bold premises with a degree of hesitation and ambiguity, his disciples relentlessly pursued and embraced the consequences of those bold premises. They did what disciples tend to do: They reduced his system to absurdity. Twardowski's writings are perhaps the first to display with full clarity the problems that arise from a wholehearted endorsement of Brentano's hesitant views on representation.

Twardowski argued that, appearances notwithstanding, no representation, not even the inconsistent ones, can be without a corresponding object. Consider the worst possible candidate, *the round square*. Superficially one might think there is no such thing. Yet the fact that sentences such as 'The round square does not exist' are meaningful entails that, in some sense, there must be a round square; we must be talking about something when we deny the existence of the object in question. One could hardly ask for a clearer example of the subject–predicate fallacy described early in this chapter and subsequently reproduced by Meinong, Moore, and Russell.[5]

At this point emerged a familiar jungle: The objects of representation, according to Twardowski, may exist or not exist and be real or unreal. A (phenomenal) tree, a grief, an act of representation are real; an absence, a possibility, and the content of a representation are not. But what is unreal may exist: We may say truly (if somewhat Germanically) that there exists a lack of money; hence, an unreal thing must exist (*On the Content and Object of Presentations,* pp. 33–4).

Austrian realism, the doctrine that objectuality is a necessary condition for realism, soon found an echo in England. It is hard to judge the extent to which Moore's and Russell's shift from idealism to Austrian realism was influenced by the Brentano school. There is no question, however, that both Moore and Russell were familiar with and admired the achievements of that movement.[6]

Moore's and Russell's rebellion against idealism in 1898 started with what Russell described as the "distinction between act and object in our apprehending of things" (*The Problems of Philosophy,* p. 42).[7] Moore's "The Refutation of Idealism" of 1903 was based largely on the idea that for all the lip service idealists pay to the content–object distinction, they treat the object as immanent to the representation. And he drew with more sharpness and emphasis than Brentano the conclusion that phenomenal qualities (sense data) have a mind-independent life. Thanks to Moore it became commonplace in Anglo-Saxon philosophy to think that physical objects are invisible and that they are best construed as highly theoretical and bold inferences out of "the given."

Russell's early writings were less concerned with matters of empirical knowledge and displayed instead the consequences of the Austrian-realist doctrine on so-called objectless representations. In fact, one of the best-known statements of the Austrian-realist position on that issue occurs in Russell's *Principles,* in which he wrote:

> *Being* is that which belongs to every conceivable term, to every possible object of thought – in short to everything that can possibly occur in any proposition, true or false, and to all such propositions themselves. . . . "A is not" must always be either false or meaningless. For if A were nothing, it could not be said not to be; "A is not" implies that there is a term A whose being is denied, and hence that A is. Thus unless "A is not" be an empty sound, it must be false – whatever A may be, it certainly is. Numbers, the Homeric gods, relations, chimeras and four-dimensional spaces all have being, for if they were not entities of a kind, we could make no propositions about them. (p. 449)

Oddly enough, this passage and many others like it appeared in a book that also contained the refutation of the view that every representation must have an object (Russell's first theory of denoting). Before we examine (in Chapter 6) this peculiar episode and its sequel, we must begin by introducing Russell's basic semantic ideas and their link with realism.

The propositional complex

In "Meinong's Theory of Complexes and Assumptions" (1904), Russell stated a thesis that he had "been led to accept . . . by Mr. G. E. Moore" and that was, in fact, one of Meinong's main contentions: "that every presentation and every belief must have an object other than itself and . . . extra-mental" (*Essays in Analysis,* p. 21). We have just met the extra-mental objects of representation; let us turn to the objects of belief.

Like Bolzano, Russell had come to realize the importance of distinguishing between propositional attitudes that pertain to psychology and their targets, which do not. But unlike his illustrious predecessors, Russell never came to grips with the fact that meaning cannot be constituted out of fragments of the real world. The psychologistic alternative to semantic monism led him (around 1910) to a self-destructive semantic standpoint that had no room for the notion of what a statement says.

Russell's semantic picture circa 1900 was, roughly speaking, what remains of Frege's when the domain of sense is dismissed and that of meaning is extended in a natural manner. Whereas Frege's semantic dualism associated with each grammatical unit two different semantic elements, a sense and a meaning, Russell always (except for a rather brief interlude between 1902 and 1905) associated a unique semantic object with each grammatical unit; when such objects joined with others to constitute a proposition, he naturally called them the "meanings" of the corresponding expressions in the sentence.

Originally, the main function of Russell's meanings was that of Frege's senses, to act as the building blocks out of which the propositional complex was constituted. Thus, Frege and Russell were in verbal agreement when they said that the proposition expressed by 'John loves Mary' has as constituents the semantic correlates (meanings, senses) of 'John', '*x* loves *y*', and 'Mary'. But Russell's meanings were not at all like Frege's senses; in fact, they were hard to distinguish from Frege's meanings (*Bedeutungen*). If 'John' and 'Mary' are proper names, their Russellian meanings are John and Mary, and the Russellian meaning of 'loves' (or '*x* loves *y*') is the concept *love* or *x loves y,* which is for Russell something as much a part of the external world as John and Mary. Like Frege, Russell thought of properties and relations as unsaturated.[8]

The central difference between Frege's "realm of meanings" and Russell's is that Frege's concepts remain forever incomplete and unsaturated, whereas in Russell's more permissive semantics the concept's need for completion is rather frequently fulfilled. Russellian concepts are saturated with objects – eventually, with objects of the appropriate types. The outcome of this process is not an evaluation (Russell's concepts are not functions in Frege's sense), but another object – a whole that displays a peculiar sort of unity in which the concept and all of the saturating objects appear as constituents. This is the Russellian proposition or propositional *complex*.

From the very beginning, in what the idealists regarded as a decisive inconsistency, Russell insisted on the element of unity that semantic complexes generally exhibit. He was particularly impressed by Meinong's "careful attempt" to deal with the very delicate problem of the unity of the complex. Russell agreed that

> a melody of four notes is not a fifth note, and generally a complex is not formed by adding an object to the constituents . . . for red, green and difference do not make 'red differs from green'. . . . It is this special and apparently indefinable kind of unity which I should propose to employ in characterizing the notion of a complex. The kind of unity in question belongs, as is evident, to all propositions. ("Meinong's Theory of Complexes and Assumptions," *Essays in Analysis*, p. 28)

Thus, what we believe, assume, or deny when we believe, assume, or deny that red differs from green is a single thing: a complex, mind-independent entity that has as constituents the meanings of the words that make up the sentence in which we express the belief, assumption, or denial. For example, in his *Theory of Knowledge* Russell examined the difference between such phrases as 'Beggars are riders' and 'Beggars would be riders':

> We may now add the question and the imperative, "are beggars riders?" and "beggars shall be riders." In all these, the relation between beggars and riders is the same; but in the first it is asserted, in the second suggested as a consequence of a hypothesis, in the third the object of a doubt, and in the fourth the object of a volition. We should not say that these four phrases "have the same meaning," yet they all have something very important in common. The word "proposition" is a natural one to use for expressing what they all have in common: we may say that they express different attitudes towards the same "proposition." (p. 107)

Compare this view with one that was part of Frege's early semantics. In *Begriffsschrift* Frege had explained that his assertion sign included symbols for two quite different operations:

> *The horizontal stroke* that is part of the sign $\vdash$ *combines the signs that follow it into a totality, and the affirmation expressed by the vertical stroke at the left end of the horizontal one refers to this totality.* Let us call the horizontal stroke the *content stroke* and the vertical stroke the *judgment stroke.* (van Heijenoort, *From Frege to Gödel*, p. 12)

By way of example Frege considered the judgment "Opposite magnetic poles attract each other," which we may represent as

$$\vdash A.$$

The expression

$$-A$$

> will not express this judgment; it is to produce in the reader merely the idea of the mutual attraction of opposite magnetic poles, say in order to derive consequences from it and to test by means of these whether the thought is correct. When the vertical stroke is omitted, we express ourselves *paraphrastically,* using the words 'the circumstance that' or 'the proposition that'. (van Heijenoort, *From Frege to Gödel,* p. 11)

Thus, if 'a' means an object and 'f' a concept,

$$\vdash f(a)$$

symbolizes the judgment that a is f, whereas

$$-f(a)$$

symbolizes the semantic content of that judgment. To make this distinction clearer, Frege considered a language in which a proposition like 'Opposite magnetic poles attract each other' is expressed as 'That opposite magnetic poles attract each other is a fact'.

> To be sure, one can distinguish between subject and predicate here too, if one wishes to do so, but the subject contains the whole content, and the predicate serves only to turn the content into a judgment. *Such a language would have only a single predicate for all judgments, namely, 'is a fact'. . . . Our ideography is a language of this sort, and in it the sign* $\vdash$ *is the common predicate for all judgments.* (van Heijenoort, *From Frege to Gödel,* pp. 12–13)

According to this, a proper translation of $\vdash f(a)$ would be

It is a fact that a is f,

or

$\vdash$ that a is f.

Frege's purpose was the same as Russell's: to detach a propositional attitude (in this, assertion) from its target. But in Frege's case the target is a complex of elements, none of which could be regarded as part of the real world. In Russell's case, matters were otherwise.

Consider, for example, the proposition that Mont Blanc is more than four thousand meters high. We may not be surprised to find the concept *height* in it. But when Russell explained to Frege his theory of the proposition in 1904, Frege was shocked to find Mont Blanc among its constituents. In his response, Frege explained in a friendly way that

> Mont Blanc itself, with all its snowfields, is not a constituent of the thought (*Gedanke*) that Mont Blanc is more than 4000 meters high. . . . The sense of the word 'moon' is a constituent of the thought that the moon is smaller than the earth. The moon itself (i.e., the meaning of the word 'moon') is not part of the sense of the word 'moon', for then it would also be a constituent of that thought. (*Wiss. Briefwechsel,* p. 245)

To this Russell replied: "I believe that Mont Blanc itself, in spite of all of its snowfields, is a constituent of what is strictly speaking asserted through the sentence 'Mont Blanc is more than 4000 meters high'" (Frege, *Wiss. Briefwechsel,* p. 250). What is Mont Blanc, snowfields and all, doing there?

In Russell's day (as in ours) there were many people who could make no sense of a question about the existence or reality of *X*s unless this was meant as a question concerning much larger contexts involving *X*. Russell was certainly not among them. He had a criterion of existence that was entirely independent of linguistic, propositional, or any other contexts: acquaintance. Russell was (often) convinced that there are many things we are not acquainted with; our knowledge of them is inferential and speculative. But he clearly thought that acquaintance was the touchstone of reality, so that the failure of acquaintance for *every* element of a category (e.g., classes or propositions) should count as strong evidence against the ontological legitimacy of the category. Under the circumstances, a natural approach to the question of propositions was to ask what it is that we are sometimes acquainted with, that we may believe and disbelieve, accept or reject, assume or question.

Moore had an answer. Propositions, he explained, are just plain facts. The proposition *Brutus killed Caesar* is simply the fact that Brutus killed Caesar, a circumstance with which at least Brutus and Caesar were acquainted. And if we think of this fact as a complex, it is natural to think also that Brutus and Caesar are themselves constituents of it. Since the fact in some obvious sense also involves the relation of *killing,* we may also say that *it* is yet another constituent of the fact. Now we have all the elements that Russell thought we should encounter in the proposition *Brutus killed Caesar.*

From this standpoint, what we say (when true) is no different from

what happens. Therefore, in his article "Truth and Falsity" for Baldwin's *Dictionary of Philosophy and Psychology,* Moore argued against the "common supposition" that the truth of a proposition consists in some relation of correspondence that it bears to reality:

> It is essential to the [correspondence] theory that a truth should differ in some specific way from the reality, in relation to which its truth is to consist. . . . It is the impossibility of finding any such difference between a truth and the reality to which it is supposed to correspond which refutes the theory. (p. 717)

Moore wondered, what could this difference be? Clearly displaying the blind spot for the domain of senses that was also characteristic of the Austrian-realist movement, Moore saw no more than two possible answers: (a) Propositions are, unlike facts, linguistic objects, that is, sentences, and (b) propositions are, unlike facts, mental copies of reality. He rightly objected that neither of these solutions would do, for the sentence merely derives its truth ground from something else, what it signifies, while the second alternative is based on a confusion between the content of a belief and its object. Thus, he concluded, "It seems plain that a truth differs in no respect from the reality to which it was supposed merely to correspond" (p. 717).

Russell thought this was very convincing, except for one small point that kept growing as time went by: If a true proposition is just the fact thought to make it true, what then is a false proposition? If truths are facts, what are falsehoods? What is it that we believe when we believe that Caesar killed Brutus? As we shall see (Chapter 8), this question would lead, by 1910, to the conclusion that there are no propositions – hence to the self-destruction of Russellian semantics.

The particular version of semantic monism endorsed by Moore and Russell was motivated primarily by a desire to avoid psychologism at all costs and by an inability to see that in semantic matters there could be anything between our subjective representations and the world.Their preferencr for realism required that propositions and their constituents be things we represent rather than representations; their blindness to the domain of Fregean sense forced the propositional constituents to be part of the ultimate furniture of the world. As it turned out, this drawback would be fatal to their realistic intentions. Like any other reasonable philosopher, Russell eventually felt the force of Kant's notion of constitution as applied to meanings; but his inability to detach meaning from the world turned that insight into another version of the old idealist project to construct the "external" world. Before looking at some of the elements in the long sequence of Russell's concessions to idealism, we should first examine briefly the nature of his original challenge to it.

Realism and holism

Few topics have elicited more heat and less light among philosophers during the past two centuries than the subject of realism. The developments we are in the process of reviewing are widely regarded as relevant to that issue. Russell and Moore are usually taken to be among the most extreme supporters of the realist position by those who challenge it. Yet Russell himself once said, "I do not regard the issue between realists and their opponents as a fundamental one" ("Logical Atomism" [1924]; *Russell's Logical Atomism,* p. 143). He may have come to think that to the extent that realism is defined as a doctrine about what is "real" or what "corresponds" to thought in the "external world," the hopeless obscurity of those phrases makes it impossible to formulate a doctrine on which it is worth taking a stand.

This difficulty is highlighted by the fact that realists and idealists can agree on virtually every verbal response they give to specific questions of reality: Russell once noted that Berkeley would surely agree that tables are real (*The Problems of Philosophy,* p. 15); Kant granted the independence of physical objects from mind and endorsed the correspondence theory of truth;[9] James endorsed the correspondence of thought and fact as the obvious essence of truth; and (as we shall see in Chapter 12) during his most intimate flirtation with phenomenalism, Carnap granted the existence and reality of every theoretical entity postulated by science. If these philosophers may legitimately talk like realists about reality and existence, what is it that divides them from the others?

Russell thought that "the fundamental doctrine of the realistic position, as I understand it, is the doctrine that relations are 'external'" ("The Basis of Realism," p. 158). And he added that "all arguments based on the contention that knowing makes a difference to what is known, or implies a community or interaction between knower and known, rest upon the internal view of relations" (p. 160). Elsewhere he explained:

> The question of relations is one of the most important that arise in philosophy, as most other issues turn on it: monism and pluralism; the question whether anything is wholly true except the whole of truth, or wholly real except the whole of reality; idealism and realism, in some of their forms; perhaps the very existence of philosophy as a subject distinct from science and possessing a method of its own. ("Logical Atomism," *Russell's Logical Atomism,* p. 154)

The heart of the logical atomism that Moore and Russell started to develop in 1898 was, in fact, a rejection of the doctrine of internal relations. What exactly were they denying, and what was its link with realism?

Russell's official version of the doctrine of internal relations described it as the unexciting view that relations are not independent, since they can be reduced to properties. Bradley had argued that every relation "essentially penetrates the being of its terms, and, in this sense, is intrinsical" (*Appearance and Reality,* p. 347), and Russell naturally took him to mean that all relations are like *love* in that they alter the terms they relate and consist, in the end, in features of those terms. The doctrine has obvious links with Leibniz's monadology, and these were amply explored in Russell's *Critical Exposition of the Philosophy of Leibniz,* in which he explained that faulty logic and an excessive reliance on the subject–predicate form are the basis of idealism. During the first decade of this century, much of Russell's philosophical activity seems to have been guided by this interpretation of the issue; in this period he devoted a great deal of time to exposing as absurd the idea that all terms have constituents that somehow correspond to the relations in which they stand.

If this were all there were to the doctrine of internal relations, both it and the realism that opposed it would deserve to be placed in the museum of thankfully forgotten ideas. But there is a second interpretation of the idealist rejection of independence that makes Russell look far less silly. It is not that relational statements are to be analyzed as of the subject–predicate form, but that both relational and subject–predicate statements are misunderstood when seen as composites built up from detachable, independent semantic parts and that the only sense these parts can make is as unsuccessful efforts to detach portions of some broader wholes.

The presence of this holistic (or, as Russell called it, "monistic") element in idealism is repeatedly acknowledged by Moore and Russell in their specific objections to British habits. For example, in a discussion of Joachim's *The Nature of Truth,* after noting with evident glee that monists call pluralists like him "crude," Russell added:

> The uninitiated might imagine that a whole is made up of parts, each of which is a genuine constituent of the whole, and is something on its own account. But this view is crude. The parts of a whole are not self-subsistent, and have no being except as parts. We can never enumerate parts *a, b, c, . . .* of a whole . . . the part *a* is not quite real. Thus W is a whole of parts all of which are not quite real. It follows that W is not quite really a whole of parts. ("On the Nature of Truth," p. 31)

In an exchange with Bradley in *Mind,* Russell formulated the main difference between them as follows: "I do not admit that [complexes] are not composed of their constituents; and what is more to the purpose, I do not admit that their constituents cannot be considered truly unless we remember that they are their constituents" ("Some Explanations in Reply to Mr. Bradley," p. 373). Bradley replied that "a relation apart from terms is to me unmeaning or self-destructive, and is an idea produced by an indefensible abstraction" ("Reply to Mr. Russell's Explanations," p. 76). A few years earlier Joachim had posed a key question: How does one have access to the constituents of the atomist's propositions?

Greenness is, for the theory [i.e., Russell's and Moore's theory] an ultimate entity in the nature of things, which has its being absolutely in itself. How, under these circumstances, greenness can yet sometimes so far depart from its sacred aloofness as to be apprehended (sensated or conceived); and how, when this takes place, the sensating or conceiving subject is assured that its immaculate *perseitas* is still preserved – these are questions to which apparently the only answer is the dogmatic reiteration of the supposed fact. ("The Nature of Truth," p. 42)

As we shall see in Chapter 7, the same question had recently been asked more artfully and powerfully by Poincaré, concerning geometric primitives. Russell's witty rejoinder to Joachim is worth recording:

Mr. Joachim alleges that the plain man is on his side. I have been tempted to ask some plain man what he thought greenness was, but have been restrained by the fear of being thought insane. Mr. Joachim, however, seems to have been bolder. Considering the difficulty of finding a really plain man nowadays, I presume he asked his scout, who apparently replied: "Well, sir, greenness is to me the name of a complex fact, the factors of which essentially and reciprocally determine one another. And if you, sir, choose to select one factor out of the complex, and to call it greenness, I will not dispute about the term, for I know my place, sir; but as thus isolated, your greenness is an abstraction, which emphatically, in itself and as such, is not *there* or *anywhere*." ("The Nature of Truth," p. 529)

Finally, we may note that in his "On the Nature of Truth," Russell noted the interconnectedness of two issues: the mind dependence of truth and the plurality of truths. And then he added that the latter question seemed to him "the more fundamental." The idealist version of monism entails "that nothing is wholly true except the whole truth, and that what seem to be isolated truths, such as $2 + 2 = 4$, are really only true in the sense that they form part of the system which is the whole truth" (p. 29). When statements are "artificially isolated they are bereft of aspects and relations which make them part of the whole truth, and are thus altered from what they are in the system" (p. 29). If we substitute references to meaning for references to truth in these remarks, we obtain a reasonably good formulation of semantic holism. The idealists themselves occasionally performed that substitution. Joachim, for example, explained that the whole truth is an organic unity or significant whole "such that all its constituent elements . . . reciprocally determine one another's being as contributory features in a single concrete meaning" ("The Nature of Truth," p. 66).

It would therefore appear that the difference between realists and idealists is that one group insists on the *independence* of certain entities

and the other rejects them as "indefensible" or "false" abstractions. The preceding remarks allow us to make a more definite statement concerning the type of independence involved. Consider a relational statement such as

(*) *a* is heavier than *b*.

Russell and Moore would say this asserts that

(**) The relation *is heavier than* holds between *a* and *b*.

Idealists (and, as we shall see, positivists in the 1920s) felt uncomfortable with this type of claim and usually refused to endorse it. But there is really no reason why they should not have. As the instances of Kant, James, and Carnap indicate, many of those who grant the idealists' basic points may accept these realistic-sounding utterances under an interpretation acceptable to them – and in later decades many would avail themselves of this possibility to endorse an idealistic philosophy and a realist rhetoric. How does the rhetoric differ from the reality?

The idealist who was the official target of Russell's atomism was, as we know, someone who explained both (**) and (*) as the attribution of certain properties to *a* and *b* – and on this debate Russell and Moore were clearly on the right side of the issue. The idealism we have seen reflected in our recent quotations does not disagree with atomism on how to analyze or reduce (*) however, but only on how to understand (**). For this second kind of idealist, (**) makes sense only insofar as it is taken to mean what (*) says. But for the atomist it is through (**) that we acquire a proper understanding of (*), at least in the sense that (**) gives us a better picture of what the claim is. Thus, both holists and atomists will agree in describing a particular situation as one in which the relation *being greater than* holds between two objects. But holists will explain this as no more than a restatement of the claim that one of those objects is greater than the other, whereas atomists will think they have uncovered an element, no less real than the objects under consideration, that is present "out there" and somehow gathers them into a particular type of factual or propositional unit. Notice that even this "atomistic" language could be endorsed by holists, provided that it were properly explained. The difference cannot be stated in any particular claim about what is or is not real, but emerges only when we consider what we might call the order of semantic explanation. (This point will be explained in greater detail in Chapter 12.)

