

FREGE'S PARADOX

by

RICHARD LLOYD MENDELSON

B.A., Cornell University
(1964)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE
DEGREE OF

DOCTOR OF PHILOSOPHY

at the

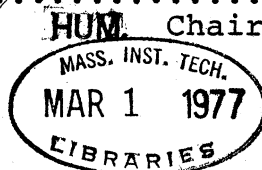
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY, 1977

Signature of Author.....
Department of Linguistics and Philosophy, January 13, 1977

Certified by.....
Thesis Supervisor

Accepted by.....
Chairman, Department Committee



FREGE'S PARADOX

by

RICHARD LLOYD MENDELSON

Submitted to the Department of Linguistics and Philosophy on January 13, 1977 in partial fulfillment of the requirements for the degree of Doctor of Philosophy

ABSTRACT

I am concerned in this thesis with Frege's problem of explaining the difference in cognitive value between $(\alpha=\alpha)$ and true $(\alpha=\beta)$.

In Chapter 1, I examine the Paradox of Identity, which purports to reveal a conflict between the view that "identity relates objects" and the fact that identity sentences differ in cognitive value. I show that this conflict results from the incorrect assumption that substitution of codenotational singular terms preserves cognitive value. Hence, the paradox exposes a problem about the notion of what is said or expressed by a sentence.

In Begriffsschrift, Frege held that cognitive value remained invariant under substitution of codenotational singular terms, so he regarded the paradox as a reductio of the view that "identity relates objects." To account for the difference in cognitive value between $(\alpha=\alpha)$ and true

$\langle \alpha = \beta \rangle$, Frege proposed that identity be taken to be a relation holding between expressions. In Chapter 2, I argue that this theory cannot be coherently maintained.

In Chapter 3, I turn to Frege's Sense/Reference theory. On this theory, truth value remains invariant under substitution of codenotational singular terms, and cognitive value remains invariant under substitution of terms having the same sense. The difference in cognitive value between $\langle \alpha = \alpha \rangle$ and true $\langle \alpha = \beta \rangle$ is attributed to the difference in sense between α and β . The main problem is Frege's ascription of sense to proper names. I develop this issue in the context of the recent debate between Saul Kripke, who attacks Frege's theory, and Michael Dummett, who defends it. I argue that Dummett's defense is insufficient.

In Chapter 4, I examine the recent attempts by Saul Kripke and Alvin Plantinga to defend Mill's theory of names, and I explore whether either offers an explanation of the difference in cognitive value between $\langle \alpha = \alpha \rangle$ and true $\langle \alpha = \beta \rangle$, when α and β are both proper names. I argue that Plantinga's explanation does not work, and that Kripke has offered only the framework of an explanation.

Name and Title of Thesis Supervisor: Richard L. Cartwright
Professor of Philosophy

ACKNOWLEDGMENTS

I should like to express my deep gratitude to Professor Richard L. Cartwright, for his good philosophical sense and for his patient and generous help in completing this thesis; to Professors George Boolos, Sylvain Bromberger, and James F. Thomson for their helpful comments on earlier drafts; to the many members of my family who assisted me: Abraham and Esther Mendelsohn, Purlaine Lieberman, Gerry Berenholz, and especially, Bobbi Lerman, who typed most of the drafts; to Lewis Schwartz and Alfred Wiener for their moral support; and finally, I should like to thank my wife, Marsha, without whose love and encouragement I could not have finished.

TABLE OF CONTENTS

Chapter		Page
1	THE PARADOX OF IDENTITY	7
	Footnotes to Chapter 1	15
2	FREGE'S BEGRIFFSSCHRIFT THEORY OF IDENTITY .	16
2.1	Introduction	16
2.2	The Begriffsschrift Theory	21
2.3	Criticism of the Begriffsschrift Theory . . .	34
2.4	A Second Interpretation of Equality of Content	42
2.5	Criticism of the Second Interpretation . . .	46
2.6	Formal Difficulties with Equality of Content	57
2.7	The Failure of the Begriffsschrift Theory . .	65
	Footnotes to Chapter 2	77
3	FREGE'S SENSE/REFERENCE THEORY	82
3.1	Introduction	82
3.2	Bestimmungsweisen and Frege's Criticism of the Begriffsschrift Theory	88
3.3	The Sense/Reference Theory	112
3.4	Frege and his Critics	128
	Footnotes to Chapter 3	158

Chapter	Page
4 Kripke's Response to the Paradox	164
4.1 Introduction	164
4.2 A Modal Paradox	170
4.3 Rigid Designators	179
4.4 Plantinga's Position	192
4.5 Kripke's Position	212
Footnotes to Chapter 4	227
BIBLIOGRAPHY	231
BIOGRAPHICAL NOTE	237

CHAPTER I

THE PARADOX OF IDENTITY

My concern in this thesis is with Frege's problem of explaining the difference in cognitive value between ' $\alpha=\alpha$ ' and true ' $\alpha=\beta$ '. In this chapter, I will say what this problem is and I will outline my plans for dealing with it in this thesis.

In "On Sense and Reference"¹ (SR for short), Frege observed that identity sentences (i-sentences, for short) differ in what he calls "cognitive value" [Erkenntniswerth]:

a=a and a=b are obviously statements of differing cognitive value; a=a holds a priori and, according to Kant, is to be labelled analytic, while statements of the form a=b often contain very valuable extensions of our knowledge and cannot always be established a priori. The discovery that the rising sun is not new every morning, but always the same, was one of the most fertile astronomical discoveries. Even to-day the identification of a small planet or a comet is not always a matter of course.²

Roughly, i-sentences of the form ' $\alpha=\alpha$ ' are true but trivial, while i-sentences of the form ' $\alpha=\beta$ ' are, if true, often interesting and informative. This observation is accurate and examples come readily to mind: 'Mark Twain = Mark Twain' is a mere truism hardly worth remarking, but 'Mark Twain = Samuel Clemens' says something of considerable literary and historical significance.