Perhaps the best way to distinguish holists from atomists is to follow them in their divergent explanatory paths: Holists will explain the attribution of relations and properties to objects in terms of the truth of the sentences stating those facts, and then they will offer an account of the truth of those sentences that does not appeal to semantic features of their constituents. They will, for example, agree that ***heavier than*** may relate only physical objects. But they will read this as a peculiar statement of the point that 'is heavier than' makes sense only when flanked by names of physical objects; they will not be tempted to conjure images of an unsaturated entity anxious to grab two physical things in order to find semantic fulfillment. Atomists, in contrast, will take the explanatory path in the opposite direction. They will explain (*) through (**) and will then have to offer some noncircular account of how they have access to the semantic correlates of the parts of (**), its "indefinables." That is why, when asking about the meaning of indefinables, idealists may appeal to the sentence but atomists may not. Atomists will have to discover a prepropositional link with the constituents of propositions. This is the problem of the semantic pineal gland, raised by Joachim and Poincaré: How do we come to grasp the semantic units that lie at the basis of the atomists' construal of the propositional complex? It was on this very question that the atomist picture of knowledge came to grief in an episode that deserves separate treatment (Chapter 8).

Russell's semantic pineal gland was acquaintance. We understand sentences because we understand their constituent phrases, and we understand ***them*** because we stand in a very remarkable direct relation to the meaning of those terms: Those meanings are as present to us as an object was for Kant, when given in sensible intuition. Acquaintance is the unmoved mover of Russell's semantics, its ultimate explanatory factor, the crucial difference between his semantics and that of idealists, and, in the end, the cause of its failure. Acquaintance remained at the heart of Russell's philosophy for decades; yet in the years between his conversion to realism and his encounter with Wittgenstein, he slowly but constantly withdrew in the direction of an idealist semantics. Considerations relating to the paradox would soon lead Russell to argue that at least some, perhaps all, properties are "false abstractions" and that they may be properly treated only in broader propositional contexts.[10] Similar concessions will be examined in the next chapter. Most of them were guided by a discovery that Russell made while studying an aspect of the problem of intentionality, to which we now turn.

6

On denoting

Can we put the problem of philosophy thus? Let us write out all we think; then part of this will contain meaningless terms only there to connect (unify) the rest. I.e., some is there on its own account, the rest for the sake of the first. Which is that first, and how far does it extend?

Ramsey, Undated manuscript (ASP)

It is natural to think that the meaning of 'blue' or of 'the taste of a pineapple' is an entity in the world with which we are sometimes acquainted, a color or a taste. For such cases it seems explanatory to say that to know a meaning is to be acquainted with what is meant. The same seems to apply also to proper names such as 'Scott'; to know what is meant by them, in the "strictest" sense, is to be acquainted with those objects.

But this simple "museum" semantics does not readily extend to most other cases. A friendship and a promise are things that, in some sense, we can be witness to; but it no longer seems explanatory to say that to understand those expressions is to be acquainted with anything in particular. And what is it that we are acquainted with when after reading Jaeger's ***Paideia*** we have grasped the sense of that Greek notion? Indeed, the terms whose semantics are not plausibly explained through acquaintance are of both general and singular types. Prominent cases among the latter are Bolzano's the 'golden mountain' and Meinong's the 'round square'. We surely understand them; but, once again, what worldly element are we acquainted with as we grasp their meaning?

One might take these difficult cases as evidence that the museum theory of meaning is simply an error, that it is not valid even in those simple cases where it seems appealing. Or one might insist on the basic idea and look for ways to reinterpret the recalcitrant areas of semantics to make them fit into the semantic monist pattern. The episodes that we are about to review constitute the first crisis within the semantic monist movement, arising when Russell recognized that Twardowski's and Meinong's solutions were more than an honest person could take. Russell's solution was not a challenge to the monist framework, however, but simply an effort to rescue it from the immense implausibility into which it had been cast by its Austrian proponents. The device he invented in order to rescue that semantic picture from its well-deserved demise is what we shall call the "incomplete symbol strategy." Russell's problems would have vanished had he chosen to distinguish meaning from world, as Frege had. But nothing would move him from his conviction that the ultimate furniture of the world was also the ultimate furniture of semantics.

The discovery of denoting: knowledge by description

Kant's opinion that there can be no knowledge without a blend of concept and intuition was a variation on a long-standing theme that would occur, once again, in the theories of propositional knowledge developed in the semantic tradition. In semantic terms, the question Kant had addressed concerned the role played by intuition (or acquaintance) in those propositions or judgments that express our knowledge.

The overwhelming weight of tradition had been in favor of a very tight link between those two elements. Aristotle had said that in sensual knowledge the sense contains the sensual object without its matter, so that, as Brentano put it, "the object which is thought is in the thinking intellect" (*Psychology*, p. 88). Aquinas claimed that the object thought is intentionally in the thinking subject, and Brentano derived from this and related ideas his own doctrine of the intentional in-existence of the objects of knowledge in its acts. Leibniz had explained that to have an idea of an object it is not necessary to be actually thinking about it but only to have the ability or "faculty" to do so. Since we can be affected by objects of which we have no idea, this faculty must be something more than mere receptivity, and, Leibniz thought, it must involve our possession of a method to generate the represented object in a way "*which not merely leads me to the thing but also expresses it*" ("What Is an Idea?" [1678], *Philosophical Papers and Letters*, p. 207). Finally, he explained, *A* expresses *B* when "there are relations (*habitudines*) which correspond to the relations of the thing expressed"; for example, "the projective delineation on a plane expresses a solid, speech expresses thoughts and truths" (p. 207).[1] Hume, in turn, believed that "it is impossible for us to *think* of any thing, which we have not antecedently *felt*" (*Enquiry*, sec. 7, pt. 1, p. 41). In Chapter 1 we saw Kant claiming that in order to represent, a representation had to be (in effect) isomorphic with its object.

Marty and Twardowski examined at length Brentano's idea of "improper representation" in order to establish that in the end Bolzano had been wrong in denying that "there corresponds to every part of the content of a representation a certain part of the object which is represented through it" (Twardowski, *On the Content and Object of Presentations*, p. 88).[2] In addition to the principle of the intentionality of consciousness, phenomenologists introduced the "principle of phenomenological accessibility" (Becker, *Mathematische Existenz*, p. 502), according to which no state of affairs is thinkable that would be in principle unknowable, in the sense of not being presentable in intuition. Finally, as we shall soon see, Wittgenstein argued that the form of an object must be contained in a proposition about it. One is not surprised to see him turn against Russell's doctrine of knowledge by description (*Philosophical Grammar*, pp. 163–71); this will be discussed in Chapter 13. In all of these cases it is being assumed that a singular representation can represent its object only if it satisfies a condition that makes knowledge by description virtually impossible; for it is required that the source and the target of the referential relation be in some respect identical – if not in content at least in structure or formally. Barring semantic miracles it would appear that this correspondence is possible only through the mediation of something like intuition or acquaintance.

This common assumption seemed reasonable so long as one did not try to spell out in detail how to make sense of standard knowledge claims upon its basis. But when, in the second half of the nineteenth century, semanticists turned to the task of explaining the character of mathematical knowledge, they found the assumption hard to live with.

One of the main problems examined in Husserl's *Philosophie der Arithmetik* was that even though mathematical knowledge is our clearest instance of sound knowledge, we can have accurate, reliable, (isomorphic) representation of its topic, numbers, only for the case of very small numbers. We can represent accurately, for example, a class of three objects by looking at or imagining three cats; but who can represent a chiliagon with similar accuracy? In view of this failure of representation, Husserl thought, one may well wonder how rigorous arithmetical knowledge is possible.

In order to solve this problem, Husserl used a distinction he had learned in Brentano's lectures, between "proper and improper or symbolic representations. I am indebted to [Brentano] for a deeper understanding of the eminent importance of improper representations for all of our psychical life, a point which no one before him had fully grasped to the best of my knowledge" (*Philosophie der Arithmetik*, p. 193). In agreement with this distinction, content can be given to us in two ways, directly or indirectly. When we see a house, for example, a content is given directly to us, and we then say that we have a proper representation of the house. But when the content is given through a description like 'the house that is on such-and-such corner', what we have is a symbolic representation.[3]

A few years later, in his "Psychologische Studien zur elementaren Logik" (1894), Husserl elaborated this distinction:

There is a distinction between those representations that are intuitions and those that are not. Certain psychic experiences, known in general as "representations," have the peculiarity that they do not include their "objects" as immanent content (hence, not in actual consciousness) but *merely intend* the object in a sense that is still to be characterized more precisely. For the time being the following definition, which is obviously suitable but deliberately too complicated, will suffice. "Mere intending" means: to aim at a content not given in consciousness by means of one that is so given. (*Aufsätze und Rezensionen (1890–1910),* p. 107)

One can see Brentano's hand in the hesitation between content and object that lends a characteristic air of ambiguity to Husserl's posture on the issue between realism and idealism. The point here, however, is that Husserl was distinguishing between intuitions that "actually embrace within themselves" the objects (or perhaps the contents) that they intend and a second strange category of improper, symbolic representations that do not embrace their contents or objects but point to them, as it were, from afar or, to coin a word, "denote" them.

Husserl's solution to the problem of arithmetical knowledge was to say that it often relies on symbolic representation. A large and remarkably obscure part of *Philosophie der Arithmetik* is devoted to an explanation of what these representations are and how they work; but the whole explanation collapses as soon as one distinguishes between psychology and logic, as Husserl would endeavor to do a few years later. In a retrospective essay written around 1913, Husserl referred with scorn to "the customary appeal in the Brentano school to [improper] representation (*uneigentliches Vorstellen*)," which, he now said, "could not help. That was only a phrase in the place of a solution" (*Introduction to the Logical Investigations,* p. 35). Frege had told that much to Husserl in 1894, in his review of the *Philosophie der Arithmetik* (*Collected Papers,* pp. 195–209). And by 1902 Russell had come to the same conclusion.

The analogy between Russell and Husserl's problems circa 1900 is quite striking. To begin with, mathematics provided in both cases the original stock of problems that led them to what both called "logical" investigations. Next, within mathematics itself, the foundational problems that initially seemed to trouble them most involved the representation of numbers and, in particular, the difficulty of representing the very large ones. As we are about to see, one of the central problems in Russell's early philosophy of mathematics was to account for our ability to have knowledge concerning infinite numbers, especially that knowledge recently uncovered by Cantor in his theory of the transfinite. Finally, in both cases the solutions offered appealed to a form of representation that was very unlike what it represented.

The differences between Russell and Husserl are, of course, equally marked. The most prominent among them is, perhaps, that Russell's strategy for dealing with his foundational problems was, from the beginning, to focus on the target of all epistemic attitudes, the proposition.

It is Russell's picture of the proposition at this time (circa 1900) that helps explain why *he* thought there was a major problem concerning our knowledge about infinity; for he had tacitly assumed that in order for a proposition to be about a particular entity, the entity itself had to be a constituent of the proposition. This doctrine of *confined aboutness,* as we shall call it, is only an extreme form of the intimate link between intuition and propositional knowledge that was described earlier in this section. It isn't hard to see how knowledge of infinity might become a problem on the assumption of confined aboutness. Consider, for example, the proposition that every natural number is either odd or even. How can we know this, Russell wondered, given that it is a proposition about an infinitely complex object and given also that our finite minds can understand only objects of finite complexity?

We know that Russell (like Brentano and Husserl) had carefully read Bolzano's *Paradoxes of the Infinite* and that he was struck by Bolzano's account of classes. According to Bolzano, classes are to be understood in terms of the concept of conjunction. To say that John and Mary are two, for example, is to assign a certain property to the class *John and Mary.* This class is a rather peculiar entity, a conjunction not of propositions but of objects. But Russell soon discovered that this account posed a problem for the case of infinite classes;[4] for if an infinite class is an infinite conjunction, it would appear that a proposition about infinite classes must be infinitely complex and hence beyond human comprehension. That is why the "theory of *and* applies practically [i.e., to human beings, though not to gods and angels] only to finite numbers" (*Principles,* p. 134).

At this point Russell turned to language for a clue. How do we, in fact, talk about infinite classes? Certainly not by means of infinitely long conjunctions. In fact, we frequently do so by means of quantification, as in 'All numbers are even or odd'. Perhaps, he thought, the solution to the problem of propositions about infinity lay in a correct theory of quantification. Thus, in 1900 Russell wrote to Moore:

Have you ever considered the meaning of *any?* I find it to be the fundamental problem for mathematical philosophy. E.g., 'Any number is less by one than another number'. Here *any number* cannot be a new concept, distinct from the particular numbers, for only these fulfill the above proposition. But can *any number* be an infinite disjunction? And, if so, what is the ground of the proposition? The problem is the general one as to what is meant by any member of any defined class. I have tried many theories without success. (Letter to Moore of 16 August 1900, Russell Archives)

Standard logic, Russell thought, was not much help. In his opinion, logicians from Bradley to Frege had given far too much importance to the role of concepts in quantification and had failed to determine what quantified propositions are *about*. Russell thought it self-evident that they are about the objects over which one quantifies and not about the correlated concepts (*Principles*, p. 90), since otherwise they would convey no knowledge about the objects in question.

Russell looked at the problem of quantification in the typical pre-Fregean fashion. For him, quantifier words such as 'all' and 'some' were to be interpreted not in isolation but in connection with the attached predicate. Whereas for Frege 'all' and 'some' were names of two specific second-order concepts, for Russell they had no meaning in isolation and were to be treated as fragments of larger grammatical units such as 'all men' and 'some men'. *These* were the grammatical units whose semantic analysis was "the fundamental problem of mathematical philosophy" in 1900. Russell's solution to that problem was his first theory of quantification, better known as his first theory of denoting. It disclosed, he thought, "the inmost secret of our power to deal with infinity" (*Principles*, p. 73).

As developed in the early drafts of *The Principles of Mathematics* (written about 1901), the doctrine of denoting is the theory of five words: 'all', 'every', 'any', 'some', and 'a', all of them quantifier words.[5] Definite descriptions were included almost as an afterthought, in the final stages of the manuscript, and there is reason to think that this thought was rather brief and uncommitted, as we shall see. What Russell had come to see in 1901–2 was that perhaps he had been wrong when he told Moore that *any number* cannot be a concept. Perhaps the reason we can deal with infinity in propositions of finite complexity is that quantifier expressions signify the presence of a very peculiar sort of concept, one that, unlike "normal" concepts, does not just sit there in the proposition, either linking its partners into a unified complex or allowing itself to be the focus of referential attention. Instead, these "denoting concepts," as Russell called them, play an altruistic semantic role in that they somehow manage to refer to (or, as he sometimes puts it, "indicate") other objects, thus allowing the proposition to be about things other than its constituents. Here is Russell's introduction of his new and revolutionary idea:

> The notion of denoting, like most of the notions of logic, has been obscured hitherto by an undue admixture of psychology. There is a sense in which *we* denote, when we point or describe, or employ words as symbols for concepts. . . . But the fact that description is possible – that we are able, by the employment of concepts, to designate a thing which is not a concept – is due to a logical relation between some concepts and some terms, in virtue of which such concepts inherently and logically *denote* such terms. It is this sense of denoting which is here in question. . . . A concept *denotes* when, if it occurs in a proposition, the proposition is not *about* the concept, but about a term connected in a certain peculiar way with the concept. If I say "I met a man," the proposition is not about *a man*: this is a concept which does not walk the streets, but lives in the shadowy limbo of the logic-books. What I met was a thing, not a concept, an actual man with a tailor and a bank-account or a public-house and a drunken wife. (*Principles*, p. 53)

At this point the stage was set for one of Russell's most celebrated and important distinctions, that between knowledge by acquaintance and knowledge by description. Even though, as Russell still thought, we can understand a proposition only when we are acquainted with all of its constituents, we can not only understand but also know propositions that are *about* things with which we are not acquainted. If I can have knowledge of other minds, of unobservable particles and abstract entities beyond acquaintance, it must be because something I am acquainted with can make some propositions refer to such things. "An object," Russell explains,

> may be *described* by means of terms which lie within our experience, and the proposition that there is an object answering to this description is then one composed wholly of experienced constituents. It is therefore [*sic*] possible to know the truth of this proposition without passing outside experience. If it appears upon examination that no *experienced* object answers to this description, the conclusion follows that there are objects not experienced. ("On the Nature of Acquaintance" [1914], p. 161)

It is easy to underestimate the extent to which this Russellian doctrine conflicts with traditional semantic dogma. Russell himself was, as usual, on both sides of the issue. But most of those who had preceded him (and not a few of those who followed him) were consistently on the wrong side of it. It is clear, of course, that Russell's theory of denoting conflicts with the thesis of confined aboutness as well as with the less specific claims concerning the link between knowledge and intuition to which we earlier alluded. A denoting concept, Russell explained, need in no way resemble what it denotes. An object of infinite complexity (such as an infinite class-as-many) may be denoted by a concept of complexity modest enough to be easily grasped by a human being:

> With regard to infinite classes, say the class of numbers, it is to be observed that the concept *all numbers*, though not itself infinitely complex, yet denotes an infinitely complex object [the class of as many numbers]. . . . infinite collections, owing to the notion of denoting, can be manipulated without introducing any concepts of infinite complexity. (*Principles*, p. 73)

It is worth emphasizing that in the theory of denotation developed in *Principles*, denoting concepts have no obligation to denote: "The denoting concepts associated with [the concept] *a* will not denote anything when and only when '*x* is an *a*' is false for all values of *x*. This is a complete definition of a denoting concept which does not denote anything" (p. 74). Thus, Russell placed himself once again with Bolzano, rather than the Austrian realists, on the issue of whether there can be representations without objects. We are therefore not surprised that he explained to Meinong in 1904 that "in such a case as that of the golden mountain or the round square one must distinguish between *sense* and *reference* (to use Frege's terms): the sense is an object, and has being; the reference, however, is not an object" (Letter to Meinong, p. 16). Nor are we surprised to read in an article written a few months before the doctrine of "On Denoting" was conceived that *the present king of France* is a "complex concept denoting nothing. The phrase intends to point out an individual, but fails to do so: it does not point out an unreal individual, but no individual at all" ("Existential Import of Propositions" [1905], *Essays in Analysis,* p. 100; see also Review of A. Meinong [1905], *Essays in Analysis,* p. 81). Thus, the new theory of denoting did a great deal more than provide a solution to the problem of our knowledge of infinity; it completely undermined the Austrian-realist approach to semantics by eliminating the implicit confusion between what must have being in order for a claim to make sense and what that claim is about. Finally, the recognition of denoting concepts was inconsistent with monism and with the principle of confined aboutness. Now, some expressions, denoting symbols, have *two* semantic dimensions equally worthy of the logician's attention: a "meaning," that is, the object they contribute to the proposition, and a denotation. In 'All numbers are finite', for example, 'all numbers' has as its meaning the denoting concept *all numbers,* while it (or its concept) denotes the class of numbers, which the proposition is about. Similarly, both *Scott is tall* and *The author of Waverley is tall* are about the author of *Waverley,* that is, Scott; but only the first has him as a constituent (on the assumption that 'Scott' is a proper name).

But then, what are we to make of the persistence of confined aboutness throughout Russell's later work and of the Austrian-realist pronouncements so prominent in Russell's *Principles?* Whitehead is said to have commented that Russell was an entire Platonic dialogue in himself. The point was not that Russell's opinions changed through the years in dialectical fashion; this would not have been worth mentioning. Like every great philosopher, Russell felt the force of conflicting intuitions. Unlike most philosophers, however, he succumbed to those temptations without much regard for consistency. It has often been noticed that Russell gallantly and frequently refuted his earlier theories; and this is admirable. But it has not been noticed often enough that he didn't always discard the theories he had refuted. His views on aboutness are a case in point; those on denoting are another.[6]

Consider, for example, this question: What would the author of chapter 5 of *Principles* ("Denoting") have to say concerning the statement printed only pages away (p. 449), quoted at the end of section 1 of the preceding chapter? It suffices to examine a proposition of the form '*A* is not', where *A* is a denoting expression – for example, 'The golden mountain is not'. Russell's argument with this particular case would go like this: 'The golden mountain is not' must always be either false or meaningless. If the golden mountain were nothing, it could not be said not to be; 'The golden mountain is not' implies that there is a term 'the golden mountain' whose being is denied and hence that the golden mountain is. Thus, unless 'The golden mountain is not' is an empty sound, it must be false – whatever the golden mountain may be, it certainly is. The golden mountain, the Homeric gods, and so on must have being, for if they were not entities of a kind, we could make no propositions about them.

These are precisely the sorts of things one would have expected Brentano's disciples to say; and one would have expected the author of chapter 5 of *Principles* to complain that this was a muddleheaded piece of reasoning. It could readily be granted that 'The golden mountain is not' would be an empty sign if 'the golden mountain' did not contribute a meaning to the corresponding proposition. But, of course, it does, regardless of whether it has a denotation or not. In general, the presence or absence of a denotation, that is, of "the logical subject," has nothing whatever to do with whether the statement in question is a mere noise or expresses a meaningful proposition.[7]

The conclusion to be drawn from this remarkable situation is that Russell's understanding of denoting at the time of *Principles* was unstable. It was *so* unstable that after putting forth a theory of the matter far better than those of his Austrian counterparts, Russell then paid scant attention to its implications.

Russell's second theory of denoting

Let us start by reviewing briefly the main elements of Russell's famous second theory of denoting symbols. Russell claimed that there are no denoting concepts, thus moving back to semantic monism. There is no particular entity, conceptual or otherwise, no "meaning," that denoting expressions contribute to a proposition. For example, definite descriptions have no meaning even though they do have (as in the old theory) "only sometimes a *denotation*" ("On Denoting" [1905], *Essays in Analysis,* p. 108). To understand the semantic character of a definite description is not to identify a certain thing, its meaning, but to give for each sentence in which it occurs a translation into another sentence in which it does not occur. In a language Russell would not have used, we could say that to "define" a denoting symbol is to give the truth conditions of all sentences in which it occurs. For the case of definite descriptions, the rule offered by Russell in "On Denoting" was this: '*F*(the ϕ)' is to be interpreted as 'Exactly one thing is ϕ and that thing is also *F*'. Thus, '*F*(the ϕ)' is not of the form '*F*(*x*)' but of the much more complex form indicated by the translation. The phrase 'the ϕ' does not contribute any simple or complex constituent to *F(the* ϕ*)*, or, to put it differently, it has no meaning. It may, however, have a denotation; indeed, it will have one precisely when there is just one ϕ (pp. 114, 108).

Russell's theory can be seen as one more step along the path initiated by Frege's theories of quantification and of number statements. The heart of Russell's theory is the observation that sentences involving 'the *F*' are very close relatives of number statements; for what 'The *F* is *G*' says is that the number of *F*s is one (a number statement) and that all of them are *G*. All one need add to derive Russell's theory is Frege's analyses of number statements and of quantification. The result is that 'The *F* is *G*' is not about the *F* (or even about *the F*) but about the concept *Fa* (and the concept *G* as well), and what we *really* say is that there is an *F* and no other (the number statement) and that *F* & *G* has instances. In the Fregean spirit, what appeared to be a statement in subject–predicate form is displayed as actually being about concepts and whether or not they are instantiated. Also in the Fregean spirit, the analysis dissolves any temptation to associate either the meaningfulness or the objectivity of the analyzed claim with the ontological features of what it might refer to.[8]

But why do we need a new theory of denoting? Russell said that the evidence for his new theory "is derived from the difficulties which seem unavoidable if we regard denoting phrases as standing for genuine constituents of the propositions in whose verbal expressions they occur" ("On Denoting," *Essays in Analysis*, p. 107).

The theories Russell aimed to refute are those that (like his earlier one) view denoting phrases "as standing for genuine constituents." He discussed two such theories: Meinong's and Frege's. Meinong assumed that an object of some sort is associated with every definite description in a way that, as Russell argued, entails the denial of the principle of contradiction. The Frege-type theory (closest to Russell's first theory of denotation) was subject to the following objection.

Consider "the cases in which the denotation appears to be absent." The "problem" is this: "If we say 'the King of England is bald', that is, it would seem, not a statement about the complex *meaning* 'the King of England', but about the actual man denoted by the meaning" (p. 108). But the same should be true of 'The King of France is bald', which "ought to be about the denotation of the phrase 'the King of France'" (p. 108). Since there is no such denotation, however, this seems to turn the proposition into nonsense (a clear echo of the Brentanophile statements of *Principles*). "Now," Russell argued,

> it is plain that such propositions do *not* become nonsense merely because [their denoting symbols do not denote]. . . . Thus we must either provide a denotation in cases in which it is at first sight absent, or we must abandon the view that the denotation is what is concerned in propositions which contain denoting phrases. (p. 109)

The former course had been advocated by Meinong and, in a different way, by Frege. Russell added, "The latter is the course that I advocate" (p. 109).

It is clear that this is no objection at all, for it simply assumes, without argument, what a Frege-type theory explicitly denies: namely, semantic monism and the doctrine of confined aboutness.[9]

Nevertheless, although Russell had no real argument for his new theory of denoting, there can be no doubt that its outcome was one of the most interesting and influential philosophical ideas of this century: the strategy of incomplete symbols. This strategy is what justifies Ramsey's claim that "On Denoting" is, after all, a paradigm of philosophy.

The incomplete-symbol strategy

A major feature of "On Denoting" is its sharp recognition of the intransparency of language and its presentation of a definite program for the elimination of this defect. One way to put this is to say, with Wittgenstein, that Russell had discovered that "the apparent logical form of the proposition need not be its real logical form" (*Tractatus*, 4.0031). That observation had been articulated much earlier and better by Frege, concerning the domain of sense, but his views had been largely ignored. What was new with Russell was the discovery of a strategy for neutralizing the semantic temptation to produce what Russell now called (echoing the idealists) "false abstractions." The application of this strategy to the case of definite descriptions in "On Denoting" was far less significant than the strategy. It would soon be applied by Russell to other domains (classes and propositions) and would, quite generally, preside over an orderly retreat in the direction of holistic idealism.

Russell's earlier writings displayed a strong inclination to assume that language is, as he once put it, "transparent," that is, that to each grammatical unit in the sentence there corresponds a (possibly complex) element, its meaning, which occurs as a constituent in the proposition it expresses. In spite of conscious efforts in *Principles* not to be misled by grammar, Russell's early atomism resembled Bolzano's in the unrecognized assumption that grammar is mostly a sound guide to semantics and ontology.

It is essential not to confuse the transparency of language that "On Denoting" was meant to deny for the case of definite descriptions with what we may call the *perfect-language thesis*, the doctrine that there is a subclass of sentences in (say) English, or a conveniently extended version of English,[10] that is powerful enough to convey all possible information and is perfectly transparent. Russell's response to the intransparency of language in "On Denoting" does not deny but, in fact, presupposes that there is a fragment of English that is both transparent and complete. The project launched in that paper was to neutralize the effect of the misleading fragments of language by identifying a (preferably effective) method to associate with *every* sentence in the given language another one in the transparent fragment of that language that expressed the very same proposition. This would be impossible, of course, if there weren't a transparent rendering of *every* proposition.

The main *technical* achievement of "On Denoting" was, as we saw, the particular translation rule intended to eliminate definite descriptions from all sentential contexts. We have already presented the basic idea in the preceding section. In a terminology closer to Russell's, we might put it as follows: Consider an arbitrary sentence S that contains one occurrence of a definite description 'the ϕ'. If we remove 'the ϕ' from S, we have a concept name 'C'. We can therefore interpret S as having 'the ϕ' as its grammatical subject and 'C' as its predicate. Then, Russell claimed, the transparent (or more transparent) version of what S says is 'There is exactly one ϕ and it is C'. Russell soon noted that this translation technique was not quite unambiguous and resolved the ambiguity only for contexts involving propositional attitudes. The purpose, however, was to present a translation method that applied to *all* sentences containing definite descriptions.[11]

Russell's conception of the link between a perfect-language sentence and its meaning can be depicted as follows: Each nonlogical word contributes an object, its meaning, to the proposition in question; and the proposition is the combination of all such meanings in a way or "form" that is itself the meaning of the framework of logical expressions involved. With this image in mind, we define a symbol as *incomplete* when it contributes no simple or complex constituent to the proposition it helps convey. Thus, if 'a' is incomplete, there is in the proposition fa no single entity, simple or complex, that counts as the contribution of 'a' to that proposition. Since the meaning of 'a' is its contribution to such propositions, if 'a' is incomplete, Russell says that it has no meaning.

Let us call a sentence transparent, or *perspicuous*, when its grammatical structure accurately reflects the structure of the proposition it expresses, so that to each grammatical unit in the sentence there corresponds an entity (its "meaning") in the proposition. Then it is clear that if a symbol is incomplete, the sentence in which it occurs is unperspicuous and therefore semantically misleading. If what we are after is a language that reflects perspicuously what we mean, a perfect language, we must detect and eliminate all incomplete expressions.

It is worth emphasizing the distinction among these three different, indeed, independent, tasks: (a) showing that an expression is incomplete, (b) providing a rule to translate all sentences containing the expression into (more) transparent versions, and (c) showing that incomplete expressions lack a denotation.

It is a widespread and dangerous error to confuse (a) and (b), that is, to think that the elimination procedure to which (b) alludes counts for Russell as a reason, perhaps even the decisive reason, for thinking that a symbol is incomplete. To show that a symbol is incomplete is to give some sort of argument to show that the symbol in question contributes no constituent to the appropriate propositions. Of course, it would be pointless to engage in (b) if one did not already have some reason to think that the symbol under consideration *was* incomplete. But it would be an error to saddle Russell with the blunder of confusing success in (b) with success in (a).[12] Russell recognized most of the time that the fact that a translation rule can be given to eliminate an expression from all contexts in no way entails that the corresponding proposition lacks a constituent contributed by the given expression: The translation might, after all, be less perspicuous than the original statement. It was in view of this fact that in "On Denoting," as well as in several other places, Russell put forth an argument designed to establish that there are no denoting concepts and, in particular, no meanings of definite descriptions. And when he applied the incomplete-symbol strategy first to classes (in 1906) and then to propositions (before 1910), he was in both cases led to do so by his discovery of reasons for thinking that the symbols for these alleged entities were, in fact, incomplete.

The argument that Russell offered to establish that definite descriptions are incomplete symbols is worth examining, since, inter alia, it offers a vivid illustration of the grip semantic monism had on Russell's thought, even at a time when he said he did not think there were propositions. Here is one of the many appearances of the famous argument:

> The central point of the theory of descriptions was that a phrase may contribute to the meaning of a sentence without having any meaning at all in isolation. Of this, in the case of descriptions, there is precise proof: If 'the author of *Waverley*' meant anything other than 'Scott', 'Scott is the author of *Waverley*' would be false, which it is not. If 'the author of *Waverley*' meant 'Scott', 'Scott is the author of *Waverley*' would be a tautology, which it is not. Therefore, 'the author of *Waver-*

ley' means neither 'Scott' nor anything else – i.e. 'the author of *Waverley*' means nothing, Q.E.D." (*My Philosophical Development*, p. 85)[13]

In other words, if the meaning of 'the author of *Waverley*' is Scott, then the proposition *The author of Waverley is Scott* is the very same proposition as *Scott is Scott*, since the very same constituents are located at the very same places. Since the latter is a tautology, so is the former. If, however, 'the author of *Waverley*' means some other thing, say, the number 2, then the proposition *The author of Waverley is Scott* has as constituents the number 2, Scott, and identity, correlated in such a way as to say something surely false (that Scott is the number 2). Notice the tacit appeal to confined aboutness; for if the "other thing" meant by 'the author of *Waverley*' were the denoting concept *the author of Waverley*, the proposition would be false only if we thought that it was about that very concept, rather than about what it denotes. Russell's argument is, of course, untenable unless one believes in semantic monism and confined aboutness; but it is the argument to which he kept appealing as the proof that definite descriptions are incomplete symbols. His later applications of the incomplete-symbol strategy were based on what he took to be weaker arguments of an inductive nature. Yet in all cases, he first came to think that a category of expressions could not be interpreted the way he had hoped in the heyday of atomism, and then proceeded to cope with the difficulty by applying the strategy.

Quite generally, what this strategy offered Russell was a way to save face as the new facts his research uncovered seemed to lead him in the direction first urged by Bradley and his followers. Definite descriptions and denoting symbols in general, then (at least) some properties and relations, then all classes, and finally propositions themselves were things that Russell's atomism had recognized as independently subsistent entities in 1900, but that by 1910 were what Bradley would call "illegitimate abstractions." The reasons that led Russell down this path were ones that Bradley's colleagues would probably not even have been able to follow, let alone discover. Yet the fact remains that Russell's project, when pursued relentlessly, seemed to lead to its own destruction. The incomplete-symbol strategy allowed Russell to preserve an image of conservative reform within his early atomist framework. The ultimate furniture of the world was still his real goal, though as years went by, one could but wonder what, besides sense data, would remain in the wake of the ontological massacre Russell was perpetrating.

7

Logic in transition

> I have long struggled against the admission of ranges of values thereby of classes; but I have found no other possibility to provide a logical foundation for arithmetic. This question is: how are we to conceive of logical objects? And I have found no answer other than this: we conceive of them as extensions of concepts or, more generally, as ranges of values of functions . . . what other way is there?
>
> Frege to Russell, 28 July 1902

> The first task in discussing the foundations of (pure) mathematics is to make precise the distinction between it and other sciences, a task which in Principia Mathematica is surprisingly neglected.
>
> Ramsey, Undated manuscript (ASP)

Logicism and the foundational crisis

In 1900 Russell underwent the one event in his intellectual life that he was willing to characterize as a "revolution": He met Peano and was struck by the capacity of Peano's work to shed light on the philosophical nature of mathematics. It was at this time that Russell conceived one of his most fruitful ideas, the logicist project.

Peano had identified a notational system, or a cluster of concepts, that seemed to have enormous expressive power. The hope was that it could be used to express all of mathematics, and Peano's school had been working for years at rewriting different fragments of mathematics in their peculiar notation. Russell suggested that its basic concepts might be reduced to purely "logical" notions, in an as yet undisclosed sense of that word, and that perhaps all the assumptions one needed were those of logic, whatever they might be.

Logicism is often defined as the thesis that mathematics is reducible to logic. This is correct as long as one understands that at this early stage, mathematics was a reality and logic a project. In Russell's practice, at any rate, the logicist motto was less a doctrine than a regulative maxim, intended as much to provide a guide to characterizing logic as to clarifying mathematics. This link between logic and mathematics was destined to have momentous implications. Among those pertaining to philosophy, none was greater than the effect it had on the role played by classes in logic.

Classes had been a part of logic since the days of Aristotle. Logic had traditionally been concerned with what we say and its internal relations; classes had been useful devices for translating talk of content into more manageable extensional terms. Thus, it was widely taken for granted that to every notion or concept there corresponded an extension, a class of things, the actual (or perhaps possible) instances of the concept in question, and that every class was the extension of some notion.

Until the nineteenth century classes remained, for the most part, the private concern of logicians, and logicians' problems were the sort that led them to question seriously the preceding assumptions. In the nineteenth century, however, classes began to take on a life of their own. It all started when Cantor and Frege brought classes to bear upon arithmetic, in their characterizations of the concept of number. This led to results that undermined the philosophical picture from which they had emerged. The first blow to the philosophers' conception of classes came from Russell's paradox (first communicated to Frege in 1902, see note 3, this chapter), which was thought to establish that the presence of a concept or intension does not guarantee the presence of the corresponding class. The second, heavier blow came in 1904 when Zermelo focused attention on the role of the axiom of choice, an assumption that implicitly questioned the necessity of intensions to guarantee the presence of classes. By the end of the first decade of our century, set theory had become a discipline that had no recognizable link with its traditional logical neighbors: concepts, intensions, and meanings. It had also become a hypothetico-deductive discipline whose main purpose was not to find a priori or even merely true assumptions but to save the mathematical phenomena.

Those who chose to preserve their commitment to the idea that this was still logic had to undergo a rather drastic change of mind as to what logic is. Frege concluded that the logicist program was dead and offered nothing in its stead. Wittgenstein eventually denounced the new turn of events as the product of unqualified confusion. Russell steered a middle course of indecision – or conflicting decisions; he joined the intensionalists on the subject of classes but, in the end, gave in to the new conjecturalist conception of logic that the mathematical community seemed, no doubt unwittingly, determined to promote.

Class struggle

By 1900 Russell had already rebelled against idealism. For him classes were no longer the outcome of classification processes but self-subsistent things, thriving mind-independently and waiting for someone to get acquainted with them. On that ground Cantor had built a famous paradise, a doctrine that was for Russell one of the greatest achievements of the human mind. Within that paradise everything seemed perfect, but for one troubling exception: Cantor's theorem that for every class there is another of higher cardinality.

Cantor's proof was extremely simple: Assume that there is a one-to-one correspondence f between a set S and its power set PS. Consider now the subset C of S whose elements are precisely those x's that are not members of their f-values (i.e., the x's such that $x \in f(x)$). Of course, C must be an element of PS and must therefore be the f-value of something in S – call that thing c. In other words, $f(c) = C$. At this point Cantor asked a question: Does c belong to $f(c)$ (i.e., to C) or not? If it does, then (since c belongs to its f-value, by the definition of C) it doesn't. And if it doesn't, then (since c is one of those x's that satisfy the defining condition of C [i.e., $x \notin f(x)$]) it does. So it does iff it does not – a contradiction. Cantor concluded that there could not be an f such as the one he had assumed; and with very little extra machinery, readily granted by Russell, he inferred his theorem that the power set of any set S has a cardinality greater than that of S.[1]

Russell was appalled. At first he thought that Cantor's theorem couldn't possibly be true. In an essay published in 1901 and devoted largely to an explanation of how Cantor's brilliant discoveries have solved centuries-old metaphysical problems, Russell pointed out that even the master is capable of making mistakes, since Cantor's theorem is false. Clearly, he thought, not *every* cardinal number has a greater one:

> There is a greatest of all infinite numbers, which is the number of all things altogether, of every sort and kind. It is obvious that there cannot be a greater number than this, because if everything has been taken, there is nothing left to add. Cantor has a proof that there is no greatest number. . . . But in this one point, the master has been guilty of a very subtle fallacy, which I hope to explain in some future work. (*Mysticism and Logic*, p. 69)

In Russell's opinion, the essence of Cantor's proof was independent of the bijective nature of his function f and of the fact that its range was the whole power set of its domain. The main point, as he explained to Frege in 1902, was that given any class A, *any* subset B of the power set of A, and *any* function f from A onto B, the class of all things in A that are not members of their f-values, cannot be a value of f. It was late in 1900 that Russell first seriously addressed the question, What is wrong with Cantor's proof? He had recently established to his satisfaction that the class of all things and the class of all classes (which he may have regarded as its power set) had the same cardinality.[2] Russell tried to follow step by step Cantor's argument as applied to this case. He first defined a many–one function f between the universe and its power set, defined as follows: $f(x) = x$ if x is a class; $f(x) = \{x\}$ otherwise (*Principles*, sec. 349, p. 367). At this point, Cantor's reasoning focused attention on the class of all things in the universe that are not members of their f-values. This class, call it R, is the class of all classes that are not members of themselves. Cantor's result – as generalized by Russell – says that R couldn't possibly be a value of f; that is, for no object t is $f(t) = R$. But this is obviously false, since by our definition of f, $f(R) = R$! There *must* be a mistake in Cantor's argument that there is no such t. Let us follow the remaining steps of the reasoning carefully.

Cantor assumed there is such a t (we know it is R) and examined the question of whether it belongs to $f(t)$ (in our case, by our definition of f, this is R again). Cantor said that we can now derive a contradiction: From the assumption that t belongs to $f(t)$ (i.e., that R belongs to R), derive the opposite. Russell saw that he was right. Let us try the other way. From the assumption that t does not belong to $f(t)$ (i.e., that R does not belong to R), Cantor said we can derive the opposite. What Russell now had in his hands was the Russell paradox.[3]

At first Russell saw in this no more than a difficulty for Cantor. But it eventually dawned on him that the contradiction he had derived depended only on assumptions he himself had been willing to grant. The shocking fact that all one needed to assume in order to derive the contradiction were premises that had been tacitly acknowledged by most mathematicians, and explicitly acknowledged by the single logician who had attempted to put the mathematical house in order, Frege. If one assumed, as Frege had in his *Grundgesetze*, that the concept *x is a class of classes not belonging to themselves* must have a class as its extension, the contradiction was inevitable

On the face of it, Russell's class R appeared no more dubious or remote than any other. But Russell's argument had made clear that its existence or subsistence was unacceptable. As this first domino fell, others started tottering. Russell noticed that it was a routine matter to generate further classes and relations in extension that led to very similar paradoxes and, therefore, could not exist.[4]

The conclusion was shocking but inescapable; it was not clear what classes were. As already noted, in December 1902, in the preface to *Principles*, Russell wrote that the discussion of indefinables "is the endeavour to see clearly, and to make others see clearly, the entities concerned, in order that the mind may have that kind of acquaintance with them which it has with redness or the taste of a pineapple" (p. xv). His struggle with the contradiction may have gone through his mind as he added: "In the case of classes, I must confess, I have failed to perceive any concept fulfilling the conditions requisite for the notion of *class*. And the contradiction discussed in Chapter x [i.e., Russell's paradox] proves that something is amiss, but what this is I have hitherto failed to discover" (pp. xv–xvi).

In appendix A to *Principles*, written after the book was completed, Russell noted that "it is very hard to see any entity such as Frege's range [*Wertverlauf*]" (p. 514), that is, Frege's "sets," but since "without a single object to represent an extension, Mathematics crumbles" (p. 515), Russell reluctantly concluded that "it would seem necessary . . . to accept ranges by an act of faith, without waiting to see whether there are such things" (p. 515).[5] Class theory was not the healthy organism that Cantoreans had taken it to be. The question was, How far does the cancer spread?

Surely not to the entire organism – or so Russell thought at first, for not only Cantor's paradise but the bulk of classical mathematics seemed to depend on classes. What was needed, he hoped, was limited preventive surgery, to be implemented by uncovering some hitherto neglected distinction that would help us understand the limit between reality and fantasy in the domain of classes.

Zermelo's axiom

The emergence of Zermelo's axiom is one of the most interesting episodes in the early history of mathematical logic. As soon as he made its use explicit, the axiom became the focus of philosophical dissent.

Borel, Poincaré, and Peano, and later Brouwer, Weyl, and Wittgenstein, either rejected the axiom or expressed serious misgivings about it. In all cases their concern was related to the broken link between extension and intension. The only alternative to their intensionalist standpoint seemed to be a picture of logic (and mathematics) as a physics of abstract objects. Indeed, their opponents seemed to think that reference to classes can somehow be taken for granted and that claims about them are like conjectures about atoms or tables. They seemed to think, that is, that for claims of logic (and mathematics), proper reference is presumed without question; truth value is therefore determined even though unknown; and believability is determined through inductive considerations entirely analogous to those that lead us to accept the existence of physical objects. From this standpoint, the way to respond to Russell's question about reality and illusion in the domain of classes is to do what Zermelo did in 1908 – to state an axiom system, a "theory" about classes, and to test it by checking whether it entails all of the desired and none of the undesired consequences.