However, Frege continues, the fact that i-sentences differ in cognitive value would appear to conflict with the generally accepted view that "identity relates objects," the view, that is, on which $(\alpha=\beta)$ is understood to express that a relation, being one and the same thing as, holds between the objects denoted by α and β . For since the relation is said to hold between the objects themselves, he argues, then what is said or expressed by $(\alpha=\beta)$ --the cognitive or informational content of the sentence--is simply that the objects stand in the given relation. $(\alpha=\beta)$ and $(\gamma=\delta)$ ($\alpha, \beta, \gamma, \delta$ not necessarily distinct) would therefore say the same thing--have the same cognitive or informational content--if the object denoted by α were one and the same as the object denoted by γ and the object denoted by β were one and the same as the object denoted by δ ; for the same relation would be said to hold between the same objects. So, Frege concludes, $(\alpha=\alpha)$ and true $(\alpha=\beta)$ could not differ in cognitive value:

. . . if we were to regard equality [Gleichheit] 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). A relation would thereby be expressed of a thing to itself, and indeed one in which each thing stands to itself but to no other thing.³

On this interpretation of '=', then, it would appear that the two i-sentences, 'Mark Twain = Mark Twain' and 'Mark Twain = Samuel Clemens', both say the same thing: that a particular object, Mark Twain or Samuel Clemens--call him

what you will--is self-identical. But everything is self-identical; and so, paradoxically, on the view that $(\alpha=\beta)$ is about the objects denoted by α and β , the i-sentence, if true, appears less a significant remark about the designated objects than a trivial rehearsal of the Law of Identity.

Let us standardize some notation. Where $S\alpha$ is a sentence containing the singular term α , $S\alpha/\beta$ results upon replacing α at one or more of its occurrences in $S\alpha$ by the singular term β ; and where η is any expression that denotes something, $d(\eta)$ is the denotation of η , i.e., the entity η denotes. Now, Frege assumes in the argument that if $(\alpha=\beta)$ is to be understood as expressing that a relation holds between $d(\alpha)$ and $d(\beta)$, then the way in which the objects are specified is irrelevant to the cognitive or informational content of $(\alpha=\beta)$: if we were to replace α by a codenotational singular term (or β by a codenotational singular term), the resultant sentence would have to have the same cognitive value as the original. But Frege is clearly proceeding here on the basis of a very general assumption about what is said or expressed by a sentence, namely, that if $S\alpha$ is genuinely about $d(\alpha)$, i.e., if $S\alpha$ ascribes a property to $d(\alpha)$ or if $S\alpha$ expresses that $d(\alpha)$ stands in a particular relation, then the cognitive or informational content of $S\alpha$ is $d(\alpha)$'s having the given

property or $d(\alpha)$'s standing in the given relation. Only the object $d(\alpha)$, not the term α , is part of the content of $S\alpha$; and so, if we were to replace α at one or more of its occurrences in $S\alpha$ by any codenotational singular term β , the resultant sentence $S\alpha/\beta$ would have to have the same cognitive value as $S\alpha$. We state this generalized substitution principle as follows:

- (I) If $S\alpha$ is about $d(\alpha)$, then if $d(\alpha)=d(\beta)$, then $S\alpha$ and $S\alpha/\beta$ have the same cognitive value.

Given this substitution principle, Frege's argument is easily shown to be valid. Suppose that $(\alpha=\beta)$ is about $d(\alpha)$ and $d(\beta)$, and consider our two i-sentences,

- (1) Mark Twain = Samuel Clemens,
and
(2) Mark Twain = Mark Twain.

If (1) is true, then $d(\text{'Mark Twain'}) = d(\text{'Samuel Clemens'})$; and since (2) is obtained from (1) by replacing 'Samuel Clemens' by 'Mark Twain', then by (I), (1) and (2) must have the same cognitive value. This argument does not depend upon any characteristic of the particular names chosen, and so, we have, quite generally, that $(\alpha=\alpha)$ and true $(\alpha=\beta)$ cannot differ in cognitive value.

This argument is known in the literature as The Paradox of Identity, a reflection of the widespread misconception that the argument exposes a problem peculiar to identity. To be sure, identity is a very special relation.

It is, after all, the only predicate generally considered to belong among the logical constants, and it does give rise to some very special and difficult puzzles of its own. But this cannot be one of them, because similar arguments are easily devised for other commonplace properties and relations. The sentence,

(3) Mark Twain wrote Innocents Abroad,

for example, would ordinarily be understood to ascribe the property of having written Innocents Abroad to Mark Twain; and so, according to (I), if we were to replace 'Mark Twain' in (3) by any codenotational singular term, the cognitive value of the sentence ought to remain unchanged. Now 'the person who wrote Innocents Abroad' is such a codenotational term for, not only is it true that Mark Twain wrote Innocents Abroad, but he was the only person to have done so; and yet,

(4) The person who wrote Innocents Abroad wrote Innocents Abroad,

if not a truism, is so nearly so as to clearly differ in cognitive value from (3). In this example, we chose the substituted singular term carefully in order to parallel Frege's argument, wherein he had transformed an informative sentence into a trivial one. But there is no need to adhere to this format if we want to show that the problem about the difference in the informativeness of i-sentences can be reproduced elsewhere. It is sufficient to note

that if (1) is an informative i-sentence, then,
 (5) Samuel Clemens wrote Innocents Abroad,
 cannot have the same cognitive value as (3): the non-trivial character of (1) goes hand-in-hand with (5)'s telling us something that (3) does not, and conversely.

To recognize that the problem is thus generalizable, however, is to recognize that the substitution principle must be false. For what these examples show is that although $S\alpha$ is about $d(\alpha)$, and although $d(\alpha)=d(\beta)$, it need not be that $S\alpha$ and $S\alpha/\beta$ have the same cognitive value. The root of the paradox, therefore, does not lie with identity: the apparent conflict between the view that identity relates objects and the fact that i-sentences differ in informativeness derives from an erroneous view about what is said or expressed by a sentence. This, as we know, is the moral Frege drew in SR. His solution there was to distinguish between that which is said or expressed by a sentence, namely, a thought [Gedanke], and that which the sentence is about, or, more accurately, that which it refers to, namely, its truth value.

However, in his Begriffsschrift,⁴ Frege had drawn a different moral from the paradox. There he had concluded that identity failed to relate objects, and instead related the names denoting objects; i.e., Frege had denied that $(\alpha=\beta)$ expressed that the object $d(\alpha)$ was identical

with the object $d(\beta)$, and he had maintained instead that $(\alpha=\beta)$ expressed that the name α denoted one and the same thing as the name β . In Chapter 2, I will show that this Begriffsschrift view of identity is thoroughly incoherent. It is widely believed that the Begriffsschrift theory fails because it incorrectly attributes the informativeness of i -sentences to the difference in the expressions used to denote objects. I will argue that the Begriffsschrift theory not only provides an inadequate explanation of the difference in cognitive value between $(\alpha=\alpha)$ and true $(\alpha=\beta)$, it provides a logically incoherent explanation.

In Chapter 3, I turn to Frege's SR solution to the paradox. As I mentioned above, his solution in SR is to affirm that identity relates objects and to deny that substitution of codenotational singular terms preserves cognitive value. In SR, Frege connects the denotation of a singular term α with the truth value of $S\alpha$, and he connects the sense of a singular term α with the thought or proposition expressed by $S\alpha$. Frege's SR solution is generally believed to be correct in the case of descriptions, but it has been questioned in the case of proper names. In Chapter 3, I will focus on this question of whether proper names have meaning. I will take up this issue within the context of a debate between Saul Kripke,⁵ who attacks Frege's theory of names, and Michael Dummett,⁶

who defends it. I shall argue that Kripke's attack on Frege is devastating, and that Dummett has failed to provide any adequate response for Frege.

In Chapter 4, I turn to Kripke's own response to the paradox. Kripke holds Mill's view of names, on which proper names are said to lack meaning; he must show how, on this view of names, $S\alpha$ and $S\alpha/\beta$ can differ in cognitive value when α and β are proper names. I will try to show that Kripke, by distinguishing between metaphysical and epistemological issues, has laid the groundwork for a coherent solution to the paradox. I will be especially interested in this chapter in making clear Kripke's position, and in correcting Dummett's misrepresentation of Kripke's views. However, there remains still the problem of filling in the framework Kripke provides before one can rightly regard him as providing a solution to the paradox. But this is a task I am not able to handle here. As such, I end this thesis not with a solution to the paradox, but with the promise of a new and interesting framework in which the issues can be joined.

Footnotes

¹Gottlob Frege, "On Sense and Reference," trans. Max Black, in Translations from the Philosophical Writings of Gottlob Frege, ed. Peter Geach and Max Black (Oxford: Blackwell, 1960), pp. 56-78. "Über Sinn und Bedeutung" originally appeared in Zeitschrift für Philosophie und philosophische Kritik, 100 (1892), 25-50; it has been reprinted in Funktion, Begriff, Bedeutung: Fünf Logische Studien, ed. Günther Patzig (Göttingen: Vandenhoeck & Ruprecht, 1962), pp. 40-65.

²Frege, "On Sense and Reference," op. cit., p. 56.

³Ibid., p. 56.

⁴Gottlob Frege, Begriffsschrift, eine der Arithmetischen Nachgebildete Formalsprache des Reinen Denkens (Halle, 1879).

⁵Saul Kripke, "Naming and Necessity," in Semantics of Natural Language, ed. Gilbert Harman and Donald Davidson (Dordrecht: D. Reidel, 1972), pp. 253-355, and "Identity and Necessity," in Identity and Individuation, ed. Milton K. Munitz (New York: NYU Press, 1971), pp. 135-164.

⁶Michael Dummett, Frege, Philosophy of Language (London: Duckworth, 1973).

CHAPTER 2

FREGE'S BEGRIFFSSCHRIFT THEORY OF IDENTITY

1. Introduction

Granting that i-sentences differ in cognitive value, the Paradox of Identity can be viewed as a reductio, either of the substitution principle (I), on which cognitive value is held to remain invariant under substitution of codenotational singular terms, or of the assumption that "identity relates objects." In SR, as we have noted, Frege chose the former course; but in his youthful logical work Begriffsschrift¹ (Bg, for short), Frege had opted for the latter. As Frege relates the story in SR, reflection on the paradox had convinced him that the information conveyed by an i-sentence could not be about the objects themselves, for then each true i-sentence would reduce to conveying the trivial information that the designated object is self-identical. If a true i-sentence of the form ($\alpha=\beta$) is to be informative, he concluded, then its informativeness would have to reside in the fact that the different expressions, α and β , turn out to stand for the same thing:

What is intended to be said by $\underline{a=b}$ seems to be that the signs or names 'a' and 'b' designate the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted.²

Thus, Frege says, he had been led in Bg to reject the view that $(\alpha=\beta)$ expresses that the relation being one and the same thing as holds between the objects denoted by α and β , and to maintain instead that $(\alpha=\beta)$ expresses that a relation holds between the expressions α and β themselves, namely, the equivalence relation denoting one and the same thing as.

Now Frege's reconstruction in SR of his Bg reasoning is misleading on two counts. First, he had explicitly denied in Bg that the informativeness of an i-sentence resides solely in the fact that different expressions stand for one and the same object; on the contrary, he had argued that what was significant was not the expressions per se, but the fact that the different expressions go along with different ways of determining [Bestimmungsweisen] the object. Second, Frege's reconstruction is superficial in that he fails to mention what was obviously the critical factor both in his denying that identity relates objects as well as in his maintaining that identity relates expressions, namely, the view he had held about what is said or expressed by a sentence: if reflection on the paradox had indeed led him to reject the view that identity relates objects, then, as our analysis of the paradox showed, this could only have been because he had held a view about what is said or expressed by a sentence that committed him to (I); and, again, if he had held that the informativeness of an i-sentence

resides solely in the fact that the different expressions stand for the same object--and, to repeat, he did not hold this--there remains a gap in his reconstruction, namely, the inference from the assumption that the informativeness of an i-sentence resides in the fact that different expressions stand for the same thing to the conclusion that the i-sentence must therefore express a relation holding between these expressions, and to justify this step, one would like to have some account of what is said or expressed by a sentence. These both have had harmful consequences: the first, of masking the connection between the Bg and SR theories of identity, and the second, of vesting i-sentences with a mystery they do not rightfully possess--or, more accurately, that i-sentences possess if, and only if, all sentences do.

Despite these errors in his reconstruction, however, Frege's description of his Bg treatment of identity is essentially correct. The equivalence relation, which he had called "equality of content" [Inhaltsgleichheit] (ec, for short), and which he had represented by the symbol ' \equiv ', was defined as follows: the judgment ' $\vdash A \equiv B$ ' means

The symbol A and the symbol B have the same conceptual content, so that A can always be replaced by B and conversely.³

The symbol for equality of content was part of the object language, and since it, unlike the symbols for negation and material implication, represented a relation that holds

between expressions instead of their contents, it required a special convention:

Elsewhere, signs are mere proxies [Vertreter] for their content, and thus any phrase they occur in just expresses a relation between their various contents; but names at once appear in propria persona so soon as they are joined together by the symbol for equality of content; for this signifies the circumstance of two names' having the same content.⁴

Hence, names in Bg were systematically ambiguous: they stood for the objects they customarily denoted everywhere save when they occurred at either end of a symbol for equality of content, at which place they stood for themselves. (This shifting between talk of 'objects' and talk of 'contents' will be explained below.)