Zermelo's reply to his critics in "A New Proof of the Possibility of a Well-Ordering" (1908, van Heijenoort, *From Frege to Gödel*) raised the epistemological issue in its clearest terms. Peano had noted that Zermelo's axiom was not derivable from the principles in his *Formulaire* and had displayed a certain skepticism concerning its truth. Zermelo's reasonable response was to ask, "How does Peano arrive at his own fundamental principles and how does he justify their inclusion in the *Formulaire?*" (van Heijenoort, *From Frege to Gödel,* p. 187). This was, no doubt, the hardest question one could ask in the field of logic at the time. As we shall see in the next section, no one had much of an answer. Zermelo's own guess was that the way to identify those fundamental principles was through the analysis of "the modes of inference that in the course of history have come to be recognized as valid and by pointing out that the principles are intuitively evident and necessary for science" (p. 187).

Peano had consigned questions of "intuitive evidence" to the domain of psychology: "Should we now state our opinion whether the proposition is true or false? Our opinion does not matter" (*Opere scelte,* vol. 1, p. 349). Zermelo thought this was a frivolous attitude, but he did not have an independent argument for the obviousness of his principle that couldn't also be used to support the obviousness of, say, Euclid's parallel postulate. He could, however, make a case for the "necessity for science" of the axiom of choice. In "A New Proof of the Possibility of a Well-Ordering," Zermelo listed a number of widely accepted results in set theory and argued that their traditional derivations tacitly relied on the axiom of choice. The principle had been an unnoticed part of mathematics. As we try to bring clarity to that discipline, our role is not to criticize but to make explicit what has been implicit: "Principles must be judged from the point of view of science, and not science from the point of view of principles fixed once and for all" (van Heijenoort, *From Frege to Gödel,* p. 189). In other words, there are (in mathematics) no standards above and beyond those implicitly recognized in scientific practice. If a principle is widely (even if tacitly) used, if it leads to widely acknowledged results that cannot be derived without it, and if it does not lead to contradictions, then all "philosophical" objections are misplaced, since they must rely on standards that do not matter to mathematics. The reasons for accepting mathematical axioms are, in a sense, inductive, in that they derive from the character of mathematical practice and not from any extramathematical form of intellectual activity.

In his reaction to Zermelo's axiom, Russell sided with the intensionalists. The axiom asserts (in the version independently discovered by Whitehead, known as the "multiplicative" axiom): Given an infinite family of nonempty classes, there is always another class that has exactly one element from each of the classes in the family. A class is said to exist even though no effort is made to establish the existence of an intension that would determine it. With what right could we make this assumption?

> Both in the case of Zermelo's axiom and in that of the multiplicative axiom, what we are primarily in doubt about is the existence of a norm or property such as will pick out one term from each of our aggregates; the doubt as to the existence of a *class* which will make this selection is derivative from the doubt as to the existence of a norm. ("On Some Difficulties in the Theory of Transfinite Numbers and Order Types" [1906], *Essays in Analysis,* pp. 162–3)

The celebrated example of an infinite class of boots was designed to illustrate this difficulty, in view of the presence and absence, respectively, of a rule to divide the class into two halves. In the latter case, he explained, "we cannot discover any property" belonging to exactly half the boots (p. 157). "If the number of pairs were finite, we could simply choose one of each pair; but we cannot choose one out of each of an infinite number of pairs unless we have a *rule* of choice" (pp. 157–8). Quite generally, he was inclined to think that "a *norm* [i.e., "property" or "propositional function"] is a necessary but not sufficient condition for the existence of an aggregate" (p. 136).[6] The idea that there might be "lawless" classes of integers appeared to him "open to doubt." "It would seem," he added, that "an infinite aggregate requires a norm, and that such haphazard collections as seem conceivable are really non-entities" (p. 163).[7]

Yet the axiom of choice appeared to be necessary for the derivation of mathematics from "logic." The philosophical half of Russell's heart was with the intensionalists, the mathematical half with Cantor's set theory. With great relief, in late 1905, Russell realized that his incomplete-symbol strategy allowed him to postpone a decision. Russell thought he might be able to do away with class commitments entirely by methods analogous to those he had used to do away with the commitment to the "meanings" of definite descriptions. Initially this was no more than a hope. It soon became a theory. Late in 1905 he wrote to Jourdain: "I believe I can now deal satisfactorily with all the various contradictions; I do so by wholly denying that there are such things as classes and relations, which I treat in the same way as I treated denoting phrases in the current *Mind*" (Grattan-Guinness, *Dear Russell – Dear Jourdain,* p. 56). The no-classes theory to which Russell referred in this passage was explained in "On the Substitutional Theory of Classes and Relations" (1906, *Essays in Analysis*) and was based on the assumption that propositions are real – a view Russell would soon abandon for reasons to be examined in Chapter 8. By 1908 he had developed the no-classes theory that would be incorporated in *Principia.*

Roughly speaking, Russell's no-classes theories treat class symbols the way "On Denoting" treated definite descriptions, by identifying a general technique for characterizing the truth conditions of all sentential contexts in which they occur. In brief, all sentences containing a class symbol are to be replaced by sentences making a corresponding assertion about an associated propositional function or property. Talk about classes is thereby translated into talk about properties (propositional functions) that have as their extension the class in question. All class paradoxes are thereby solved, since as far as Russell's logic could tell, there are no classes. The corresponding intensional paradoxes did remain, of course; they would be handled through type theory. Also the new intensionalized interpretation of mathematics required axioms that were hardly distinguishable in syntactic structure from those that had generated so much dissent when conceived as referring to classes. Russell might have prided himself on noting that the axioms were now intelligible, however, and that they did not presuppose a dubious ontology. (After all, we are acquainted with intensions but not with classes.) But the difficult question remained, What reason is there to think that these axioms are logically true? Indeed, what reason is there to believe that they are true at all?

Conjectural logic

As the logicist project evolved, Russell's "logic" took shape, but the shape it took was rather unexpected. A year after the publication of *Principles*, Russell was still explaining, "The view advocated by those who, like myself, believe all pure mathematics to be a mere prolongation of symbolic logic, is, that there are no new axioms at all in the later parts of mathematics, including among these both ordinary arithmetic and the arithmetic of infinite numbers" ("The Axiom of Infinity" [1904], *Essays in Analysis*, p. 256). Yet as he joined Whitehead in the effort to fulfill the reductionist promises of *Principles*, he was forced to change his mind.

The first major surprise was the one we have just talked about. In 1904, while revising one of Whitehead's proofs of a certain statement that was required to develop (inter alia) the theory of cardinal multiplication, Russell discovered that the proof was circular. He tried other ways to prove the formula in question and found none. Eventually, he came to think that this "multiplicative axiom," as they started calling it, was, indeed, an axiom, an assumption without which mathematics could not be reduced to "logic."[8]

The next surprise was the axiom of infinity. In *Principles* Russell had followed Bolzano and Dedekind in taking this "evident" truth to be capable of proof from standard logical assumptions (sec. 339, p. 357). In response to a challenge by Kayser to the effect that an extra assumption was involved, Russell insisted in "The Axiom of Infinity" that there is no need to state an axiom of infinity. Yet shortly thereafter he changed his mind and, in fact, concluded (probably due to type-theoretic considera-

tions) that the axiom asserts the existence of infinitely many individuals and is therefore a purely empirical assumption.[9]

These new axioms were, perhaps, required to deduce classical mathematics; but were they part of "logic"? What, exactly, was this "logic" to which Frege and Russell had decided to reduce mathematics? Even those who do not care about such questions might still wonder whether these axioms, logical or otherwise, are true; or whether there is any other reason to accept them.

Russell once said that he felt about the contradiction the way an honest Catholic would feel about a wicked pope. His attitude toward these "logical" axioms could not have been much warmer, especially when he noticed that he could avoid the contradiction but not the axioms. After a few efforts to dissolve the contradiction, Frege must have seen what was happening to logicism; he acknowledged the fact that it had been refuted. Russell saw the same events as an opportunity to redefine the project. He had sided with philosophers on the matter of classes, but when he saw no way to provide a "foundation" for mathematics without assuming these peculiar axioms, he decided it was time to revise those philosophical standards in the light of which logicism appeared to have failed. The idea was to search for a conception of logic that would render the new developments a confirmation of logicism. And Russell quickly found it:

> In fact self-evidence is never more than a part of the reason for accepting an axiom, and is never indispensable. The reason for accepting an axiom, as for accepting any other proposition, is always largely inductive, namely that many propositions which are nearly indubitable can be deduced from it, and that no equally plausible way is known by which these propositions could be true if the axiom were false, and nothing which is probably false can be deduced from it. (*Principia*, p. 59)

A few years earlier, in response to one of Poincaré's attacks against "logistic," Russell noted that the claim for absolute certainty in the choice of a logical basis was no part of his logicist project. This is a common misconception concerning the nature of the evidence on which logistic relies, he explained, and added in a footnote, "Indeed, I shared it myself until I came upon the contradictions" ("On 'Insolubilia' and Their Solution by Symbolic Logic" [1906], *Essays in Analysis*, p. 193). "The method of logistic," he went on,

> is fundamentally the same as that of every other science. There is the same fallibility, the same uncertainty, the same mixture of induction and deduction, and the same necessity of appealing, in confirmation of principles, to the diffused agreement of calculated results with observation. The object is not to banish "intuition," but to test and systematise its employment. . . . In all this, logistic is

exactly on a level with (say) astronomy, except that, in astronomy, verification is effected not by intuition but by the senses. (p. 194)[10]

This is the new perspective on logical matters that Russell framed in response to the challenge to logicism. In his hands, logic began to resemble not only mathematics but even physics. There was, however, still an effort to preserve a link with the old ideals of apriority and certainty. Some propositions, Russell insisted, have a character of "inherent" or "intrinsic obviousness." This domain of "instinctive beliefs," as he also called them, "gives necessarily the basis of all other knowledge. . . . In the natural sciences, the obviousness is that of the senses, while in pure mathematics it is an a priori obviousness" ("The Regressive Method of Discovering the Premises of Mathematics" [1907], *Essays in Analysis*, p. 279). Intrinsic obviousness is the "basis of every science" (p. 279). Yet apart from telling us that intrinsic obviousness is not infallible, Russell had virtually nothing to say by way of elucidation or justification of this "basis of every science." The foundationalist hopes were still there; but the project they inspired had yet to produce anything that might be taken as confirmation of those hopes.

What is logic?

The semantic tradition offered an image of logical and other a priori knowledge that was far superior to all earlier accounts. But it had not completed that image with an explanation of what it is that makes a priori knowledge a priori known, nor had it fared successfully in its modest (if sometimes lengthy) attempts to explain the distinction between a priori and a posteriori. It was at the threshold of this very problem that the semantic tradition encountered its limits.

Kantians and semanticists differed from positivists in the enormous significance they assigned to the a priori. In earlier chapters we examined semanticists' undoing of the Kantian theory and their own constructive contributions. In one way or another, they faced three different issues, which it is now time to distinguish. We shall call them the questions of the *extent*, *intent*, and *ground* of a priori knowledge. In the case of logic, for example, to ask for the extent of logical truth is to ask for the class of statements that qualify as such. To ask for its intent is to ask what it is that makes them logical truths, or why the distinction was worth drawing in the first place. Finally, to ask for the ground of logic is to ask what reason we have to believe that a logical truth is true.

In the preceding section we described the struggle of Russell and others with the problem of how to make sense of the logic that was taking shape in their own hands. Overwhelmed by the complexity of the problems generated by the logicist project, Russell concluded that our

access to a priori truths is no different in kind from our access to the rest. He also argued that there is nothing in the meaning of a priori claims that might allow us to distinguish them from their a posteriori counterparts. What, then, justifies their distinction? What is the intent of logical truth, and what is its ground?

Almost every major nineteenth-century philosopher felt the need to take a stand on the topic of synthetic a priori knowledge, but few thought there was much point in discussing its *analytic* counterpart. One might have hoped that the inadequacy of this attitude would have become clear as logic moved well beyond the principles of identity and contradiction. But even though the members of the semantic tradition were uniquely responsible for the creation of mathematical logic, they had little or nothing of significance to say on the questions of its nature and ground. There can be no doubt that logic, as we now know it, was born in the writings of Frege and Russell. In *Begriffsschrift*, *Grundgesetze*, and *Principia Mathematica*, they told us far better than anyone before which statements and patterns of inference are sanctioned by logic. But why does logic sanction them, and on what ground?

Frege's work was aimed at providing a logical foundation for arithmetic. There is no indication that he ever seriously worried about the foundations of logic itself, however. It isn't simply that it is hard to find an explicit statement in his writings about the intent and ground of logic. Rather, Frege did not even intend to explain why we are justified in believing logic. As we shall see in Chapter 8, Wittgenstein criticized Frege's appeal to rules of inference; he complained, for example, that the rule modus ponens does not provide a justification for the inference from '$A \Rightarrow B$' and 'A' to 'B'. The point is correct, but Wittgenstein was wrong in assuming that Frege intended the rule to play a justificatory role. The idea of a rule of inference was Frege's contribution to what he called the "ideal of a strictly scientific method." This ideal, he said,

which I have here attempted to realize, and which might indeed be named after Euclid, I should like to describe as follows. It cannot be demanded that everything be proved, because that is impossible; but we can require that all propositions used without proof be expressly declared as such, so that we can see distinctly what the whole structure rests upon. . . . Furthermore, I demand – and in this I go beyond Euclid – that all methods of inference employed be specified in advance; otherwise we cannot be certain of satisfying the first requirement. (*The Basic Laws of Arithmetic*, p. 2)

Thus, the purpose is systematic rather than foundational: to localize and minimize the required assumptions. And this purposeful avoidance of foundational questions extends to the case in which the axiomatized theory is logic itself. Frege's logical project is to display explicitly the logical axioms and logical rules, *not* to explain why one should accept them: "The question why and with what right we acknowledge a law of logic to be true, logic can answer only by reducing it to another law of logic. Where that is not possible, logic can give no answer" (p. 15). Evidently, Frege did not believe that "logic must take care of itself." What should? Frege did not know:

> If we step away from logic, we may say: we are compelled to make judgments by our own nature and by external circumstances; and if we do so, we cannot reject this law – of Identity, for example; we must acknowledge it unless we wish to reduce our thought to confusion and finally renounce all judgment whatever. I shall neither dispute nor support this view; I shall merely remark that what we have here is not a logical consequence. What is given is not a reason for something's being true, but for our taking it to be true. (p. 15)

This is all Frege had to say on the topic. A few months before his death, he wrote a piece on the sources of knowledge; there are, he said, three sorts: sense perception, the "logical source of knowledge," and the geometric and temporal sources. There is little doubt about what the first and third items are, but Frege offered no explanation of the second. He did say that "a source of knowledge is what justifies the recognition of truth." But besides telling us that the logical source "is wholly within us," all of his remarks on that topic concerned the extent to which ordinary language may be responsible for errors in logic. The conclusion seems inevitable: The father of modern logic had no opinions on the ground of logical truth.

What about Russell? *Principia Mathematica* was, unquestionably, the most complete codification of logical truths to date. Yet in his review of Ramsey's *Foundations of Mathematics*, Russell acknowledged that "at that time I had no definition of mathematical [i.e., logical] propositions" ("Review of *Foundations of Mathematics* by F. P. Ramsey," p. 477).[11] To put it bluntly, he had no clear idea of what it was that he had codified.[12] Unlike Frege, however, he chose to display in public his private hesitations.

As we have already seen, Russell explained that the reasons we have for accepting a formula as a truth of logic are "inductive." In other words, our reasons for accepting axioms of logic are of the same type as those that lead us to accept axioms of geography: Either the axioms are intrinsically obvious, or we can deduce from them (and from no reasonable alternative) some intrinsically obvious claims. One problem is that for some (e.g., Frege) the intrinsically obvious included propositions of Euclidean geometry, while for others (e.g., Russell) it included propositions about tables and chairs – or about corresponding sense data.

Even more problematic was Russell's view that a priori propositions are not necessary. Since merely empirical claims and purely a priori propositions are entirely alike insofar as their relation to facts is concerned, "it seems impossible to distinguish, among true propositions, some which are necessary from others which are mere facts" ("Meinong's Theory of Complexes and Assumptions," *Essays in Analysis*, p. 26).[13] For example, "the law of contradiction is . . . a fact concerning the things in the world" (*The Problems of Philosophy*, p. 89). This denies that the law is about thought, but it also asserts that its topic is a mere fact, entirely lacking in modal force. The principle of contradiction is, inter alia, about trees, and it states "that if the tree *is* a beech, it cannot at the same time *be* not a beech" (p. 89).[14] As far as Russell's semantics could tell, the principle of contradiction and the statements of geography have the same basic semantic structure: They assert a certain fact that happens to be the case. Nothing in their "meaning" allows for that elusive distinction that we are after: "The difference between an *a priori* general proposition and an empirical generalization does not come in the *meaning* of the proposition" (*The Problems of Philosophy*, p. 106). From whence *does* it come?

Plato provided one standard way to account for the distinction between the a priori and the a posteriori: One thinks of them as claims referring to radically different domains. The a priori concerns certain rigid, immutable objects; the a posteriori deals with the changeable world of experience. Since the days of Aristotle, however, this view has been deemed by some to be too extravagant for belief. From then on, most epistemology has consisted of footnotes to Plato and Aristotle.

One of these footnotes was written by Wittgenstein in 1915. "My method," he wrote, "is not to separate the hard from the soft, but to see the hardness of the soft" (*Notebooks*, p. 44). It is tempting to think that he had Russell in mind, for three years earlier Russell had offered the following solution to the problem of the a priori: "The fact seems to be that all our *a priori* knowledge is concerned with entities which do not, properly speaking, *exist*, either in the mental or in the physical world" (*The Problems of Philosophy*, pp. 89–90). These entities are the meanings of certain very general expressions in our language; they are what Russell called "universals" and "forms":

> The world of universals, therefore, may also be described as the world of being. The world of being is unchangeable, rigid, exact, delightful to the mathematician, the logician, the builder of metaphysical systems, and all who love perfection more than life. The world of existence is fleeting, vague, without sharp boundaries, without any clear plan or arrangement. (*The Problems of Philosophy*, p. 100)

The recognition of this world "solves the problem of *a priori* knowledge" (p. 100). Or so Russell hoped.

The trouble with Platonism had always been its inability to define a priori knowledge in a way that made it possible for human beings to have it. Identifying the *topic* of a priori knowledge, saying that "*all* a priori *knowledge deals exclusively with the relations of universals*" (*The Problems of Philosophy,* p. 103), is not enough to explain how we *know* it. One must add an explanation of how we have access to such universals and their relations: What is the semantic pineal gland that links the world of universals and forms with merely human epistemology? That is why any Platonism that is more than just a colorful rephrasing of common-sense beliefs must postulate *both* an incredible world and an incredible faculty of access to it.

Proponents of the chemical picture of representation had never been good at explaining how we come to possess the ultimate simples from which representational complexity emerges. The little they had to say on this matter had already been said by Hume: "Complex ideas," he had written,

> may, perhaps, be well known by definition, which is nothing but an enumeration of those parts or simple ideas, that compose them. But when we have pushed up definitions to the most simple ideas, and find still some ambiguity and obscurity; what resource are we then possessed of? By what invention can we throw light upon these ideas . . . ? Produce the impressions or original sentiments, from which the ideas are copied. (*Enquiry,* sec. 7, pt. 1, p. 41)

Except for Hume's restriction to the domain of sense impressions, Russell would not have objected. Russell's solution to the problem of indefinables, his "semantic pineal gland," was intuition. He rarely used the term, since its Kantian resonances clearly displeased him, but that is the right word for what he called "acquaintance" and for its propositional correlate, self-evidence. In philosophy, Russell explained in 1900, "the emphasis should be laid on the indefinable and indemonstrable, and here no method is available save intuition" (*A Critical Exposition of the Philosophy of Leibniz,* p. 171). As we have already seen, he explained in *Principles* that

> the discussion of indefinables – which forms the chief part of philosophical logic – is the endeavour to see clearly, and to make others see clearly, the entities concerned, in order that the mind may have that kind of acquaintance with them which it has with redness or the taste of a pineapple. (p. xv)

Ten years later he was still explaining that all knowledge starts with undefined terms and unproved propositions, and that "the undefined terms are understood by means of acquaintance. The unproved propositions must be known by means of self-evidence" (*Theory of Knowledge,* p. 158). Acquaintance and a Platonistic ontology were Russell's twin answers to the problem of the a priori. His philosophy of logic was the immediate consequence of this. It is worth looking briefly at his last attempt to articulate his conception of the matter, just before Wittgenstein explained that his whole approach was ill-conceived.

Russell explained in *Theory of Knowledge* that the ultimate furniture of the world contains three different categories of things: particulars, universals, and forms. The first two are the constituents of propositions (if there are such things); a proposition is a form, "the way in which the constituents are combined in the complex. It is such pure 'forms' that occur in logic" (p. 98).

Russell had always thought of logic as (inter alia) the most general of all sciences. If a proposition mentions anything specific, whether particular or universal, it cannot be logical. A "touchstone by which logical propositions may be distinguished from all others" is that they result "from a process of generalization which has been carried to its utmost limit" (*Theory of Knowledge,* p. 97). We may therefore think of reaching the logical through a process of elimination, by removing every single constituent from propositions. What remains is, of course, not a constituent of the proposition, nor does *it* have any constituents;[15] but it is nonetheless something and, indeed, something with which we are, in fact, acquainted as soon as we understand a proposition *of that form.*

Russell thought that we are certainly acquainted with forms, since otherwise he could not explain the fact that we can understand propositions we have never seen before. They are rearrangements of familiar objects in familiar forms. Since acquaintance with x entails the reality of x, it follows that forms are objects and not symbolic fictions like classes or propositions (p. 129). What objects could they be? The question echoes a similar one Russell had asked a decade earlier of propositions, and once again, he decided the answer should be: some sort of fact. Consider, for example, the complex *Socrates precedes Plato.* Its form must be an object with which we are acquainted, and it must be related to what we obtain by removing all constituents from that complex. Russell's choice was: *something has some relation to something.* Such facts are the subject of logic; and what distinguishes them from the others is that, apparently, it is enough to understand the claims that express them in order to know that they are true. It is enough because for statements of this type there is no distinction, as far as Russell can see, between understanding and acquaintance (p. 130) – hence no distinction between understanding and recognition of truth – since acquaintance is with the very fact that makes the "claim" true.[16] His argument for this: "I am unable introspectively to discover any difference" between acquaintance and understanding in the cases under consideration (pp. 130–1).