The Bg theory of identity has been by and large ignored in the philosophical literature. It is the SR theory that has proved of greatest interest, and since the Bg theory is a nonstandard theory of identity, and since, moreover, Frege himself came to repudiate the Bg theory (replacing it with the SR theory), it has been viewed as an early aberration happily forgotten. What criticism there has been of the Bg theory reduces to the following three charges:

(a) It has been alleged that the information contained in an i-sentence, when interpreted in the manner of Bg, can only be the trivial information that the linguistic community has adopted such-and-such conventions, not the sub-

stantial information embodied in a genuine discovery about the world. (This is derived from Frege's own criticism of the Bg theory in SR. Linsky,⁵ Kneale⁶)

(b) It has been alleged that the Bg theory is circular or that it involves a vicious infinite regress. (Russell,⁷ Wiggins,⁸ Kneale⁹)

(c) It has been alleged that the Bg theory is flawed by use/mention confusion. (Church¹⁰)

Of these criticisms, the first is, on my view, not only unpersuasive, but unheedful of what Frege actually says in Bg; the second is based on a misunderstanding of the Bg theory; and the third, though accurate, has never been fully spelled out--it is a passing remark in one of Church's footnotes--and when this criticism has been picked up by others, it has been construed as a purely formal problem. Thus, such an able commentator as Montgomery Furth has remarked:

[The Bg theory of identity] has the merit of accounting for the interest of true "A=B" as against the unformativeness of "A=A". But the price is exorbitantly high, for the device renders it practically impossible to integrate the theory of identity into the formalized object-language itself; e.g., to state generally such a law as that if $F(a)$ and $a=b$ then $F(b)$.¹¹

I agree with Furth's criticism of the formal aspect of the Bg theory, but he is much too generous in allowing that Frege can account for the difference in cognitive value between ($\alpha=\alpha$) and true ($\alpha=\beta$). The Bg theory is an utter failure. It is so thoroughly confused and unappealing that

I suspect Frege himself to have been dissatisfied with it.¹² Many years later, Frege attributed the error of his Bg account of identity to two factors: (a) insufficient attention to the distinction between sign and thing signified, and (b) confusion in the Bg notion of content [Inhalt].¹³ I agree with Frege's diagnosis. I shall develop all of these points in this chapter.

Before turning to the Bg theory, however, I wish to note a problem about interpretation. In SR, Frege indicates that his Bg theory of identity was prompted by reflection on the paradox. But he does not discuss the paradox in Bg, nor does he acknowledge there that he is offering a nonstandard account of identity. (Rather, his primary concern in Bg is to justify the inclusion of a symbol for identity in his system of notation. More on this in the next chapter.) Hence, in regarding the Bg theory of identity as a response to the paradox, I shall have to extrapolate from the text and rechannel his remarks towards a solution to the puzzle.

2. The Begriffsschrift Theory

We have so far given a brief description of the Bg theory of identity. What has yet to be explained is why Frege had adopted it.

We have drawn attention to the following point: i-sentences appear to derive their significance from the

fact that there is quite often more than one expression in a given language that stands for an object. This was an important (though not overwhelming) consideration for Frege. But it is only in conjunction with the substitution principle, (I), that it justifies denial of the view that $(\alpha=\beta)$ is about $d(\alpha)$ and $d(\beta)$. That is, there is no inconsistency in holding that an i-sentence, $(\alpha=\beta)$, is about $d(\alpha)$ and $d(\beta)$ even though its informative character stems from there being different expressions on either side of the identity sign--if one does not accept (I). This is the position preferred by some philosophers today. Quine, for example, remarks in Methods of Logic that "['Cicero = Tully'] is informative, because it joins two different terms."¹⁴ And he later adds:

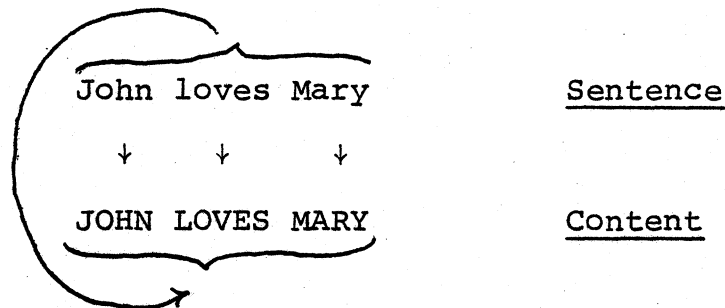
. . . since the useful statements of identity are those in which the named objects are the same and the names are different, it is only because of a peculiarity of language that the notion of identity is needed. If our language were so perfect a copy of its subject matter that each thing had but one name, then statements of identity would indeed be useless.¹⁵

Quine, however, does not accept (I), and with no evident discomfort, he is a vigorous defender of the view that $(\alpha=\beta)$ is about $d(\alpha)$ and $d(\beta)$:

But to say that the need of identity derives from a peculiarity of language is not to say that identity is a relation of expressions in language. On the contrary, . . . what are identical are the objects with themselves and not the names with one another; the names stand in the statement of identity, but it is the named objects that are identified.¹⁶

So, the point we must explain is Frege's acceptance of (I), or something very much like (I), in Bg.

We turn, then, to the semantic theory Frege held in Bg. On that theory, a sentence corresponds to its content [Inhalt]. This might be understood as a thought, but only if we are careful not to assimilate this use of "thought" to Frege's more familiar technical notion of Gedanke from the SR theory. The sentence stands for or denotes its content--the sentence is, in Frege's words, a Vertreter (agent, proxy, substitute, representative) for its content--and the parts of the sentence, in turn, stand for or denote corresponding parts of the content of the sentence. The structure of the sentence, then, mirrors the structure of the thought. This situation might be pictured so,



where the arrows indicate the denoting relation. If an expression in a sentence functions solely as proxy for its content, then we can see rather easily from the diagram that were we to replace any one of the parts of the sentence by any other which has the same content, then the

resultant sentence would have the same content as the original.

Of course, some mirrors are more accurate reflectors than others. Frege developed his Begriffsschrift to provide a more perspicuous representation of thought than is afforded by ordinary language. It was to serve as a tool for the abstract sciences, and especially mathematics, which required, he felt, a system of notation commensurate with the rigorous standards of proof and precision of thought essential to them. Natural language, he said, "proves to be deficient in the matter of protecting our thinking from error."¹⁷ Its primary fault is ambiguity. We find many expressions in natural language which have more than one meaning. 'Bear', for instance, is used in one sense when we speak of a woman's bearing children and in quite another when we speak of a person's bearing up under adversity. Ambiguity becomes especially pernicious when logical differences are covered up. Consider, for example, the two sentences

(6) The horse is a roan,

and

(7) The horse is a mammal,

where (6) speaks of an individual horse while (7) does not. Although (6) and (7) appear to have the same grammatical structure, they represent thoughts which have radically

different logical structures. Using Frege's later terminology, (6) expresses the circumstance of an object's falling under a concept while (7) expresses the subordination of one concept to another. This crucial logical difference between the contents of (6) and (7) is masked by the superficial similarity between the grammatical structure of (6) and (7). In Bg, of course, these contents receive rather different representations.

Just as ordinary language sometimes fails to reflect logical differences, it sometimes marks distinctions which are of no logical importance. For instance, in a language like German which has case inflexions there is great flexibility in the ordering of words in a sentence. But this choice of word order is largely a matter of force or topic, and so only of psychological or pragmatic importance. This is precisely Frege's reason for not reproducing the subject/predicate distinction in his Begriffsschrift:

In language the place occupied by the subject in the word-order has the significance of a specially important place; it is where we put what we want the hearer to attend to specially. . . . This may, e.g., have the purpose of indicating a relation between this judgment and others, and thus making it easier for the hearer to grasp the whole sequence of thought. All such aspects of language are merely results of the reciprocal action of speaker and hearer; e.g. the speaker takes account of what the hearer expects, and tries to set him upon the right track before actually uttering the judgment. In my formalized language there is nothing that corresponds; only that part of judgments which affects the possible inferences is taken into consideration. Whatever is needed for valid inference is fully expressed; what is not needed is for the most part not indicated

either; no scope is left for conjecture.¹⁸

Protecting thought from error is not the sole interest of natural language. Natural language is a versatile and flexible instrument, capable of serving a community of speakers in a wide variety of circumstances. The fact that it must serve so many interests, however, means that it must compromise the needs of each particular one; and so, when we consider the specialized task of representing inference, we ought not be surprised to find that natural language is less than ideal:

Language is not in such a way dominated by logical laws that compliance with grammar would of itself guarantee the correctness of thought processes. The forms in which inference is expressed are so diverse, so unstable and elastic that it is easy for premises to slip in unnoticed which are then overlooked in listing the necessary conditions for the validity of the conclusion, thereby lending it a greater validity than it should by rights have.¹⁹

A strictly delineated group of forms of inference is just not present in language with the result that it is not possible on the basis of linguistic form to distinguish a flawless argumentation from one in which steps have been left out. It can, in fact, be said that the former is almost not to be found in language, that it runs counter to the feel of language, because it would be bound up with an insufferable prolixity. Logical relationships are almost always only hinted at by language and left to guesswork but not properly expressed.²⁰

Protecting thought from error, however, was to be the sole interest of Frege's Begriffsschrift, and, therefore, it was to have the vital properties mentioned in the passages just cited which natural languages lack. In particular, it

was to be governed by logical laws in such a way that mere adherence to grammar would guarantee the formal correctness of thought processes.

Only that portion of the content of a sentence which counted for inference was of any interest to Frege, and this he called the "conceptual content" [Begrifflichen Inhalt] of the sentence:

. . . let me observe that there are two ways in which the content of two judgments may differ; it may, or it may not, be the case that all inferences that can be drawn from the first judgment when combined with certain other ones can always be drawn from the second when combined with the same other judgments. The two propositions 'the Greeks defeated the Persians at Plataea' and 'the Persians were defeated by the Greeks at Plataea' differ in the former way; even if a slight difference of sense is discernible, the agreement in sense is preponderant. Now I call the part of the content that is the same in both the conceptual content. Only this has significance for our symbolic language; we need therefore make no distinction between propositions that have the same conceptual content.²¹

Since each sentence is inferable from itself, with or without any additional premises, Frege's condition for sameness of conceptual content comes to this: two sentences have the same conceptual content if, and only if, they are mutually inferable. Hence, a sentence in Bg was viewed as a transparent substitute for its conceptual content, and any other sentence having the same conceptual content could serve in its stead.

Both sentences and parts of sentences were said to have contents. The content of a sentence was called an

"assertible content" [Beurtheilbar Inhalt]; the content of any other kind of expression--and so, of a singular term--was called a "nonassertible content." The quoted condition for sameness of conceptual content--mutual inferability--only applies to assertible contents, and nowhere in Bg had Frege explicitly stipulated the conditions under which two singular terms were to have the same conceptual content. In keeping with the general semantic structure of Bg, however, the indicated course would be to take two singular terms as having the same conceptual content if, and only if, replacement of one by the other in any given sentence results in another sentence having the same conceptual content as the original. We ought to note that Frege's practice in Bg was to take two singular terms as having the same conceptual content if, and only if, they denote the same object; and thus he identified the conceptual content of a singular term with the object for which it stands. Since the content of a part of a sentence is said to be part of the content of the whole sentence, we find that objects themselves are parts of thoughts. This hybrid notion of 'content' reveals a confusion in Bg which Frege later attempted to rectify by distinguishing between the sense and the reference of an expression.

Frege's brief remarks on sameness of conceptual content require some further comment. The two sentences,

(8) The Greeks defeated the Persians at Plataea,
and

(9) The Persians were defeated by the Greeks at Plataea,
differ only as active and passive; and since, according to Frege, these two sentences have the same conceptual content-- which is to say that whatever inferential relations the one enters, the other enters too--no corresponding distinction is made in his Begriffsschrift. Since there is no active/passive distinction to be found in Bg, and since (8) and (9) have the same conceptual content, then the point of not differentiating between the two sentences is that any symbolization of (8) serves equally as a symbolization of (9). For example, if we were to symbolize (8) as a two-place relational sentence, 'F(a,b)', according to the translation scheme:

'a': 'The Greeks'

'b': 'The Persians'

'F(ξ , η): ' ξ defeated η at Plataea'

then, since this is also a symbolization of (9), we should more accurately give 'F(ξ , η)' as

'F(ξ , η): ' ξ defeated η at Plataea' or ' η were defeated by ξ at Plataea'

Hence, the fact that (8) and (9) have the same conceptual content is not proved in his Begriffsschrift, but rather it is assumed at the outset. Frege is here obviously relying on a person's intuitive knowledge to recognize

that (8) and (9) have the same conceptual content. Were a person unaware that sentences related to one another as (8) and (9) have the same conceptual content, Frege's Begriffsschrift would not help him to find out that they are. For, if this person were unaware that (8) and (9) did have the same conceptual content, he would not symbolize each in the same way, but would, perhaps, take ' ξ defeated η at Plataea' as ' $F(\xi, \eta)$ ' and ' η were defeated by ξ at Plataea' as ' $G(\eta, \xi)$ '. In this case, he would not be able to prove, using only the rules of Frege's Begriffsschrift, that ' $F(a, b)$ ' and ' $G(b, a)$ ' have the same conceptual content, because

$$F(a, b) \equiv G(b, a)$$

is not formally valid.