If this was Russell's view on the nature of logic, there is serious doubt that *Principia Mathematica* had much to do with that discipline. The claimed link between understanding and truth for logical propositions posed grave problems for the new axioms that were required. Within a few years Carnap would remind Wittgenstein's allies that if they seriously thought understanding in mathematics entailed recognition of truth, they would have to conclude that they did not understand even the statement of Fermat's last theorem.

Logical propositions had acquired a mysterious status. Russell noted how strange it was that these "facts" had no constituents whatsoever. He was pleased to note that they had "all the essential characteristics required of pure forms" (*Theory of Knowledge,* p. 129), but he also wondered, "Why, if pure forms are simple, is it so obviously inappropriate to give them simple proper names, such as John or Peter?" (p. 130). Doesn't this betoken a sort of complexity? What, exactly, is going on?

What is going on is that an approach to logic had finally reduced itself to absurdity. The point did not escape the attention of the one student to whom Russell showed the manuscript from which we have been quoting. Wittgenstein's criticisms (see Chapter 8) led Russell to drop plans to publish the book and, for a while, to abandon philosophical speculation on fundamental, "logical" problems.[17]

We are witnessing here the death of Cartesian dreams:

> It must be taken as a fact, discovered by reflecting upon our knowledge, that we have the power of sometimes perceiving such relations between universals, and therefore of sometimes knowing general *a priori* propositions such as those of arithmetic and logic. (*The Problems of Philosophy,* p. 105)

"It must be taken as a fact . . ." and yet the path that logic had taken under Russell's leadership had made it harder than ever to see this "fact" as anything more than a refusal to face the fundamental problems, an appeal to ancient hopes that the discoveries of the semantic tradition had now made obsolete.

This type of Cartesianism had seemed attractive when compared with what Kantians had done to philosophy. But finally it came time to recognize that the semanticists' project was in need of major revision. The first clear indication of what the new approach might be emerged from unexpected territory: from a meditation on the foundations of geometry. Ironically, the most outspoken opponents of the development that would eventually lead to an understanding of the nature of logic were Frege and Russell.

What is geometry?

Their desire to avoid the psychologism they saw in the Kantian answers and their commitment to semantic monism left most semanticists with no choice but the very Platonism that Kant had justifiably dismissed as dogmatic metaphysics. And when they asked for the ground of the a priori, the only answer they had was, in effect, intuition. It is ironic that after all those complaints against Kant's pure intuition, at the culmination of this revolutionary process we find an appeal to a form of intuition that Kant himself would have regarded as extravagant. As the semanticist position revealed its inability to cope with this problem, a new standpoint emerged that would combine the insights of the semanticists with a concession to Kantianism. The transition to this new perspective was initiated near the turn of the century, and its conflict with the semantic tradition was first displayed in a debate concerning the foundations of geometry.

Russell and Poincaré had been allies in the struggle to remove Kantian intuition from the field of geometry and transform it into a purely conceptual discipline; but once that war was won, they turned against each other. Their quarrel concerned a seemingly trivial topic: How do we have access to the basic, indefinable, geometric concepts? At about the same time, Frege and Hilbert were also examining that question. They did not realize that the issue they were debating held the key to the question of the nature of a priori knowledge, even in the field of logic.

Since the 1880s Poincaré had been advocating a remarkable doctrine concerning the nature of geometry, according to which no properly geometric axiom expresses either "an experiential fact or a logical necessity or a synthetic a priori judgment" (Poincaré, "Analyse de ses travaux scientifiques," p. 127). What else could an axiom be? Poincaré's often repeated answer was that they are "definitions in disguise." In 1899 Hilbert published a monograph on the foundations of geometry in which he adopted a closely related standpoint, also describing geometric axioms as definitions (*Erklärungen*). When Russell found out about Poincaré's views and Frege about Hilbert's, their responses were, revealingly, the same: Geometers are thoroughly confused on the nature of definitions, and they need to be enlightened.

In fact, Frege and Russell had totally misunderstood the geometers' point; for they had even missed the problem their remarks were aimed to solve. Let us see how and why.

Russell and Poincaré

In 1897 Russell had published his fellowship dissertation under the title *An Essay on the Foundations of Geometry.* The *Revue de Métaphysique et de Morale* soon published a very enthusiastic review of it by Couturat in which he remarked that the work revealed a mind endowed with "a vast mathematical erudition" and an equally vast "understanding of philosophical problems." Amidst a flood of extravagant praise, Couturat added "That such a mind could not be found in France may be cause for regret but not for surprise" ("Essai sur les fondements de la Géométrie par Bertrand Russell," p. 354). The next volume of the *Revue* included a long, careful, devastating discussion of Russell's book by Poincaré ("Des fondements de la géométrie").

Of the many problems raised by Poincaré, the one that concerns us here is the one Russell called the "most important and difficult issue," that of the "definition" of geometric primitives. As part of his defense of conventionalism, Poincaré challenged Russell to explain what, in his view, were the meanings of a number of primitive notions. Russell replied:

> M. Poincaré requests "a definition of distance and of the straight line, independent of (Euclid's) postulate and free from ambiguity or vicious circle" (Sec. 20). Perhaps he will be shocked if I tell him that one is not entitled to make such a request since everything that is fundamental is necessarily indefinable. And yet I am convinced that this is the only philosophically correct answer. Since mathematicians almost invariably ignore the role of definitions, and since M. Poincaré appears to share their disdain, I will allow myself a few remarks on this topic. ("Sur les axiomes de la géométrie," pp. 699–700)

There are, Russell explained, two kinds of definitions: mathematical and philosophical. Mathematical definitions (soon to become "knowledge by description") merely identify an object as the one and only that stands in a certain relation to already known concepts or objects (p. 700). For example, if we define the letter *A* as the one that precedes *B*, or the number 1 as that which precedes 2, what we have given is a mathematical definition of those objects:

> But these definitions are not definitions in the proper and philosophical sense of the word. Philosophically, a term is defined when its *meaning* is known, and its *meaning* cannot consist of relations to other terms. It will be readily granted that a term cannot be usefully employed if it does not mean something. Its meaning can be complex or simple. In other words, either it is composed of other meanings or it is one of those ultimate elements that are constituents of other meanings. In the first instance one defines the term philosophically by enumerating its simple elements. But when it is itself simple, no philosophical definition is possible. . . . Definition is an operation analogous to spelling. One can spell words but not letters. M. Poincaré's request places me in the uncomfortable situation of a student who has been asked to spell the letter A without allowing him to use that letter in his reply. . . . All these truths are so evident that I would be ashamed to recall them, were it not that mathematicians insist in ignoring them. (pp. 700–1)

Applying these remarks to geometry, Russell concluded:

> These observations apply manifestly to distance and the straight line. Both belong, one might say, to the geometric alphabet; they can be used to define other terms, but they themselves are indefinable. It follows that any proposition, whatever it may be, in which these notions occur, is either an axiom or a theorem and not a definition of a word. When I say that the straight line is determined by two points, I assume that *straight line* and *point* are terms already known and understood, and I make a judgment concerning their relations, which will be true or false, but in no case arbitrary. (pp. 701–2)

Russell's remarks would have been devastating if Poincaré had meant by giving a 'definition' what Russell thought he *should* have meant, that is, analysis. But Poincaré merely meant what the dictionary instructed him to mean, that is, a process through which meaning is assigned to an expression. As Mill had put it, a definition is a "proposition declaratory of the meaning of a word" (*Logic*, bk. 1, chap. 8, sec. 1, p. 133). On Russell's view, to define a word is to construct (synthesize) new meanings from old, the latter being regarded as constituents of the former. The dictionary sense of 'definition', in contrast, allows processes which do not assume that meanings are available prior to the definition (as in 'ostensive definition', 'coordinative definition', and the like). Thus, when Poincaré asked Russell to define geometric primitives, he was not asking for the analysis of the unanalyzable; he was requesting a sufficiently definite and geometrically acceptable characterization of what those primitive terms mean.

Russell could not see the point of the question. In addition, quite independently, he thought it obvious that one could establish a priori that the geometric axioms could not be employed in the process of assigning meanings to terms involved in their formulation. His reasoning involved an appeal to a principle that may be called the *thesis of semantic atomism.*

This principle says that if a sentence *S* is to convey information (or, as Russell or Frege would put it, to express a proposition), then its grammatical units must have a meaning *before* they join their partners in *S.* Recognition of the meaning of the constituent phrases must be independent of and prior to the acceptance of the statement in question. The harmless appearance of this principle will vanish as we begin to see the dominant role it came to play in these geometric debates. The basic fact to bear in mind is that *all* participants in the debates endorsed this principle. The philosophers used it to infer (by modus ponens) a conception of geometry that the geometers could not accept, while the geometers used it to infer (by modus tollens) a picture of geometric knowledge that the philosophers would not take seriously. Whether one was inclined to move up or down the argument chain depended entirely on one's attitude toward the character of indefinables.

At the end of our quotation from Russell's "Sur les axiomes de la géométrie" (pp. 701–2), we caught a glimpse of the train of thought that led him to regard as absurd the idea that axioms can be used to give definitions. His argument was, in effect, this: Since *obviously* the axioms of geometry express propositions (convey information), by the thesis of semantic atomism the geometric primitives must *somehow* acquire a meaning before they can contribute to the expression of the appropriate propositions. Poincaré was not satisfied with "somehow"; he wanted to know *how*. How, he asked, are we supposed to decide whether this or that entity is a point, a straight line, a plane? How are we to tell what distance is?

In a reaction typical of the worst tendencies in the semantic tradition, Russell promptly concluded that Poincaré was confusing epistemology with semantics. How we find out whether something is the case, he thought, can have no bearing on what that "something" is:

> If there are such quantities as distance and angle, their measure can only be arbitrary so far as concerns the choice of a unit, and any different measure must be simply false; while if there are no such quantities, then they can have no measure at all. . . . How we discover two actual spaces to be equal is no concern of the geometer; all that concerns him is the existence of equal spaces. . . . The whole confusion appears to be due to not distinguishing between the process of measurement, which is of purely practical interest, and the meaning of equality, which is essential to all metrical Geometry. ("Geometry, Non-Euclidean," p. 671)

But clearly neither Poincaré nor his logical positivist followers meant to challenge (as others eventually would) the idea that before we can raise the question of testing, we must have solved the question of meaning. On the contrary, Poincaré agreed with Russell that the latter had to be solved before the former. It was Russell's answer to the question of meaning that Poincaré found entirely untenable, and because of this turned to his new interpretation of geometric axioms.

Though in his debate with Poincaré Russell refused to state and defend his own answer to this question of meaning, we know what it was: Geometric indefinables are first given to us in acquaintance.[18] Poincaré also knew what Russell had in mind. In "Des fondements de la géométrie" he suggested that Russell's answer would be as follows:

> There is no need to define [the indefinables] because these things are directly known through intuition. I find it difficult to talk to those who claim to have a direct intuition of equality of two distances or of two time lapses; we speak very different languages. I can only admire them, since I am thoroughly deprived of this intuition. (p. 274; see also "Sur les principes de la géométrie," p. 75)

The emptiness of Russell's appeal to acquaintance becomes clear when regarded in the context of the relevant geometric facts.[19] By the end of the nineteenth century, the only reason anyone could possibly have for

saying that the ultimate distinction between Euclidean and hyperbolic *distance* is given by acquaintance was the weight of a dead philosophy. One might insist that once we understand these notions, we have become acquainted with the concepts in question. But this could be viewed only as a linguistic ploy, designed to obscure the fact that acquaintance could not be asked to play in geometry the role that atomism had assigned to it. Circa 1900 it was no longer possible to suppose that acquaintance plays the specific semantic *explanatory* role in geometry that it was supposed to play in the atomist picture of knowledge, that is, that the construction of geometric theory would start with acquaintance, then proceed to a construction of claims, and perhaps conclude with the testing of these claims.

Poincaré's conventionalism was based on the idea that in order to understand geometry, one must stand Russell's argument on its head: Since geometric primitives do not acquire their meaning prior to their incorporation into the axiomatic claims, such axioms do *not* express propositions in Frege's or Russell's sense. A passage in his reply to Russell clearly displays the semantic dimension of Poincaré's views. Struggling to explain why he thought it is a mistake to conceive of the axioms of geometry as bona fide propositions, Poincaré said:

> If an object has two properties A and B, and if it is the only one that has the property A, this property can be used as a definition; and since it will suffice as a definition, the property B [i.e., the attribution of B] will not be a definition: it will be an axiom or a theorem. If, on the contrary, the object is not the only one that has the property A, but it is the only one that has both properties A and B, A no longer suffices to define it, and the property B will be a complement of the definition, not an axiom or a theorem. In a word, in order for a property to be an axiom or a theorem it is necessary that the object that has this property has been completely defined *independently of this property*. Therefore, in order to have the right to say that the so-called distance axioms are not a disguised definition of distance, one should be able to define distance in a way that does not involve an appeal to those axioms. But where is that definition? ("Des fondements de la géométrie," p. 274)

Poincaré made the same point in a discussion of free mobility. Russell had argued that the axiom of free mobility is a priori, and he had formulated it as follows: *"Spatial magnitudes can be moved from place to place without distortion. . . . Shapes do not in any way depend upon absolute position in space"* (*An Essay on the Foundations of Geometry*, p. 150). Poincaré asked:

> What is the meaning of "without distortion"? What is the meaning of "shape"? Is shape something that we know in advance, or is it, by definition, what does not alter under the envisaged class of motions? Does your axiom mean: In order for measurement to be possible figures must be susceptible of certain motions and

there must be a certain thing that will remain invariant through these motions and that we shall call shape? Or else, does it mean: You know full well what shape is; well, in order for measurement to be possible it is necessary that figures can undergo certain movements that do not alter their form. I do not know what is it that Mr. Russell has meant to say; but in my opinion only the first sense is correct. ("Des fondements de la géométrie," p. 259)

With minimal sharpening, the argument is this: The axioms of geometry are widely regarded as statements that convey information about certain elusive geometric entities. If they do, then by the principle of semantic atomism it should be possible to "define," that is, to identify in some intersubjective manner, the meanings of their geometric primitives before they are incorporated into the axiomatic sentences. Up to this point the geometer and the philosopher agree; but now Poincaré brought in a new premise, the lesson that geometers had learned from the evolution of non-Euclidean geometry: There is really nothing we can say about the meaning of geometric primitives beyond what the axioms themselves say. There is, of course, nothing to prevent us from *deciding* to circumscribe those meanings further so that the terms in question would refer only to certain physical objects (light rays, etc.); but there is no particular meaning of that sort or of any more ethereal Platonic sort that geometry attaches to its primitives prior to its construction. Geometry does not depend on geometric *objects*, whether they be Platonic straight lines or Millian light rays; all it needs in order to have a life of its own is geometric concepts or meanings. And such meanings are constituted roughly in the way in which Kantians used to think that we constitute experience or its objects, through the employment of rules or maxims whose adoption is prior to and the source of the meanings in question.

Therefore, according to Poincaré, all we can say about the meanings of geometric primitives is what geometric axioms say. Under the circumstances, the thesis of semantic atomism prevents those axioms from conveying any sort of factual (nonsemantic) information. No wonder that they are neither analytic (in Kant's first sense) nor synthetic, since they are not propositions. No wonder either that they had always been regarded as extraordinary claims, endowed with a particularly strong sort of truth. The error was to think that they convey a privileged sort of information or information about some extraordinary domain. Their distinguishing feature is that they determine, to the extent needed in geometry, the meanings of geometric primitives. The conviction that they are necessary emerges from the fact that we would be talking about something else or, better yet, meaning something different from what is intended, if we denied them. Geometric axioms are definitions disguised as claims, and what they define are the indefinables.

Frege and Hilbert

At the same time that Poincaré was crossing papers with Russell, Hilbert was releasing a monograph that would soon become one of the landmarks of nineteenth-century geometry, his *Grundlagen der Geometrie*. O. Blumenthal reports that as early as 1891, commenting on a lecture by H. Wiener, Hilbert had said that "it must always be possible to replace [in geometric statements] the words, 'points', 'lines', 'planes', by 'tables', 'chairs', 'mugs'" (Hilbert, *Gesammelte Abhandlungen*, vol. 3, p. 403). Several years later he decided to put the idea to work. In the winter term of 1898–9 he offered a course on the foundations of Euclidean geometry, on which *Grundlagen* was based.

Frege had had a deep interest in geometry from very early on. For reasons he never explained and may never have seriously scrutinized, he thought that geometry was a clear instance of a priori knowledge based on pure intuition. He read Hilbert's monograph as soon as it was published, and his immediate reaction was disappointment. He wrote to a friend that the book was "a failure" (Letter to Liebmann [1900], *Wiss. Briefwechsel*, p. 148) and started corresponding with Hilbert in order to set him straight on the relevant logical issues.

It is hard to avoid a sense of déjà vu when one sees that Frege's main complaint was that Hilbert did not seem to understand the nature of definition. As is well known, Hilbert had started the *Grundlagen* stating what he called a definition, which turned out to be the axioms of his formulation of Euclidean geometry. Frege was appalled. "I think it is high time," he wrote, "that we came to an understanding about what a definition is and what it is supposed to accomplish. . . . It seems to me that at the present time complete anarchy and subjective inclination reign supreme" (Letter to Hilbert [1899], in *Wiss. Briefwechsel*, p. 62). What followed was a masterful and patronizing explanation of the classical picture of knowledge. The totality of the sentences of a theory, Frege explained, is to be divided into two groups, those in which something is asserted and those in which something is stipulated. The former are the axioms and theorems of the theory, the latter are the definitions:

It is absolutely essential for the rigor of mathematical investigations that the difference between definitions and all other sentences be maintained throughout in all its sharpness. The other sentences (axioms, principles, theorems) must contain no word (sign) whose sense and meaning (*Sinn und Bedeutung*) or (in the case of form words, letters in formulas) whose contribution to the expression of the thought is not *already* completely settled, so that there is no doubt about the sense of the sentence – about the proposition expressed in it. Therefore it can only be a question of whether this proposition is true, and on what its truth rests. Therefore, it can never be the purpose of axioms and theorems to establish the meaning of a sign or word occurring in them; rather, this must *already* be established. (pp. 62–3; my italics)

Clearly, a central role is played by the thesis of semantic atomism; and equally clear is Frege's conviction that *every* sentence of a theory that is not a definition must express a proposition and thereby convey (true or false) information. Given these two assumptions, Hilbert's axioms in sections 1 and 3 of the *Grundlagen* should be such that "the meanings of the words 'point', 'straight line', and 'between' are not given but are presupposed as known" (p. 61).

It is in this conclusion that Hilbert located "the crux of the misunderstanding":

I do not want to presuppose anything as known. I see in my explanation in section 1 the definition of the concepts point, straight line, and plane, if one adds to these all the axioms of group i–v as characteristics. If one is looking for other definitions of point, perhaps by means of paraphrase in terms of extensionless, etc., then, of course, I would most decidedly have to oppose such an enterprise. One is then looking for something that can never be found, for there is nothing there, and everything gets lost, becomes confused and vague, and degenerates into a game of hide-and-seek. (Letter to Frege [1899], Frege, *Wiss. Briefwechsel,* p. 66)

Compare Hilbert's statement about preaxiomatic procedures for capturing the indefinables of geometry with Poincaré's remarks about those who claim to be acquainted with (to intuit) them. For both of them, and eventually for all geometers, the preaxiomatic search for the indefinables is "a game of hide-and-seek" in which "everything gets lost, becomes confused and vague" because, in the end, "there is nothing there."

Besides semantic atomism, Frege and Russell shared a further doctrine, which we shall call "propositionalism." The propositionalist notes that all branches of knowledge, including logic and geometry, formulate their claims in statements that, syntactically speaking, do not seem to differ significantly from regular factual claims. The statement that no claim can be both true and false and the statement that two points determine a unique straight line seem to differ only in topic and degree of certainty from the statement that this table is brown. For the propositionalist this syntactic uniformity is as it should be, for all these statements are seen as playing essentially the same syntactic role: They tell us how things stand. According to the propositionalist, logic, geometry, physics, and ordinary talk all call for the same type of semantic analysis; in all cases we are dealing with "propositions" in the minimal sense that takes them to be vehicles of information. And only propositions can be regarded as the targets of what Russell called the "propositional attitudes" (assertion, assumption, belief, etc.) and as the subject matter of logical operations such as inference and proof.[20] For Frege and Russell, as for their predecessors, propositionalism was not so much a conscious assumption adopted after exploring the alternatives; they simply would not have known how to begin making sense of the claim that the most basic principles of knowledge, such as the laws of geometry or logic, are really not the sorts of things that say anything and therefore are not the sorts of things that could be true or false. In their view one might conceivably argue that geometric propositions make weaker claims than previously thought or even, at the limits of sanity, that they make purely logical claims. Frege thought this suggestion was ludicrous, but at least it had the virtue of being intelligible.