On the other hand, we are able to construct different strings of symbols in Frege's Begriffsschrift which have the same conceptual content and which, moreover, can be shown to have the same conceptual content using only the rules of his Begriffsschrift. For example (using modern notation),

$$(10) \quad ((p \supset q) \supset p) \supset p,$$

and

$$(11) \quad p \supset (q \supset p),$$

are both theorems, and so, since they are mutually derivable, they have the same conceptual content. On the translation scheme:

'p': 'war is hell'

'q': 'hope springs eternal'

(10) and (11) would be symbolizations, respectively, of

(12) If if if war is hell then hope springs eternal then
war is hell, then war is hell,

and,

(13) If war is hell, then if hope springs eternal then
war is hell

Hence, (12) and (13) have the same conceptual content, and unlike (8) and (9), the fact that they have the same conceptual content can be proved in Frege's Begriffsschrift. Of course, we might imagine there to be a person who recognizes intuitively that (12) and (13) have the same conceptual content, and so symbolizes (12) and (13) in the same way. But for most of us, the fact that (12) and (13) do have the same conceptual content is not intuitively obvious, and so Frege's Begriffsschrift is of great help.

Obviously, then, some distinctions between propositions that have the same conceptual content are made: some sentences of the form $(\phi \equiv \psi)$ are theorems in his Begriffsschrift and others are not. The Begriffsschrift was intended to aid in evaluating inferences, and so to aid in determining whether two sentences are mutually inferable. Hence, there appears to be no theoretical reason for not enriching Frege's artificial notation to include

an active/passive distinction, adding on an axiom to the effect that any two strings of symbols which differ only in that one is active and the other passive (however this is to be marked) have the same conceptual content. In this enriched language, (8) and (9) could be shown to have the same conceptual content on formal grounds alone. As far as I can tell, the reason for not doing so is a pragmatic one: it is just intuitively obvious that sentences like (8) and (9) have the same conceptual content, and there is no point in cluttering up the symbolic notation to mark a distinction which can be more easily seen without its aid.

We return, now, to the main point of our excursion into the Bg semantic theory, which is to seek an analogue to the substitution principle, (I). If we assume that the notion of "cognitive value" in SR coincides with the Bg notion of "conceptual content," and if we take advantage of the fact that every sentence in Bg is said to stand for or denote its conceptual content, then we have our analogue in

(II) If $S\alpha$ is about $d(\alpha)$, then

if $d(\alpha) = d(\beta)$, then $d(S\alpha) = d(S\alpha/\beta)$.

How does this principle fit into the Bg theory? Frege's Begriffsschrift is intended to be an artificial notation which so accurately reflects the structure of thoughts

that the symbols themselves can be treated as if they were the contents represented, and so inference, which is actually a matter pertaining to thought, is reduced to a manipulation of symbols. Frege's Begriffsschrift, however, is only the grammatical framework of a language. Had it been a full-fledged language, and were we to imagine 'John loves Mary' to be a sentence of the language, then as we saw from the diagram a few pages back, (II) is immediate. In fact, we fill the framework provided by his Begriffsschrift by associating the symbols with sentences from a natural language. But natural languages are imperfect representations of thought, and so in assigning a particular analysis to a given sentence from some natural language, we must be sure that the sentence does take the proposed analysis. And here is where (II) comes in, namely, to help test whether a given analysis accurately represents the conceptual content of the sentence. If we have a sentence containing a singular term, $S\alpha$, and we wish to determine whether $S\alpha$ is about $d(\alpha)$, then by trying out different singular terms in place of α we slowly eliminate the likelihood that the conceptual content of $S\alpha$ depends upon any other feature of these terms than the object they stand for if the conceptual content of the sentence remains unchanged.

If I have reconstructed the Bg semantic theory correctly, then it is just such an application of (II)

which led Frege to reject the view that an i-sentence of the form ' $\alpha=\beta$ ' is about $d(\alpha)$ and $d(\beta)$. He believed that ' $\alpha=\alpha$ ' and true ' $\alpha=\beta$ ' differed in conceptual content, and so, since $d(\alpha)=d(\beta)$, but $d(S\alpha)\neq d(S\alpha/\beta)$, ' $\alpha=\beta$ ' could not express a relation which holds between $d(\alpha)$ and $d(\beta)$.

3. Criticism of the Begriffsschrift Solution

How does Frege's Bg theory solve the problem of the difference in cognitive value between ' $\alpha=\alpha$ ' and true ' $\alpha=\beta$ '? According to the theory, the two i-sentences, (1) and (2), are understood as

$$(14) \quad a\equiv b,$$

and

$$(15) \quad a\equiv a,$$

respectively, where the singular terms on either end of the ec-symbol, by the convention laid down, stand for themselves. Let us assume that (14) is about that which the terms on either end of the ec-symbol stand for--and similarly for (15). And let us assume further, that (14) and (15) are both true and yet differ in conceptual content, i.e., that $d('a\equiv b')\neq d('a\equiv a')$. Since (15) results from (14) by replacement of 'b' by 'a', then by (II), 'a' and 'b' also differ in conceptual content, i.e., $d('a')\neq d('b')$. But this is no longer puzzling, because 'a' and 'b' are different expressions.

Note, however, that (14) says that the symbol 'a' has the same conceptual content as the symbol 'b', i.e., that $d('a')=d('b')$; and (14) was assumed to be true. Yet the conclusion of the argument is that $d('a')\neq d('b')$.

On reflection, we ought not be surprised at this result. We cannot properly speak of the denotation of an expression, for an expression might denote different objects in different contexts, namely, its usual content or itself. It is precisely this ambiguity which was traded on to produce the contradiction. The easiest way of getting around this difficulty would be to drop Frege's awkward device of having an expression stand now for its content, now for itself, and adopt, instead, the convention of forming the name of an expression by enclosing that expressing within single quotes. The whole, following Quine, is called a quotation. Each expression in the language now always stands for its content--never itself.

We would represent (14) and (15), respectively, as

(14') 'a'≡'b',

and

(15') 'a'≡'a'.

A quotation is, from the logical point of view, a simple name: the occurrence of 'a' in ''a'' is, in Quine's phrase, a mere orthographic accident, because $d(''a'')$ is not uniquely determined by $d('a')$; and the unfortunate tendency to overlook this fact, that is, to unwittingly assume that

the structure of the notation reflects logical structure, would be avoided had we adopted a different, but equally suitable, convention. At any rate, a consequence of a quotation's being a simple name is that substitution within single quotes is illegitimate. We cannot construe (15') as resulting from (14') by replacement of 'b' by 'a', but only as resulting by replacement of ''b'' by ''a''. Continuing the argument, since (14') differs in conceptual content from (15'), i.e., $d('a' \equiv 'b') \neq d('a' \equiv 'a')$ and since (15') results from (14') by replacement of ''b'' by ''a'', then by (II), ''a'' and ''b'' also differ in conceptual content, i.e., $d('a') \neq d('b')$. No contradiction arises, for (14'), assumed true, says that 'a' and 'b' have the same conceptual content, i.e., $d('a') = d('b')$; and 'a' is not the same sign as ''a''.

It would appear, then, that with this slight notational emendation, the Bg theory solves the problem about the difference in cognitive value between $(\alpha = \alpha)$ and true $(\alpha = \beta)$. By taking an i-sentence to be about the terms occurring in the sentence, (II) is satisfied and the puzzle dispelled. But this is not so.

To begin, note that the assumption of the truth of (14') was never used in the argument. In fact, the truth value of (14') was irrelevant to the argument; for, the difference in conceptual content between (14') and (15') was found to consist solely in the difference between the

singular terms occurring in (14') and (15'). However, a difference in conceptual content between (14') and (15') might arise from the fact that (14') and (15') differ in truth value, or, it might arise from the fact that (14') and (15'), while agreeing in truth value, contain different singular terms. But the Bg theory does not distinguish between these two sorts of cases. Hence, it is misleading--if not false--to say that the Bg theory accounts for the fact that there is a difference in cognitive value between $(\alpha=\alpha)$ and true $(\alpha=\beta)$.

There is, however, a more serious problem with the theory. By using the single quote method of name formation, we have been able to reestablish a uniformity in the notation: an expression always goes proxy for its content, never itself. We find, however, that ec, although a relation which holds between expressions, holds between these expressions qua contents. That is, it is not the expressions on either end of the ec-symbol which are now said to have the same content, but the content of those expressions, i.e., what those expressions stand for. Hence, it would be incorrect to describe ec as being a relation which holds between expressions and not contents; for it does hold between contents, albeit contents of a special sort--contents which themselves have content, namely, symbols.

Now, for a given expression, as, in general, for any given object, there is quite often more than one expression which uniquely identifies it. For example, 'the word 'four'' and 'the first word of the Gettysburg Address' both denote the same word. And it is therefore possible for two i-sentences, $(\alpha=\alpha)$ and true $(\alpha=\beta)$, to differ in cognitive value when the object denoted by 'a' and 'b' happens to be a word. For example, compare 'the word 'four' = the first word of the Gettysburg Address' with 'the word 'four' = the word 'four''. The latter is a truism, but the former is not. Using this fact, it is relatively easy to show that the same difficulty which Frege found with the view that $(\alpha=\beta)$ is about $d(\alpha)$ and $d(\beta)$ also burdens his own.

'Mark Twain' is (identical with) the pen name of Samuel Clemens. So, the conceptual content of 'Mark Twain' is the same as the conceptual content of 'the pen name of Samuel Clemens'. Consider, now, the two ec-sentences:

(16) 'Mark Twain' \equiv the pen name of Samuel Clemens,

and

(17) 'Mark Twain' \equiv 'Mark Twain'.

(17) is a truism, but (16) is more subtle: since 'Mark Twain' is one and the same word as the pen name of Samuel Clemens, and since a word denotes whatever it denotes, then 'Mark Twain' must stand for the same thing as the pen

name of Samuel Clemens. (17) results from (16) by replacement of 'the pen name of Samuel Clemens' by 'Mark Twain'. Since, as has been previously noted, 'the pen name of Samuel Clemens' has the same conceptual content as 'Mark Twain', i.e., since $d(\text{'the pen name of Samuel Clemens'}) = d(\text{'Mark Twain'})$, then by (II), (16) and (17) must also have the same conceptual content, i.e., $d((16)) = d((17))$. But, it is not at all obvious that (16) and (17) do have the same conceptual content--(16) is much more informative than (17). There seems to be just as much reason for supposing that (16) and (17) differ in cognitive value as there was for supposing that 'Mark Twain is identical with Mark Twain' and 'Mark Twain is identical with Samuel Clemens' differ in cognitive value.

As long as there is more than one expression which can uniquely identify a given object, we are faced with the problem of difference in cognitive value. Such was the case for identity, and such, we have just found, is the case for ec. This fact had been concealed from us in the emended argument in the previous section because of the way in which that argument had been set up. In using the single quote method of name formation, we had surreptitiously assumed that the names on either end of the ec-symbol named different expressions. For a quotation actually contains its denotation as a proper part; so, two quotations differ if, and only if, their denotations (which are, in fact, their quoted interiors) differ.

In the emended argument, the expressions on either end of the ec-symbol in (14') differed because the expression on either end of the identity sign in (1) differed. This was not due to our use of single quotes to form names, but rather to the fact that quotations are names: the same would hold whatever names--whether proper names, quotations, or definite descriptions--were used to refer to the singular terms in (1). Distinct objects require, on pain of ambiguity, distinct names; and we had, essentially, adopted the simplifying assumption that no name names more than one object. Of course, this assumption does not justify our having the same expression on either end of the ec-symbol in (15') because the same expression occurred on either end of the identity sign in (2). For, in general, it need not hold that any two distinct names name distinct objects.

Were we, then, to let 'a' and 'b' be arbitrary names of expressions and consider,

$$(18) \quad \underline{a} \equiv \underline{b},$$

and

$$(19) \quad \underline{a} \equiv \underline{a},$$

we would see clearly that ec runs afoul of (II). For, let us suppose that (18) and (19) are both true and yet differ in conceptual content, i.e., that $d(\underline{a} \equiv \underline{b}) \neq d(\underline{a} \equiv \underline{a})$.

Since (19) results from (18) by replacement of 'b' by 'a', then by (II), 'a' and 'b' differ in conceptual content, i.e.,

$d('a') \neq d('b')$. But it is, as we have just seen, possible that a and b be one and the same expression, i.e., that $a=b$, in which case 'a' and 'b' could not differ in conceptual content, i.e., $d('a')=d('b')$.

The fact that the problem of difference in cognitive value can be reproduced for ec is certainly damaging to the Bg theory. How damaging becomes evident from the following considerations.

On the Bg theory, an i-sentence is understood to be (really) saying something about the singular terms occurring on either end of the identity sign, namely, that they have the same conceptual content. There are not, on this account, two relations--identity and ec--the one holding between objects and the other holding between names of objects. There is only one relation, for the view that an i-sentence ($\alpha=\beta$) is about $d(\alpha)$ and $d(\beta)$ has been rejected since it conflicts with the fact that i-sentences differ in cognitive value. The conceptual content of an i-sentence, then is that the singular terms occurring in the sentence have the same conceptual content, so when a given i-sentence is translated into Bg notation, the singular terms flanking the ec-symbol are understood to stand for the singular terms flanking the identity sign. But, since to a given singular term there might correspond more than one name which uniquely identifies it, there is no unique

correlation of i-sentences with ec-sentences. This would raise no serious difficulties if each of the ec-sentences correlated with a given i-sentence had the same conceptual content. But, as we have just shown, this is not so. Since the various ec-sentences correlated with a given i-sentence might not have the same conceptual content, we shall not be able properly to speak of the conceptual content of an i-sentence. That is, the conceptual content of an i-sentence would not be determined so much by the singular terms occurring on either end of the identity sign as it would be by the names used to denote these singular terms. Hence, the conceptual content of an i-sentence which Frege sought to capture with his Bg theory, namely, that two singular terms stand for or denote one and the same thing, is hopelessly masked.