In his second letter to Hilbert, Frege made his first attempt to interpret what Hilbert was doing. "It seems to me," he wrote,

that you want to detach geometry completely from the intuition of space and to make it a purely logical discipline, like arithmetic. If I understand you correctly, the axioms that are no doubt usually considered the basis of the whole structure, on the assumption that they are guaranteed by the intuition of space, are to be carried as conditions in every theorem; not, to be sure, expressed in their full wording but as contained in the words 'point', 'straight line', etc. (Letter to Hilbert [1900], *Wiss. Briefwechsel,* p. 70)[21]

Frege said no more on this matter until 1906, when he raised the issue once again in the second part of "Über die Grundlagen der Geometrie." Hilbert's axioms and theorems, he explained, are not propositions but "improper sentences," that is, sentences from which one or more meaningful terms have been removed and replaced by variables. Hilbert did not really mean to assert either his axioms or his theorems, but only certain implications whose antecedents are conjunctions of his "axioms" and whose consequents are each of his "theorems." In fact, we reach the domain of sense – that is, of propositional knowledge – only when we quantify universally the free variables in each of these implications. Thus, "what Mr. Hilbert calls a definition will in most cases be an antecedent improper sentence, a dependent part of a general theorem" ("Über die Grundlagen der Geometrie" [1906], *Kleine Schriften,* p. 303). Part 2 of Frege's "Über die Grundlagen der Geometrie" was devoted entirely to an elaboration of if–thenism that was, as might be expected, much more thorough than any Russell ever gave.[22] Frege examined, for example, an "alleged" proof of a Hilbertian theorem, showing in painful detail how to reconstruct it as a proof of an implication of the appropriate sort. It should be emphasized, however, that while Russell put forth if–thenism as his own view of the matter, Frege never endorsed this doctrine but stated it only as the best sense he could make of Hilbert's words. And he thought it was obvious that this much sense was not sense enough.[23]

The discovery of syntax

In the spirit of tolerance, most who have examined these debates have argued that the outcome was a draw. It is said that the participants neglected to make the distinction (later made by Russell) between pure and applied geometry; had they done so, they would have noticed that they were talking past each other. Hilbert and Poincaré were surely talking about uninterpreted geometries, so were, of course, right in denying that there were any propositions in them. But Frege and (initially) Russell were talking about interpreted geometry. They were therefore right in viewing them as sets of true or false propositions.

The problem with this "resolution" of the conflict is that it creates the impression of a compromise or of a higher synthesis while, in fact, granting everything to one side and nothing to the other; for to endorse the interpreted–uninterpreted distinction as a *sufficient* account of the character of geometric claims is to grant Poincaré's and Hilbert's whole case. If all there is to geometry besides its uninterpreted form is the populous democracy of geometric models, then the noble class of propositions that constitutes *real* geometry according to Frege and the young Russell is lost among the uncountable multitude of swindlers that pass for interpretations. The idea that propositions about his watch could be part of "geometry" was, for Frege, unspeakably silly. Any account of geometry that did not include a way to distinguish between propositions about points – *real* points, that is – and propositions about watches was, in his view, hopelessly inadequate.

The real issue was whether geometric axioms have to be understood as expressing propositions. Poincaré and Hilbert saw more clearly than anyone else at the time that the propositionalist reading was inadequate. Moreover, their efforts to convince the propositionalists of the peculiar role of geometric axioms were the first serious attempts to acknowledge a distinction that was destined to have a long and illustrious history in the twentieth century. Wittgenstein's domain of showing, his later grammar, Carnap's syntax, Sellars's categorial frameworks, and Kuhn's paradigms are some well-known members of the continuing series of attempts to find the right way of looking at that peculiar kind of knowledge that seems necessary and not vacuous, yet at the same time does not quite state any factual claims. Poincaré may have officially started the search when he observed that geometric axioms present themselves "in disguise," pretending to be claims but really being something else; and he also fixed the broad category in which they belong, since by calling them definitions, he clearly intended to assign to them a role in the determination of meaning.

Poincaré's conventionalism adumbrated, however indecisively, the first promising alternative to Kant's guess concerning the nature of apriority and necessity. Convention is sometimes thought to conflict with necessity. Yet convention, semantically interpreted, is merely the opposite side of necessity. In the range of meanings, what appears conventional from the outside is what appears necessary from the inside. The "linguistic" (better, semantic) theory of the a priori that would emerge decades later in the writings of Wittgenstein and Carnap would simply say that all necessity is semantic necessity, that all a priori truth is truth *ex vi terminorum,* that when a statement is necessary, it is because its rejection would be no more than a misleading way of rejecting the language (the system of meanings) to which it belongs. Thus, in the case of a priori claims, one and the same linguistic form may be seen as playing two radically different linguistic roles: When regarded from outside a linguistic framework, it must be seen as part of the definition of one such framework, a definition in disguise; when regarded from within the defined framework, that very sentence now expresses a claim, one true in virtue of the constituted meanings and therefore necessary.

Hilbert's formalism was inspired largely by the same motives as Poincaré's conventionalism. This convergence is hard to see because of longstanding prejudices about the character of formalism, which have also helped distort our understanding of the syntacticist stage in logical positivism. We will have to face those prejudices at the appropriate place, but we might as well say a word about them right now.

Hilbert's project was the culmination of developments starting with Pasch's celebrated *Vorlesungen über neuere Geometrie* of 1882. Pasch still thought that in order to understand the meaning of geometric primitives, there was no other way but to exhibit the corresponding empirical correlates, and that the meanings of axioms similarly depend on their correlation with certain figures. But what distinguished him from most predecessors was his insistence that "the process of inference must be independent in all its parts from the *meaning* of the geometrical concepts, just as it must be independent from the diagrams" (p. 98). Thus, he endeavored to produce axiomatic formulations that provided a sufficient basis for geometry even when judged by standards as exacting as Frege's. Yet Pasch completely misunderstood what Poincaré clearly saw: the role played by meaning in his own considerations. By failing to think through his ideas on the role of ostensive definition, he did not recognize the untenability of geometric empiricism. But, more importantly, by tacitly granting to Kantians and positivists an intimate link between intuition and meaning, he misdescribed his own project as that of the elimination of meaning from geometry, thereby promoting the confusion between the purely formal and the meaningless. That Pasch's actual achievement could have been interpreted as the elimination of meaning from geometry is an indication of the extent to which Kantian presuppositions had become common ground in the intervening century. If Kant was right, concepts without intuitions are empty, and no geometric derivation is possible that does not appeal to intuition. But by the end of the nineteenth century, Bolzano, Helmholtz, Frege, Dedekind, and many others had helped determine that Kant was not right, that concepts without intuition are not empty at all. The formalist project in geometry was therefore designed not to expel meaning from science but to realize Bolzano's old dream: the formulation of nonempirical scientific knowledge on a purely conceptual basis. Once the Kantian prejudice was removed, one could see the hidden message of formalism concerning the meaning of geometric primitives: It is not that meaning is given at the beginning, in order to be immediately taken away so that geometers can do their work properly; rather, as Poincaré and Hilbert argued, meaning is first given by the very axioms that constitute the discipline.

To be sure, in these geometric writings one finds no more than glimpses, oblique adumbrations of things to come. The fog would not start to lift until three decades later; and even then, the myth that geometric formalism conceives of geometry as marks on paper would become the myth that syntax relates not to meaning but only to marks on paper. It would take a long time to realize that besides the suspect sense of meaning offered by the Platonist tradition, there is the sense that 'meaning' has in the English language, and that in this sense, formalism and syntax have a great deal to do with meaning. The next major step toward a clarification of these matters was taken in one of the strangest books ever written.

8

A logico-philosophical treatise

> Whenever I have met unbelievers before, or read their books, it always seemed to me that they were speaking and writing in their books about something quite different, although it seemed to be about that on the surface. . . . Listen, Parfyon. You asked me a question just now; here is my answer. The essence of religious feeling does not come under any sort of reasoning or atheism, and has nothing to do with any crimes or misdemeanors. There is something else here, and there will always be something else – something that the atheists will forever slur over; they will always be talking of something else. – Prince Myshkin.
>
> Dostoevsky, *The Idiot*

> The logic of mysticism shows, as is natural, the defects which are inherent in anything malicious. While the mystic mood is dominant, the need of logic is not felt; as the mood fades, the impulse to logic reasserts itself, but with a desire to retain the vanishing insight, or at least to prove that it *was* insight, and that what seems to contradict it is illusion.
>
> Russell, *Our Knowledge of the External World*

It isn't easy to decide whether Wittgenstein should be included among the members of the semantic tradition or among its most ferocious enemies. On the surface, at any rate, Wittgenstein's problems and techniques were those of the semanticists; beneath the surface, however, things are less clear. The difficulty is not so much that Wittgenstein's purposes were quite different from those of Frege, Russell, and their fellows, but that their philosophical hopes seemed to be Wittgenstein's fears; their projects, Wittgenstein's targets; their enemies, Wittgenstein's friends. Nor is it that Wittgenstein had a radically different view of the nature of logic, mathematics, and science than everyone else in that group; there had always been a large degree of pluralism within the tradition. But for all semanticists, scientific knowledge had been a model, a source of inspiration and of spiritual comfort. For Wittgenstein, however, it was of only secondary interest, to be dealt with much as Kant had dealt with pure theoretical reason: Set limits on it in order to leave room for something more substantial. Even Kant was too much a rationalist for Wittgenstein's taste, since in spite of their limitations, Kant remained enthusiastic about science and about rationality in general. If Wittgenstein was a fifth column among semanticists, it is because from the very beginning his heart was with one of the most romantic, unrational versions of idealism. If it ever looked otherwise, it is in part because he joined the enemy camp in order to display its failure from within and in part, also, because he was more successful at scoring points against his favorite team than against his opponents.

Wittgenstein's philosophy went through a number of stages. Two of them influenced logical positivists decisively. The first stage, characterized by the doctrines of the *Tractatus Logico-Philosophicus,* is the topic of this chapter; the second will be discussed in Chapters 16 and 17.

The *Tractatus* may well be the most difficult philosophical book written in this century. Two facts conspired to produce this result: The thoughts in it are very hard to explain – we are told; and Wittgenstein was singularly uninterested in or incapable of explaining his views to others. Almost everything he wrote was in the nature of a diary, a record of his thoughts, a conversation with himself or with God – hence, he did not feel the need to meet a potential interlocutor even halfway.

Ortega once wrote a piece entitled "In Defence of the Theologian and against the Mystic" (a project Wittgenstein would have detested), in which he noted a somewhat disturbing feature of mystics' dealings with religious subjects. The writings of mystics, he noted, are often classics in their own languages. With unsurpassed eloquence mystics take us stage by stage up the circuitous path leading to their mystical experiences. But when the hour of truth comes, when the real substance of their stories is about to emerge, all of their eloquence vanishes. They let go of our hands and say, "Words elude me at this point. . . . I must proceed now to my silent mystical experience; I wish you luck in getting your own." Something not entirely unlike this happens in the *Tractatus.* Its final aphorism tells us that "whereof one cannot speak, thereof one must be silent," and as Wittgenstein explained to von Ficker in 1919, it is what comes (or should have come) after this statement that really mattered to him: "My work consists of two parts: the one presented here plus all that I have *not* written. And it is precisely this second part that is the important one" (*Prototractatus,* p. 15).

Perhaps it will not be inappropriate to approach the *Tractatus* with a strategy fit for mystical writings, by attempting to ascend through stages to the ultimate vision awaiting the lucky ones right after proposition 7.

The first circle: links with the past

The semantics described in Wittgenstein's *Tractatus* is somewhere between Frege's and Russell's both in problems and in solutions, but on both scores, it is closer to Russell's. Even though the *Tractatus* talks extensively of *Sinn* and *Bedeutung,* these expressions do not designate, as in Frege, two semantic categories that apply to every linguistic unit. They refer, instead, to specific semantic elements: Not everything has a *Bedeutung,* only names do; and not everything has a *Sinn,* only pictures do. Wittgenstein did not think about semantics in Frege's way. From the very beginning he thought that the ultimate constituents of what we say are also the ultimate furniture of the world – the objects.[1] It is therefore worth considering more closely the links with Russell.

The Russellian conception of analysis and the associated doctrine of a perfect language are clearly part of Wittgenstein's view. "In statements (*Sätze*)," Wittgenstein asserted,

> thoughts can be so expressed that to the objects of the thoughts correspond the elements of the sentence (*Satzzeichen*). These elements I call 'simple signs', and the statement 'completely analyzed'. The simple signs employed in statements are called 'names'. The name means (*bedeutet*) the object. The object is its meaning (*Bedeutung*). . . . The name represents the object in the statement. (*Tractatus* 3.2–3.22)

In 1915 he wrote in his notebook:

> It is clear that the constituents of our statements can and should be analyzed by means of definitions, if we want to approximate the real structure of the statement. *In any case, then, there is a process of analysis.* . . . Analysis makes the statement more complicated than it was; but it cannot and ought not to make it more complicated than its meaning (*Bedeutung*) was to begin with. When the statement is as complex as its meaning, then it is *completely* analyzed. But the meaning of our statements is not infinitely complicated. The statement is the picture of the fact (*Tatsache*). (*Notebooks,* p. 46)

Frege would have gladly endorsed this picture-theoretic conception had it been associated with the domain of senses. But Russell had shifted the picture theory to the range of *Bedeutungen,* and Wittgenstein followed.

The link between Russell and Wittgenstein should not be exaggerated, however. Russell had been swept from one philosophical standpoint to another, with little consistency of purpose through the changes. In fact, when Wittgenstein first met him, Russell was in the process of unwittingly dismantling the core structure of the semantic project that he had eagerly promoted ten years earlier. The problem concerned Russell's theory of the proposition.

Since 1898 Russell had been trying to give shape to a theory of the proposition, with little fortune. Russell's early enthusiasm for relations had led him to conclude at first that if we believe (assume, etc.) a proposition, there must be an ego, a relation of belief (assumption, etc.), and a thing, the proposition in question. As we know, he had begun by acknowledging Moore's identification of propositions with facts; but as he tried to determine what a false proposition might be, he slowly came to think that there could be no such thing. If there were false propositions, there would have to be not only things such as *that Napoleon was a General* but also things like *that Napoleon was defeated at Marengo*. The former objects are surely there, but where are the latter?

For a while Russell entertained the possibility that the ultimate furniture of the world included not only facts but also "objective non-facts" or "fictions" (see "On the Nature of Truth," p. 46). In his 1904 review of Meinong's *Über Annahmen* Russell explained that there are two basic approaches to knowledge: One says that "knowledge is the affirmation of a true complex, error that of a false one" ("Meinong's Theory of Complexes and Assumptions," *Essays in Analysis*, p. 63); the other contends that there are no false complexes, and therefore "judgment has no object except when the object is a *true* proposition" (p. 63). Russell thought it obvious that there were true propositional complexes (Moore's facts). He also thought that the analysis of a propositional attitude should not depend on the truth value of its target. If, for example, belief is a two-place relation between an ego and a proposition when the proposition is true, it must also be a dyadic relation when the proposition is false. Otherwise there would be intrinsic features of belief that would allow us to detect all falsehoods a priori. So he was at first inclined to go along with the first approach.

But what Russell was pleased to describe as his "vivid sense of reality" (*Russell's Logical Atomism*, p. 79) prevented him from holding such a view for long. A few years later he explained:

> [Meinong's] view is, that there is an entity, namely the 'proposition' (*Objektiv*), to which we may have the dual relation of assumption or the dual relation of belief. Such a view is not, I think, strictly refutable, and until I had discovered the theory of 'incomplete symbols' I was myself willing to accept it, since it seemed unavoidable. Now, however, it appears to me to result from a certain logical naivete, which compels us, from poverty of available hypotheses, to do violence to instincts which deserve respect. (*Theory of Knowledge*, p. 108)

The instinct that deserved respect is that there is no such thing as, for example, *that Napoleon was defeated at Marengo*. So there are no false propositions. Therefore, propositional attitudes may not be analyzed as relations to propositions; therefore (*sic*), there are no propositions.

In order to mitigate the effects of this "final solution" in the field of semantics, Russell recalled that the incomplete-symbol strategy helps make sense of situations in which the category of symbol plays a semantic role even though (in Russell's sense) it "means" nothing. By 1910 he had concluded that propositions are incomplete symbols and that "some context is necessary before the phrase expressing a proposition acquires a complete meaning" (*Theory of Knowledge*, p. 109). What he meant by this is hard to explain without confusing use and mention or something else; but, roughly speaking, the idea was as follows.

Russell thought that the question "What is a proposition?" suffers from the same problem as "What is a class?" or "What is the round square?" It cannot be answered. The closest we can get to finding an answer involves looking at a broader context. The idealists were, once again, right in thinking that Russell's and Moore's propositional complexes were "false abstractions." In *Principia* Russell explained that

> a "proposition," in the sense in which a proposition is supposed to be *the* object of a judgment, is a false abstraction. . . . That is to say, the phrase which expresses a proposition is what we call an "incomplete" symbol; it does not have meaning in itself, but requires some supplementation in order to acquire a complete meaning. This fact is somewhat concealed by the circumstance that judgment in itself supplies a sufficient supplement, and that judgment in itself makes no *verbal* addition to the proposition. Thus, "the proposition 'Socrates is human'" uses "Socrates is human" in a way which requires a supplement of some kind before it acquires a complete meaning; but when I judge "Socrates is human," the meaning is completed by the act of judging, and we no longer have an incomplete symbol. (p. 44)[2]

Thus, the proper context from which they cannot be detached is judgment or belief. That is why the closest thing there is to a theory of the proposition is a theory of judgment.

The new theory was given its first full presentation in 1910. Its basic ideas are two: There is no more to propositions than there is to judgment, and judgment is not a relation between the judging mind and the proposition but a many-termed relation between the mind and the diverse constituents of what was thought to be the proposition. Russell still believed, with Brentano, that "in all cognitive acts . . . the mind has *objects* other than itself to which it stands in some one of these various relations" ("On the Nature of Truth and Falsehood" [1910], *Philosophical Essays*, p. 150; my italics). Since judgment or belief is listed as one of those cognitive relations, the principle applies to it as well. When I understand, entertain, believe, judge, or assert that John is tall, I do not stand in a relation of any sort to Bolzano's *Satz an sich*, or to the Fregean sense of 'John is tall', or to Meinong's *Objektiv*, or to Russell's and Moore's old proposition *that John is tall*. I stand in relation only to what used to be regarded as the constituents of that proposition (and are still regarded as the constituents of the fact, if the judgment is true). What obtains, then, is a relation not between me and *that John is tall* but between me, and John, and tallness.

This approach is at odds with the purposes that had inspired Russell's philosophy in the first place, for it immediately entails the surrender to psychologistic semantics. Russell's original project was to oppose psychologism and subjectivism in logical or semantic matters. The strategy had been to develop the basic intuition that what we believe and know has a certain type of independence from the human mind; that what two people believe (even when false) might be the same; that logical relations of inference, consequence, and the like may – indeed, should – be analyzed without reference to the judging mind. The central tenet of the project had been that the nature of and the relations among the things that we may say are quite independent of whether anyone has said or will ever say them or hold any propositional attitude toward them.

If Russell's new theory of the proposition was correct, then these were vain dreams. Logic could not be regarded as a theory concerning the inferential relations of arbitrary propositions but, at best, of those that have and will be judged. No sense could be made of the idea that some propositions are a priori and others not; whatever sense there might be in this must be derived from properties of judgment – an idea that idealists had been advocating all along. Truth and falsehood could not exist in a world without minds. Russell did not miss this consequence. In 1912 he noted that it was "fairly evident that if there were no beliefs there could be no falsehood, and no truth either" (*The Problems of Philosophy,* p. 120). If so, no sense could be made of the truth value of an unasserted antecedent of an implication. Even if one extends Russell's doctrine beyond judgment to other propositional attitudes, it still follows that logic makes sense only when there are minds to judge or to engage in propositional attitudes. If truth and falsehood are mind-dependent, the type of semantics that Gödel and Tarski would develop in later years (Chapter 16) would have to be based on philosophical psychology and involve a theory of propositional attitudes. By 1910 Russell had driven himself full circle back to the psychologistic stage from which Brentano had initiated the march of the Austrian-realist tradition.

Russell seems to have sensed that something was amiss. In the manuscript of his *Theory of Knowledge* that he showed Wittgenstein in 1913, he let his old instincts loose for a moment when he recalled that "it is fairly obvious that the truth or falsehood which is attributed to a judgment or statement is derivative from the truth or falsehood of the associated proposition" (*Theory of Knowledge,* p. 108). But in the next two paragraphs he explained that "in my opinion" neither true nor false propositions are entities, and both must be regarded "as alike unreal, i.e. as incomplete symbols" (p. 109). A few pages later Russell's opinion was bracketed in an attempt to make these two doctrines consistent.

Russell's attempt to breathe some life into his defunct propositions began with the psychologistic assumption that the proposition must emerge from propositional attitudes – not necessarily from judgment but, preferably, from the one propositional attitude presupposed in all others: understanding. Russell analyzed our understanding that xRy as

$$U(S, x, R, y, \phi),$$

where U is the relation of understanding, S is the subject, x, R, and y are the constituents of the (quasi-resuscitated) proposition that xRy, and ϕ is the form of that complex. From this mind-dependent notion Russell defined the proposition that xRy as

There is a U and an S such that $U(S, x, R, y, \phi)$.

This, he concluded, "is the same for all subjects and for all propositional relations which we should regard as concerned with the same proposition. Thus there is no formal obstacle to defining this as the proposition" that xRy (p. 115).

Even though there is no "formal" obstacle (whatever that might mean), the definition fails to serve the purposes propositions were intended to serve in the semantic tradition – as Russell proceeded to point out. The main problem, he explained, "is that we cannot be sure that there are propositions in all cases in which logic would seem to need them" (p. 115). We would like logic to tell us about the inferential relations of propositions, quite independently of whether anyone has thought of them or has been acquainted with their specific constituents – a condition Russell imposed on understanding. Russell concluded that "we cannot know of the existence of propositions other than those that have been actually thought of" (p. 116). With that he let the matter drop.[3]

This is the point at which Wittgenstein entered. When Russell gave his young student the manuscript of his *Theory of Knowledge,* the topic on which Wittgenstein focused his criticism was the theory of judgment. We know little about the specific nature of his objections (see the subsection on types and forms, this chapter); but we do know that shortly after raising them, Wittgenstein wrote to Russell, "I am very sorry to hear that my objection to your theory of judgment paralyses you. I think it can only be removed by a correct theory of propositions" (Wittgenstein, *Letters,* p. 24). And when, a few months later, Russell insisted on his theory of propositions as incomplete symbols, Wittgenstein impatiently replied that "the proposition . . . is *of course* not an incomplete symbol" (*Letters,* p. 35).

By 1913 Russell had left semantic monism in ruins. On Wittgenstein's view, the first thing that had to be done in order to reconstruct semantics was to find the right answer to the question, What is a proposition? The key to that question, he thought, was a notion that neither Russell nor Frege had taken seriously enough: form.

The second circle: objects, facts, and their forms

Like Frege and (the early) Russell, Wittgenstein thought that as we analyze statements, we usually find that other statements occur as their constituents (as in 'John believes *p*' or '*p* and *q*'). The analysis is not concluded until we reach the ultimate simples, which are, of course, not statements but something else. Therefore, before we reach the endpoints of propositional analysis, there must come a stage at which we encounter the simplest complexes that convey information. This frontier between the range of statements and that of its constituents, between what we can say and mere names, is what Wittgenstein called the elementary statements (*Elementarsätze*). The radical character of Wittgenstein's reductionism is displayed in his doctrine that *all* information that could possibly be conveyed is already present at that level. Anything that can be said, can be said by means of elementary statements. From this it follows that semantics need not worry about anything beyond the domain of elementary statements and their semantic correlates.

According to Wittgenstein's reductionism, the only symbols that contribute constituents to propositions (that are not incomplete, as Russell would put it) are the constituents of elementary statements,[4] the "names," as Wittgenstein called them, whose semantic correlates are the ultimate constituents of the world, the "objects." In order to understand what an elementary statement is, we must understand what its form is. This, in turn, depends on understanding the form of its constituent names and their links with the forms of the objects they name.

As we know, Frege and Russell were struck by the power of a distinction between those constituents of what we say that can stand on their own two feet and those that cannot. Russell thought that the constituents of propositions are either concepts or objects; Frege, that they are either saturated (objects) or unsaturated (senses of concept words or relation words). Wittgenstein's decision to use a single word, *'Gegenstand',* for *all* ultimate constituents reflects his conviction that there is something wrong in the bipartition that both Frege and Russell had accepted. One of the favorite questions among Wittgenstein scholars used to be whether the *Tractatus* was nominalistic or realistic, which was sometimes translated into the question of whether Wittgenstein admitted only Fregean objects as constituents of facts or whether he allowed concepts as well. There can be little doubt that had the question been put to Wittgenstein, his answer would have been that *neither* concepts *nor* objects in Frege's sense are constituents of anything in his semantics.

Objects

Frege's distinction between concept (function) and object was, as we know, introduced as an *alternative* to the traditional subject–predicate analysis; so was Russell's. Few people have been more vocal than those two philosophers in their denunciation of the dangers and confusions implicit in the subject–predicate picture of semantics. It is therefore puzzling to find Wittgenstein saying that the concept–object distinction must be rejected because it is identical to the old subject–predicate distinction (*Philosophical Remarks,* pp. 119, 136; *Philosophical Grammar,* pp. 202, 205). Fregean concepts, he thought, were the properties of substrata (*Philosophical Remarks,* p. 120; *Philosophical Grammar,* p. 202), though, of course, he did not say why he thought so. He found this unacceptable.

A hint of what may have been in Wittgenstein's mind emerged during one of his conversations with his Viennese audience in 1929 (Waismann, *Vienna Circle,* pp. 41–2). Basically, what he told them is this: As long as we think only of the misleading picture of things implicit in ordinary language, we will take the subject–predicate form as dominant, and we will naturally be tempted to conclude that the world consists only of two sorts of things, those that are designated by subject expressions and those that are designated by predicate or relation expressions. *That* is the basic intuition behind the Frege–Russell distinction between concept and object.[5] As soon as we look beyond ordinary language to other forms of representation, Wittgenstein explained, the appeal of this dichotomy vanishes. For example, ordinary language would describe this room by alluding to tables, chairs, and so on, as well as to relations among them. But consider the following alternative description: The *phenomenal* room is now to be described by means of a two-dimensional surface given by an analytic equation and by assigning phenomenal colors (à la Carnap) to each of the points on the surface. This method of representation, according to Wittgenstein, is no less faithful to the basic facts than the first one, since it is surely closer to the level of "primary language" (as discussed later). When this form of representation is chosen, the temptation to talk about objects and concepts vanishes. Which are the objects and which the concepts? There must, of course, be ultimate elements of analysis, but it is not correct to partition them into the two Fregean categories.

These ultimate elements of analysis, these *indefinables,* are precisely Wittgenstein's objects.[6] The fact that he chose to assign all of them to a single category reflects his disagreement with Frege's and Russell's dichotomy. But it might also be taken to imply that the ultimate constituents of propositions are not of two radically different sorts but only of one, [7] that is, that Frege and Russell erred in finding too many varieties of indefinables. The truth is rather the opposite. Wittgenstein's main complaint against the Frege–Russell bipartition was that it sins even in trying to establish a priori the variety of categories of forms that there may be. This point emerges quite clearly when we turn to the central philosophical feature of Wittgensteinian objects, their form.

The form of objects

Wittgenstein said virtually nothing directly about the character of objects. There is no example of an object in the *Tractatus* and not even a hint of what these things might be.[8] But there is a great deal of talk about something *of* objects, their form. Understanding this may be the closest we can hope to come to understanding what objects are.

Perhaps the best way to approach Wittgenstein's doctrine of the form of objects is to compare it with Frege's doctrine of the unsaturatedness of first-level concepts. Even though Wittgenstein's objects were quite unlike Frege's, they were like Frege's concepts in that they were also unsaturated and in need of completion. Unlike Frege's concepts, however, they possessed a much more discriminating nature, a greater unsaturatedness, as it were. Let us examine each of these two features in turn.

It used to be said that the Tractarian picture of a proposition as a sequence of names was indefensible since 'John, Peter, Mary', for example, does not say anything. If we put a number of Fregean objects together, what we get is not a claim but a bunch of objects. But Wittgensteinian objects have "holes," like Frege's concepts, so that when we put them together, there may emerge a unit of a new, nonobjectual sort. Wittgenstein preferred a different analogy for expressing the same point: He compared objects to the links of a chain (*Tractatus,* 2.03), and he explained to Ogden that the point of the metaphor "is *that there isn't anything third* that connects the links but that the links *themselves* make connexion with one another" (*Letters to C. K. Ogden,* p. 23). In a similar vein, in 1929 he examined the question of how to analyze a proposition of the type *color c is located at place p* and observed:

> It is clear that there is no relation of 'being at' obtaining between a color and a place, at which the color "is located." There is no in-between term between color and space. Color and space saturate (*sättigen*) each other. And the way in which they penetrate each other makes the visual field. (Typescript of vols. 1–4, p. 5, Wittgenstein Papers, 208)[9]

There is a crucial difference between Frege's concepts and Wittgenstein's objects, however. The holes in Frege's first-level concepts have a unique form, while those in Wittgenstein's objects have a plethora of forms. This metaphor is the reflection of a point Wittgenstein first made in connection with the doctrine of types. We must pause here to remind ourselves of how Russell came upon that doctrine and what its point was.

Types and forms

Russell's paradox, as stated for classes, could be solved by means of the no-classes theory: If there are no classes, there are no class paradoxes.[10] The contradiction was derived from the assumption that there is a certain class, and the conclusion was therefore that there is no such class. But for the parallel case of the intensional contradiction, the relevant assumption is not one of existence but of meaningfulness; all that is assumed is that $F(F)$ makes sense, that it is either true or false. What we seem to have here is the first conclusive proof that some sentences that appear to make sense are, in fact, senseless.

In Wittgenstein's view, Russell's intensional paradox was not so much a dreadful problem to be solved for the sake of consistency, but rather an extraordinarily fortunate accident that brought to the surface a previously undetected confusion in traditional conceptions of language. If $F(F)$ is not, in fact, significant, there must be a host of other similarly misleading expressions, equally lacking in meaning. What we need, Wittgenstein thought, is *not* a trick to avoid the paradox, like Russell's theory of types, but a precise diagnosis of the disease upon which Russell had accidentally stumbled.

As Wittgenstein read it, Russell's theory of types attempted to solve the difficulty by identifying a system of rules that prevented the construction of expressions of the delinquent (meaningless) sort. It was, he thought, as if someone had decided to suspend the use of modus ponens in a region of argument in order to prevent the derivation of a contradiction; the theory attempted to prescribe, where what we had to do was more in the nature of describing. We did not need to impose rules of types. God had already done that. All we had to do was recognize His works in an appropriate symbolism, one that would make the type rules superfluous. Wittgenstein's first statement of this attitude appeared in a letter to Russell of January 1913; its point must be seen in the context of the views Russell was holding at the time.

In his *Theory of Knowledge* Russell had explained that if we understand *aRb,* then must we be acquainted not only with *a, b,* and *R,* but also with a fourth elusive element (not a thing, not a constituent of the proposition, but something nonetheless), the form of dyadic relational complexes, represented by '*xZy*' (where '*x*', '*Z*', and '*y*' are variables of the appropriate types). Here is Russell's account of why we understand propositions we have never seen before:

> Let us suppose that we are acquainted with Socrates and with Plato and with the relation "precedes," but not with the complex "Socrates precedes Plato." Suppose now that someone tells us that Socrates precedes Plato. How do we know what he means? It is plain that his statement does not give us *acquaintance* with the complex "Socrates precedes Plato." What we understand is that Socrates and Plato and "precedes" are united in a complex of the form "*xRy*", where Socrates has the *x*-place and Plato has the *y*-place. It is difficult to see how we could possibly understand how Socrates and Plato and "precedes" are to be combined unless we had acquaintance with the form of the complex. (*Theory of Knowledge,* p. 99)

As usual whenever a question about "understanding" arose, Russell answered in terms of acquaintance. To understand a word is to know its meaning, and to know a meaning is to be acquainted with something, that meaning (e.g., *The Problems of Philosophy,* p. 104). This might be understood in a harmless manner, as a colorful rephrasing of the claim that we understand something. Russell did not mean it that way. For him, references to acquaintance were intended to be *explanatory:* To understand a meaning is to relate to it the way I relate to a sense datum when it is present in consciousness. And this was true not only of particulars and universals, the constituents of propositions and facts, but also of forms.

It was a natural concomitant of this view that even though forms are not constituents of propositions, they are quite independent of the entities that could be correlated "in those forms." Thus, "we might understand all the separate words of a sentence without understanding the sentence. . . . In such a case we have knowledge of the constituents, but not of the form. We may also have knowledge of the form without having knowledge of the constituents" (*Our Knowledge of the External World,* pp. 52–3). One may gather from scattered information that before 1913 Wittgenstein held a similar view (*Notebooks,* p. 117; *Letters,* pp. 13, 19). He thought that objects (particulars) and relations (in the old sense of these terms) were constituents of propositions, and independently of these things, and somehow "above" them, there was something else that Russell called the "form of the complex" and Wittgenstein the "copula."

In his letter to Russell of January 1913, Wittgenstein reported a change of mind on atomic complexes (roughly, in the direction of Frege's analysis) and then added:[11]

Every theory of types must be rendered superfluous by a proper theory of the symbolism: For instance if I analyze the prop Socrates is mortal into Socrates, Mortality and $(Ex, y)\,\epsilon_1(x, y)$ I want a theory of types to tell me that "Mortality is Socrates" is nonsensical, because if I treat "Mortality" as a proper name (as I did) there is nothing to prevent me to make the substitution the wrong way round. *But* if I analyze [it] (as I do now) into Socrates and $(Ex)x$ is mortal or generally into x and $(Ex)\phi(x)$ it becomes impossible to substitute the wrong way round, because the two symbols are now of a different *kind* themselves. (*Letters,* p. 19)[12]

Wittgenstein was telling Russell that he no longer agreed with the arguments in *Principles* that concepts can occur both in subject and in predicate position; revealingly, his point concerned substitution, which was, as we shall soon see, the most apparent feature of the form of objects. He was offering what would later become his definitive objection to Russell's theory of judgment (or, better yet, of the proposition): If there were no features intrinsic to the constituents of a complex that prevented them from being interchangeable, then in a complex such as

(*) that this book is on top of this table,

it should be possible to put a penholder where the relation *is on top of* stands. The result is the complex

(**) that this book penholders this table.

Equivalently, the sentence displayed in (**) and expressing the complex in question should be meaningful; it clearly is not (*Notebooks,* p. 96).

As Wittgenstein saw it, we do not need type rules to cope with things such as (**); rather we need to recognize that there simply are no such things as the complex allegedly described in (**). There are no such things because the constituents of that alleged complex have forms; that is, they are so constituted as to allow for certain sorts of links but not others, and *consequently,* syntactical appearances notwithstanding, there is no statement that represents it (*Notebooks,* 2.9.14, p. 2). If a linguistic system is like English in that it allows for configurations of symbols that appear to express the alleged complex (**), then what we need is not to add to that system a theory of types, but to throw that system away and replace it by one that does not allow such misleading symbolic configurations.

The idea was illustrated at a meeting with Vienna circle representatives, when Wittgenstein agreed that the following was a good explanation of the intended point: In a normal language such as German or English, not only can we formulate the statements 'A is north of B' and 'B is north of A' but we can also formulate their conjunction, and someone could be misled into thinking that there is a conceivable circumstance that this new configuration represents. The Russellian strategy for avoiding this difficulty would be to add a rule forbidding the introduction of conjunctions of that sort. The Wittgensteinian strategy was to adopt a different system of representation, in this case a map, in which one could still say the meaningful, that A is north of B, or that B is north of A, but could no longer display a configuration of symbols corresponding to the old 'A is north of B and vice versa' (Waismann, *Wiener Kreis,* pp. 79–80).[13]

The preceding considerations indicate how, from Russell's contradiction together with his attempted solution of its intensional version, Wittgenstein was led to think that objects, the ultimate constituents of propositional analysis and the ultimate furniture of the world, must have forms. These forms might be regarded as certain propensities to attach to or reject other objects, and in this respect, Wittgenstein's objects were very much like Fregean concepts. But we have also said that in a sense, they are unlike them. It is now time to see why.

It is ironic that several authors have seen in Frege's modest hierarchy of concepts an anticipation of the theory of types. There is, in fact, no philosopher whose thought has been more hostile to type-theoretic thinking than Frege. In the famous hierarchy of concept levels, he paid attention only to the first two stages and expressed a clear reluctance to go beyond it. Much more to the point, Frege's first-level concepts, the ones that apply to objects, are strikingly undiscriminating – absolutely any object in the universe can fit into a concept and saturate it. From the first inaccessible cardinal to the president of France, from a particular shade of color to truth values and sneezes, every object in the universe is a proper argument for the concept *is greater than 3*. If, as Anscombe once put it, Frege's concepts have holes, the first thing to say about these holes is that they all have the same "form" and, correspondingly, all objects must also have the same "form." Indeed, in Wittgenstein's technical sense of form, it turns out that this can be proved; for even though Wittgenstein said next to nothing about what form *is*, he did explain himself on what it is to have the same form: Two objects have the same form when, upon replacement of one of them by the other in an arbitrary state of affairs, the result is still a state of affairs. Thus, explaining to Waismann and Schlick that there cannot be one subject–predicate form, he noted that "if there were only one, all substantives should be intersubstitutable [*salva significatione*] and so should all adjectives, for all intersubstitutable words belong in a single class" (Waismann, *Wiener Kreis*, p. 46).[14]

Similarly, in his *Theory of Knowledge* (which Wittgenstein read and, by some accounts, "demolished"), Russell wrote, "two complexes have the same form if the one becomes the other when the constituents of the other are successively substituted for the constituents of the one" (p. 113). Two decades later this notion of form would make a notorious appearance in a different branch of the semantic tradition. In his monograph on truth Tarski explained:

> [The] concept [of semantical category], which we owe to E. Husserl, was introduced into investigations on the foundations of the deductive sciences by Lesniewski. From the formal point of view this concept plays a part in the construction of a science which is analogous to that played by the notion of type in the system *Principia Mathematica* of Whitehead and Russell. But, so far as its origin and content are concerned, it corresponds (approximately) rather to the well-known concept of part of speech from the grammar of colloquial language. Whilst the theory of types was thought of chiefly as a kind of prophylactic to guard the deductive sciences against possible antinomies, the theory of semantical categories penetrates so deeply into our fundamental intuitions regarding the meaningfulness of expressions, that it is scarcely possible to imagine a scientific language in which the sentences have a clear intuitive meaning but the structure of which cannot be brought into harmony with the above theory. ("The Concept of Truth in Formalized Languages," p. 215)

He then offered the following explication for the notion of 'same semantical category':

> Two expressions *belong to the same semantical category* if (1) there is a sentential function which contains one of these expressions, and if (2) no sentential function which contains one of these expressions ceases to be a sentential function if this expression is replaced in it by the other. . . . By applying the principle of abstraction, all the expressions of the language which are parts of sentential functions can be divided into mutually exclusive classes. (p. 216)

If Tarski's sentential functions are replaced by Wittgenstein's elementary statements, we have the linguistic correlate of the notion of form: To each semantic category there corresponds a form and (in the ideal case) vice versa.[15]

The form of statements

When objects saturate each other in the fashion appropriate to their forms, a certain whole emerges. This whole is the closest thing in Wittgenstein's semantics to Russell's propositional complex.[16] It is what Wittgenstein called a *Sachverhalt* (henceforth Sv), a (possible) elementary circumstance,[17] "what corresponds to an Elementarsatz if it is true" (Letter to Russell of 19 August 1919, *Letters*, p. 72).

Like his predecessors, Wittgenstein had distinguished between language and what it represents, and he sought through analysis the constituents of linguistic and ontic complexes. The minimal molecule at the level of language is the elementary statement, and at the level of semantic it is the Sv. The constituent atoms of these molecules are names and objects, respectively. A particular statement and a particular Sv are configurations or "structures" (as Wittgenstein called them) of their constituent atoms. None of this seriously challenged tradition. But pre-Wittgensteinian monistic semantics was based on the idea that each linguistic configuration should map isomorphically the semantic configuration it aims to express and that the essence of linguistic representation lies in preservation of structure. This, Wittgenstein thought, is a superficial view of the matter, for it ignores form. At both the linguistic and the "worldly" level, the forms of the constituents induce a form on the new holistic units they constitute. "The logical form of the statement must already be given in the forms of its constituents" (*Notebooks*, 1 November 1914 (6), p. 23) and "If I know an object I also know all the possibilities of its occurrence in elementary circumstances" (*Tractatus*, 2.0123).

In 1929 Wittgenstein explained to Waismann, "A statement may be varied in as many dimensions as there are constants in it. And that is the number of dimensions of the space in which the statement lies" (Waismann, *Wiener Kreis,* p. 91). For example, if the statement under consideration is the concatenation of the names *a* and *b,* then the logical (or semantic) space in which it can be represented is two-dimensional, and the coordinate axis associated with each of the objects (or names) associates with each of its points an object of the same form. If John and *is tall* were objects and they were the only objects in *John is tall,* then this would be represented by a point in a two-dimensional semantic space; another point would represent, for example, *Mary is short,* assuming that Mary has the same form as John and that *is tall* has the same form as *is short.* If, following Wittgenstein, we call "the connection of the elements of a picture . . . its structure" (*Tractatus,* 2.15), then the form of a statement is the class of all structures that may be generated from it by appropriate substitutions, or if we want to insist on an aphorist style, we might say of the connection of the elements of a picture that "its possibility is its form of representation" (2.15) or, better yet, "form is the possibility of structure" (2.151).

Thus, a particular Sv displays a particular structure; the nature of its constituent objects determines a class of other structures obtained by substituting equiform objects. This class of all the *possible* structures associated with the original structure displays the form of the Sv, which underlies its visible structure. It is this form, rather than the surface structure, that a correct linguistic representation must match. Not only must a propositional picture be capable of offering an isomorphic map of the specific arrangement of the objects in the Sv depicted; it must also be modifiable, in conformity with its form, so as to represent isomorphically all other structures of the same form. To the possible connections of objects there must correspond isomorphic possible connections of their names. What Wittgenstein called "multiplicity" was, apparently, this formal element. The requirement that a symbolic system and a corresponding reality have the same multiplicity demands that the symbolic system and its objective correlate have exactly the same numbers of elements and that these are capable of exactly the same structural arrangements.

The motivation for the requirement that an appropriate symbolism have the same multiplicity as what it symbolizes is that the other two alternatives have evident drawbacks. If the multiplicity of the symbolic system is smaller than that of what it represents, there will be possible circumstances we will not be able to describe. If the multiplicity is *greater,* the problem is more familiar – it is called 'philosophy'. All of philosophy (up to, and perhaps including, Wittgenstein) had consisted of attempts to say things that cannot be said. Good philosophy attempts to say what can be shown, the *sinnlos;* bad philosophy attempts to say what cannot even be shown, the *unsinnig,* the utter nonsense. Most philosophy had been bad philosophy, based on confusions concerning language. These confusions were roughly of the sort displayed by Russell's paradox: The language we use has a greater multiplicity than what it talks about. We can therefore form expressions whose syntactic appearance is like that of perfectly meaningful claims, but from them we are led to some form of chaos. *All* bad philosophy is the use of expressions like Russell's '*F*(*F*)'. A "correct *Begriffsschrift*" would make bad philosophy impossible; it would also display the way to show and only show what good philosophy tries to say.

Form is the key to meaning. Indeed, it may be said that Wittgenstein held a correspondence theory of meaning: Statements need not correspond to reality the way true sentences do, but there must nonetheless be some element in the symbol that is identical with some element in the circumstances symbolized, in order for the former to symbolize the latter; this element is form. A name, a symbol for an object, can represent that object only if it has the same form as the object: "The forms of the entities are contained in the form of the proposition which is about these entities . . . the proposition contains the form of an entity which it is about" (*Remarks on the Foundations of Mathematics,* p. 36). And a thought can represent a reality only if they have the same form: "What the picture and reality must have in common in order to be able to represent it after its manner – either truly or falsely – is its form of representation" (*Tractatus,* 2.17; see also *Notebooks,* 20 November 1914, p. 15).

Let us pause to survey this territory from the vantage point we have now achieved. There are basically three things that concern the Wittgensteinian logician in the treatment of statements and what they represent: their constituents, their form, and the associated facts. At the level of language, the reductionist analysis took us first from arbitrary statements down to elementary statements, and beyond them to their simple constituents, the names. Then we moved to their worldly correlates – the objects – examined their form, and started moving upward toward Sv's. Yet the world is neither objects nor forms but everything that is the case, the facts. Thus, we have to move upward beyond elementary circumstances.

One of the seven Tractarian pillars (*Tractatus,* 2), says, "What is the case, the fact, is the obtaining of Sv's"; not Sv's, but their obtaining. Names and statement forms are essential in that they provide the scaffolding required for the purposes of communication, but those purposes are ultimately fulfilled by something else, by statement facts in which those names are constituents and those forms are displayed. The reason for this is that what we want to communicate is what we regard as *fact,* and Wittgenstein thought that this could be depicted only by means of other facts. Thus, objects are represented by names (which may have been endowed with a form of sorts through grammatical rules), forms by forms, and facts by facts. If, *per impossibile*, John and Tall were objects, they would be represented by names like 'John' and 'is tall', their forms by the forms induced on 'John' and 'is tall' through linguistic rules (although the ideal symbolism, as in the map analogy, would not require such rules), and finally, the fact that John is tall would be represented by some fact concerning the symbols 'John' and 'is tall'. A correct symbolism would allow for substitutions of all names of the same form for 'John' in 'John is tall', but not for 'tall'. The symbol 'John is tall' would have the same form as the fact that John is tall, *because* the symbols 'John' and 'is tall' would have the same form as their corresponding objects. In this language there would be no symbol for forms. There would be no symbol for the form of the fact that John is tall, for example. But that form would nonetheless be displayed by the language. The proposition that says John is tall would, at the same time, show what the form of that fact is, since its form is precisely the form of that fact: Its possibilities of meaningful rearrangement are a perfect replica of the possibilities of rearrangement of the corresponding objects.

We have been talking about the representation of facts by means of linguistic facts, but what we want to be able to represent includes things such as

that Hitler was wise and benevolent.

The word 'fact' may not be appropriate here. That is why in the *Tractatus* Wittgenstein did not talk of linguistic facts representing nonlinguistic facts, but said instead that only facts can represent *senses* (*Sinne*), since senses are the possibilities of the holding or not holding of an Sv. The remaining question is, What is sense? The answer depends on the notion of a *Sachlage* (henceforth Sl).

The third circle: Wittgenstein's *Sinn*

In the beginning were the objects and their forms. Having decided on which objects to create, God still had more work to do, since the world is the totality not of objects but of facts.

Sv's are of two sorts: those that obtain (*Bestehende* Sv) and those that do not (*Nichtbestehende* Sv). A fact (*Tatsache*) is the obtaining of a circumstance (*Tractatus*, 2) – not a circumstance, or even one that obtains, but the obtaining of one. An Sl is the holding or not holding of circumstances – not necessarily a fragment of the *actual* holding or not holding of circumstances chosen by God in the act of creation, but any possible one (2.11). Thus, an Sl can be seen as a fragment of a world that

God could have created or, equivalently, as a class of possible worlds (all the worlds agreeing on that particular fragment).

This notion provides the key to Wittgenstein's theory of the meaning or sense of a sentence. He explained that a picture presents (*darstellt*) a possible Sl in logical space (*Tractatus*, 2.202), and what the picture presents is its sense (2.221). Since he also thought that every statement is a picture of reality (4.01), we seem justified in concluding that the sense of a statement is always a possible Sl.[18]

Further clues to the nature of sense are given in *Tractatus*, 4.2, which offers an indirect way of identifying the sense: "The sense of the statement is its agreement and disagreement with the possibilities of the obtaining and non-obtaining of Sv's." Unscrambled the claim appears to be that if we were given all the possibilities of obtaining and nonobtaining of Sv's, all "maximal Sl's" we might call them, or all possible worlds, then to know the sense of a statement is just to know with which of those possible worlds it agrees and with which it does not – or, equivalently (by 2.222), in which of those possible worlds the statement is true. This is just another way of identifying the appropriate Sl, since the Sl in question is simply the element common to all the maximal collections of Sl's with which the statement agrees. To put the matter in more pedestrian terms, we might represent an Sl as a fragment of what Carnap called a state description, and what I have called a maximal Sl as a state description. The sense of a statement would then become the Sl in question or, alternatively, the class of all possible worlds that have it as a part. This alternative would be the one described in *Tractatus*, 4.2. Once again, one may notice the link between these ideas and semantic monism since, like Russell, Wittgenstein concludes that the object *a* occurs in the sense of '*fa*'.

To understand a statement, we said, involves knowing its form and therefore the form of what it represents. But to understand a statement *is* to know (*kennen*) its sense (4.021). When I understand a statement, I know the Sl it presents (*darstellt*). This ties up with the famous doctrine that to understand a statement is to know (*wissen*) what is the case when it is true, for the Sl that is its sense is precisely that – what is the case if the statement is true. It also explains why the sentence shows how things stand if it is true (4.022), since what the sentence shows is its sense (4.022).[19]

Closely connected with these points is one of Wittgenstein's earliest ideas, what he called at first the "bipolarity business," the root of his later verificationism. Already in a text from 1913 Wittgenstein had made the observation that "what we know when we understand a proposition is this: we know what is the case if it is true and what is the case if it is false"

(*Notebooks*, pp. 93–4). He noticed shortly thereafter that far from it being a mere accidental correlation, "the being true or false actually constitutes" the having of sense (p. 112). If an alleged statement is to be a genuine vehicle of information, if it is to be a real statement, it must have two poles, the truth pole and the false pole; that is, there must be *both* circumstances that make the statement true *and* circumstances that make it false. The absence of *either* of those poles guarantees the absence of both, because it guarantees the failure of sense. (Since sense is the possible facts common to the possible worlds in which the statement is true and since no possible fact is present in all possible worlds, both tautologies and contradictions have nothing at all as sense.)

Now we know what sense is, we know the "general form of the proposition," we know what it is to *say* something, what conditions are to be fulfilled in order for an apparent vehicle of information actually to be such. By implication, we have also determined what lies outside the sphere of sense.

The fourth circle: showing and saying

What, one may well ask, is the purpose of this baroque semantic edifice? In the semantic tradition feats of this sort had always been performed for the greater glory of science, to cleanse it from blemishes and give it as solid a foundation as possible. Why engage in this extraordinary search for the essence of the proposition? Where was the fire that fueled this amazing enterprise?

Most philosophers will agree that Wittgenstein was, after all, human, and hence was likely to be inspired by reasons that reason might not always comprehend. It is therefore worth observing that, like Prince Myshkin, Wittgenstein came to believe very early on that the truly important things in life are not among those whereof one may speak. When Parfyon Roghozin asked Myshkin whether he believed in God, the gentle prince would not give a direct answer. Some commentators (e.g., Guardini) have explained that the answer was too obvious; but it seems more plausible to think that in Myshkin's mind *any* answer to that question would have been wrong. The atheist professor who answers one way is no more confused than the Catholic theologian who contradicts him. The sort of reasoned discourse in which both engage, the domain of saying, is fit only for the lower purposes of science and other pragmatic matters. Everything that really counts, that gives meaning to life and thereby to the world, is entirely unsayable.

The *Tractatus* tells us that "all propositions are of equal value," that "in the world . . . there is no value . . . there can be no ethical propositions," that "propositions cannot express anything sublime," that "even if all *possible* scientific questions be answered, the problems of life still have not been touched at all," and that "the sense of the world must lie outside the world" (6.4–6.421, 6.52). Besides what we can say, "there is indeed the inexpressible. This *shows* itself; it is the mystical" (6.522). Even though these doctrines appear at the end of the *Tractatus*, there can be little doubt that they stood at the very beginning of the train of thought that inspired it. Wittgenstein *wanted* to establish with classical rigor that there is a special, superior place for something like that moment of insight and peace at the very beginning of Myshkin's epileptic fit, when he was overcome by "a feeling, unknown and undivined till then, of wholeness, of proportion, of reconciliation, and of ecstatic devotional merging in the highest synthesis of life." In the end, Wittgenstein did come up with a distinction between what language says and a different domain of things that we may aim to say but always unsuccessfully.

In 1919, in response to some questions Russell had raised about the *Tractatus*, Wittgenstein wrote that Russell had missed the "main contention" of the book: "The main point is the theory of what can be expressed (*gesagt*) by prop[ositions] . . . and what can not be expressed by prop[ositions], but only shown (*gezeigt*); which, I believe, is the cardinal problem of philosophy" (Letter of 19 August 1919, *Letters*, p. 71). Whether one agrees with Wittgenstein or not, there is, as we shall see, a recognizable and insightful train of thought leading from the analysis of form presented earlier to the conclusion that we cannot say what modus ponens "says," or what is "said" by anything involving formal concepts as a topic. Whether there is any link at all between the view we are about to explain and God, or the meaning of life, or any of the mystic images that Wittgenstein was hoping to rescue from the claws of Reason, is a matter I must leave for better minds to determine. The purpose here merely is to make some sense of what Wittgenstein showed about showing in the field of logic.

Two decades before the *Tractatus*, Lewis Carroll had presented a striking paradox stemming from the traditional way of thinking about logic. In his celebrated dialogue between Achilles and the Tortoise, he had the latter arguing that a valid inference can be drawn only after we have completed infinitely many tasks, for we must acknowledge infinitely many premises before we are entitled to endorse the conclusion. His point was basically this: We all agree that from 'If A then B' and A, B follows logically. But Carroll notes that if we *fail* to grant *either*

(1) $A \;\&\; A \Rightarrow B$

or

(2) $(A \& A \Rightarrow B) \Rightarrow B$,

we will not be entitled to infer

(3) B.

"Hence," he wrote, "*before* granting (3), I must grant [not only (1) but] (1) and (2)." But, in fact, if I do not grant *also* that (1) and (2) entail (3), I am still not entitled to infer (3). "Surely, my granting (3) must *wait* until I have been made to see the validity of this sequence" (Carroll, *Symbolic Logic,* p. 472). And so on ad infinitum.

The paradox is entirely spurious because of the role time plays in it. Carroll's reasoning is, in fact, a revealing illustration of the way psychology and logic were fused before objective content was detached from the mental acts related to it. Whether an inference is justified does not, of course, depend on whether we have noticed anything at all. But when the reference to time and psychology is excluded, there still remains a problem for the semantic construal of logic. If the justification of the inference from (1) to (3) requires an appeal to the logical law (2), why don't we need to appeal to further logical laws in order to justify the inference from (1) and (2) to (3); and so on ad infinitum? Doesn't the justification of logical inference involve an infinite regress?

Since Aristotle, several philosophers had recognized something like this regress, and it had been acknowledged that the regress must somehow be broken. A favorite way to break it, recently revived by Russell, was by an appeal to self-evidence or intuition. The inference from one claim to another must be backed up by a logical law; but the regress stops at the law, because we now say that it is known immediately in intuition.

In *The Problems of Philosophy,* while explaining his views on logic, Russell invited his readers to imagine a discussion between two men trying to determine what day it is:

> One of them says, 'At least you will admit that *if* yesterday was the 15th to-day must be the 16th'. 'Yes', says the other, 'I admit that'. 'And you know', the first continues, 'that yesterday was the 15th, because you dined with Jones, and your diary will tell you that was on the 15th'. 'Yes', says the second; 'therefore to-day *is* the 16th'. (p. 71)

Everyone will recognize this as an instance of the soundest sort of reasoning there is, logical reasoning. But *why* is it so solid? What are the root and the justification of its infallible accuracy? Russell explained that the argument is acceptable *only because* it can be subsumed under a very special sort of law: Its conclusion

> depends for its truth upon an instance of a general logical principle. The logical principle is as follows: 'Suppose it known that *if* this is true, then that is true. Suppose it also known that this *is* true, then it follows that that is true'. . . . This principle is really involved . . . in all demonstrations. . . . If anyone asks: 'Why should I accept the results of valid arguments based on true premisses?' we can only answer by appealing to our principle. In fact, the truth of the principle is impossible to doubt, and its obviousness is so great that at first sight it seems almost trivial. (*The Problems of Philosophy,* pp. 71–2)

In the model of logical representation developed by Frege and Russell and now widely regarded as standard, one thinks of a proper axiomatization of a discipline as involving, besides the specific assumptions or axioms of that discipline, a set of logical axioms and logical rules of inference. Both the logical and the nonlogical axioms are thought to occur as part of the same language and to differ only in the scope of their validity and perhaps also in the abstractness of the domain with which they are concerned. Thus, a correct and complete account of the inference from $A \vee B$ and $-A$ to B will involve logical laws (perhaps $((A \vee B) \& -A) \Rightarrow B)$) and logical rules (perhaps modus ponens). All of this machinery must be mobilized for the purpose of inferring B, and without any of it, the inference would not be justified. This is the point that Russell was trying to exemplify in the preceding quotation (taking account of the fact that, unlike Frege, Russell seldom paid attention to the distinction between logical rules and logical laws). In Russell's case, at any rate, it is clear that the appeal to laws is intended to play an explanatory and justificatory role: "If anyone asks: 'Why should I accept the results of valid arguments based on true premisses?' we can *only* answer by appealing to our principle" (i.e., modus ponens; *The Problems of Philosophy,* p. 72; my italics).

The problem that emerges from this perspective is, What justifies the logical laws and rules? The standard answer was intuition. But we have recently examined the most detailed account of what this would involve in the field of logic, and we have seen the view collapse into virtual incoherence. In Wittgenstein's view, an accurate diagnosis of the situation has been prevented by the propositionalist prejudice, by the insistence on looking at logic and other a priori disciplines as being expressed in statements that convey facts, just like any a posteriori statement – except that the facts in question are somehow otherworldly. If logical laws were essentially like all other statements, they would require some sort of justification, as every statement does. The best among Wittgenstein's predecessors had tried to provide that justification, but with no success. Wittgenstein's solution was to say that logic, and the a priori in general, has no justification, since it cannot even be conveyed by means of claims. There isn't a "hard" domain of a priori truths and a "soft" domain of empirical truths; all the truths there are, are empirical. That does not eliminate the a priori but rather locates it as a "hardness in the soft," as what one might perhaps call a formal aspect of all meaningful, factual discourse.

In Wittgenstein's view, the depsychologized version of Carroll's paradox suffices to establish the bankruptcy of all attempts to provide a foundation for inference on the basis of more general statements: If B will not follow from $(A \vee B)$ & $-A$ unless we explicitly add the assumption that $((A \vee B) \& -A) \Rightarrow B$, then by the same token B will not follow from *these three assumptions* unless we add the further assumption that all of these premises, properly generalized, entail B; and so on ad infinitum (see *Lectures, 1930–32,* p. 56). To the pseudo-Fregean alternative that deductive inference is justified by rules, Wittgenstein's response would be to ask for a non-question-begging justification of those rules. In later years it would become apparent that any such "justification" would presuppose the very claim it was supposed to establish and therefore also lead to an infinite regress or a circle.

Wittgenstein had focused on this problem from early on. In 1912 he had written to Russell: "Logic must turn out to be of a TOTALLY different kind than any other science" (*Letters,* p. 10). Two years later he found a way to articulate his point. He had come to think that logic must differ from other theories already at the level of how it conveys its message, of how it says what it says. He concluded that he had to draw a distinction between regular "saying," the sort of thing for which it matters how things stand, and something that superficially seems like saying but is entirely different from it. To see what he meant we must examine his way of thinking about logic.

Wittgenstein's new solution started by going back to the first stage of the problem: Do we need to "know" something besides $A \vee B$ and $-A$ in order to conclude that *B? If* Carroll and Russell were right in thinking that the inference to B requires an appeal to the corresponding logical law, and also right in thinking that the law is independent of those premises even to the point that it could be rejected while the premises are accepted, then it seems that there is no way to stop the infinite regress. Russell's appeal to self-evidence put forth as a model for the explanation of logic the idea of a special kind of relationship to a special kind of claim. Wittgenstein, however, put forth the idea that logic is fully explained by an appeal to a familiar kind of relationship to the most familiar kinds of claims. The so-called logical law that is said to justify the inference to B is not a law, and it does not justify anything at all. It is, in a sense, a necessary statement, but it is quite unlike all other statements in that what it "says" bears no relationship to how things might stand and relates only to the nature of language. For example, the logical law that was thought to justify the inference to B is something we see when we understand the premises A and $A \Rightarrow B$. In this picture of things, logic

emerges not from the intuition of extraordinary claims but from our understanding of the ordinary ones. Logic "shows itself" because it is something we recognize as soon as we understand the language we are talking, something that in no sense depends on how anything stands in the world or in the mind, but only in the understanding of language:

> If p follows from q, I can conclude from q to p; infer p from q. The mode of inference is to be understood from the two propositions alone. Only they themselves could justify the inference. "Laws of inference" which – as in Frege and Russell – are supposed to justify the inference, are senseless and would be superfluous. (*Tractatus,* 5.132)[20]

That the truth of one proposition follows from that of another is something "we see from the structure of the propositions" (*Tractatus,* 5.13). If so, it would be an error to think that the conclusion of Russell's date example "depends for its truth upon an instance of a general logical principle." The *truth* of the conclusion (as Russell knew) depends only on the facts of the matter; and our inferential knowledge of it depends only on the recognition of the particular claims that have been accepted as premises, as well as an understanding of what it is to accept something as true. Similarly, to say that Russell's version of modus ponens is "really involved . . . in all demonstrations" is as plausible as saying that the restatement in the course of an argument of a previously endorsed claim "strictly speaking" involves an appeal to an instance of the law of identity (that p implies p).[21]

What, then, about logic and its laws? "My fundamental thought," Wittgenstein explained in the *Tractatus,* "is that the 'logical constants' do not represent (*vertreten*). That the *logic* of facts cannot be represented" (4.0312). This means that we cannot think of logic as the most general laws of a *Begriffsschrift.* The "old logic," Wittgenstein had explained in 1914, "gives so-called primitive propositions; so-called rules of deduction; and then says that what you get by applying the rules to the propositions is a *logical* proposition that you have *proved*" (*Notebooks,* p. 108). The tacit presupposition is that, as in the model case of Euclidean geometry, we can also identify in logic some primitive truths from which one can deduce all others. This is the root of the misunderstanding.

Logic is radically different from every other type of knowledge because its "justification" lies not in how things stand but in the understanding of language. As I understand the language in which $A \vee B$ and $-A$ are formulated, I ipso facto recognize that whatever A and B might be, if those two statements were true, B would also be true. One might perhaps say that I "see" the law that $((A \vee B \& -A) \Rightarrow B)$, for all As and Bs; but this way of putting things is the one that makes us think of logical laws as very general truths and thereby to appeal to intuition or self-evidence: "The self-evidence of which Russell talks so much can only be discarded in logic by language itself preventing every logical mistake. That logic is a priori consists in the fact that we *cannot* think illogically" (*Tractatus,* 5.4731).

The basic point is that in this new picture of things, the focus is not on the "seeing" of a certain very general and a priori truth but on the recognition of certain meanings, on *understanding.*[22] In general, it would appear that *everything* that "shows itself" emerges entirely from our understanding of language.

What shows itself is not restricted to the field of logic. When we go beyond the molecular structure of the proposition and examine its internal constitution, we find the source of further "claims" that appear to be a priori but that also belong to the range of what shows itself once the essence of a language is recognized. These "laws" emerge from the character of formal concepts. Here bipolarity is no longer of any use in helping us detect the idiosyncratic character of these claims, since in the case of most statements about formal concepts, we seem to be able both to assert and to deny the apparent claim. We seem to be able to assert *and* to deny that numbers are objects or that there are more than two things in the universe. The reason for the unsayability of these claims emerges from the fact that 'object', 'number', and the other basic notions they involve are not really concepts but formal concepts; thus, they do not designate properties an object might or might not have, but features of the very essence of the entities involved.

Ontologically speaking (which was Wittgenstein's way of speaking at this stage), it is very hard to make any sense of all this. But if the "internal," the "formal," is interpreted as that which is determined entirely by our understanding of language, then we may perhaps interpret these as remarks concerning what we need to know in order to understand what we are talking about, as opposed to what we need to know in order to be aware of those properties it has and it might not have had.

Now we see that the decision to identify the bounds of sense, to chart the limits of language from within, was inspired not by an interest in the charted territory but in what lies beyond it. Beyond the range of meaningful discourse lies, first of all, the vast and arid land of utter nonsense, inhabited philosophy, and much else. The discovery of this domain caused great excitement among positivists in Vienna. Wittgenstein, who did not share their antimetaphysical zeal, did not much care about that outcome of his endeavor. What he did care about was the remaining domain, not the *sinnvoll* or the *unsinnig,* but the *sinnlos,* what shows itself. All good metaphysics, all of the truly important things in philosophy, had finally been given their proper place – right outside language. Language is the key to all metaphysical knowledge; it is the vehicle

through which we recognize it, but not in any previously imagined manner. These "truths" emerge not from the acknowledgment of facts – these could always be otherwise – but through the recognition of meaning. From logic to solipsism, everything that is a priori is what we must see when we know how to communicate. Language and meaning had become the very heart of metaphysics.