4. A Second Interpretation of Equality of Content

One response to the problem would be to step to the next level and seek a relation which holds between names of names of expressions. Aside from the evident implausibility of this scheme--Are we to suppose that an i-sentence says something about (some? all?) names of the singular terms occurring on either end of the identity sign?--it is simply a dead end, for there can be little doubt that we shall once again run square against the problem of difference in cognitive value for this new relation.

Another--and more promising--response would be to attempt to save the Bg theory by placing a restriction on the language such that each expression in it can be uniquely identified by no more than one name. In this way, the argument in the previous section would be blocked: for, now, two names of expressions, 'a' and 'b', would be such that $\underline{a}=\underline{b}$ if, and only if, 'a' \equiv 'b' (i.e., $d(\underline{a})=d(\underline{b})$). And this, I think, is the kind of thing Frege actually does in Bg.

In the last section, we had regarded the terms on either end of the ec-symbol as names of the corresponding expressions occurring on either end of the identity sign in the correlated i-sentence. However, in the text, Frege understands the terms occurring on either end of the identity sign to stand for themselves, and so chose likewise for the terms on either end of the ec-symbol. By doing this, the names are, so to speak, given directly and there is thus no problem about determining whether we have the same or different sign as we had last section when the names flanking the identity sign were given indirectly. On this interpretation, the singular terms occurring on either end of the identity sign are not used to stand for their customary denotations; in this context, the singular terms are used to stand for themselves (hence, they are both used and mentioned), and the claim of identity is so

understood that the mentioned singular terms are said to have the same conceptual content (i.e., to stand for the same thing). In

(1) Mark Twain = Samuel Clemens,

the expressions 'Mark Twain' and 'Samuel Clemens' stand for themselves, and '=' is understood to express a relation between these names. In Bg notation, this i-sentence, (1), is represented as

(20) Mark Twain \equiv Samuel Clemens,

where, again, 'Mark Twain' and 'Samuel Clemens' stand for themselves.

In order to avoid confusions which might arise from both using and mentioning a given expression in a given context--i.e., of using an expression autonomously--our own procedure was to insert single quotes around the expressions occurring on either end of the ec-symbol. So, (20) became

(21) 'Mark Twain' \equiv 'Samuel Clemens'.

However, we had overemphasized the naming aspect of quotation and forgotten the particular virtue of single quotes which makes it preferable to other methods of name formation, namely, that a quotation exhibits or displays its denotation. In a sense, a quotation is a perfectly transparent name: for anyone familiar with the single quote method of name formation, acquaintance with a quotation is tantamount to acquaintance with the object for which the

quotation stands. (Indeed, this is the paradox of quotation: whereas the value of using single quotes lies in our being able to look inside the quote marks and see the expression denoted, the logical condition of a quotation's being a name is that what appears inside the quote marks counts for naught.)

The importance of so exhibiting or displaying an expression becomes apparent when we consider, again,

(16) 'Mark Twain' \equiv the pen name of Samuel Clemens.

On the view just outlined (16) turns out to be nonsense: it is not a well-formed (i.e., syntactically coherent) sentence. The expression 'the pen name of Samuel Clemens' certainly names the expression 'Mark Twain', but it fails to exhibit or display 'Mark Twain' or, indeed, any other expression. Hence, not (16), but

(22) 'Mark Twain' \equiv 'Mark Twain'

corresponds to the i-sentence

(23) Mark Twain = Mark Twain.

And, again, not (16), but

(24) 'Mark Twain' \equiv 'the pen name of Samuel Clemens'

corresponds to the i-sentence

(25) Mark Twain = the pen name of Samuel Clemens.

(This latter example is a bit more obvious, because (16) is true, but (24) and (25) are both false.) (16) is not well-formed because, in an ec-sentence, the expressions said to have the same conceptual content must actually appear in

the sentence. Quote marks serve not so much to form names as they do to underscore the fact that the quoted expressions, in that context, stand not for their ordinary contents, but for themselves. In effect, then the single quotes are mere explicit reminders that the quoted expressions are occurring autonomously.

There is, thus, a 1-1 mapping from the set of i-sentences onto the set of ec-sentences which is described by the following algorithm: given an i-sentence, replace the identity sign by the ec-symbol and place single quote marks around the expressions flanking the ec-symbol. To reverse the procedure, simply delete the outermost pairs of single quote marks occurring around the expressions flanking the ec-symbol and replace the ec-symbol by the identity sign.

5. Criticism of the Second Interpretation

I think this interpretation is much closer to Frege's intentions in Bg than is the account provided in section 3. But there remains, still, a question about the logical status of our use of single quotes in ec-sentences: Are these quotations to be regarded as names or not? Frege, I think, would wish to say that they are names; for, his own device was to have the terms on either end of the ec-symbol occur autonomously. But it is doubtful whether quotations can be logically regarded as names on this

interpretation.

First, quotations do not belong to the same syntactic category of expressions as do proper names and definite descriptions. For, whereas $'a' \equiv 'b''$ is well-formed when 'a' and 'b' are singular terms in the language, were we to replace either $'a''$ or $'b''$ by a proper name or a definite description--not a quotation--the resultant string would no longer be well-formed. The case is entirely analogous to replacing the wedge in $'(p \vee q)'$ by either a left parenthesis or a propositional variable, to obtain, respectively, $'(p(q)'$ and $'(p r q)'$. Parentheses, logical connectives, and propositional variables belong to three distinct syntactic categories.

Second, if we are to regard our single quote convention as a species of name formation, we shall have to recognize that the requirements of our notation severely limit the alternative conventions open to us: we can only adopt a convention wherein the mentioned expression is displayed or exhibited. So, for example, instead of single quotes, we might have adopted a procedure Geach once proposed, namely, underlining the mentioned expressions, i.e., writing 'Mark Twain' instead of $'Mark Twain''$.²² However, we would be unable to use the most widely known alternative to quotations, namely, Tarski's structural-descriptive names.²³ On this convention, an expression is

described by spelling it: indicating the letters making up the expression (using names of these letters, of course) and specifying the ordering of these letters. So, for example, 'Mark Twain' would be denoted by the structural description:

the expression consisting of two words, the first of which is composed of the letters em, ay, ar, and kay (in that order), and the second of which is composed of the letters tee, double-u, ay, eye, and en (in that order).

Or, adopting the symbol '^' to represent concatenation, and using '#' to represent a word boundary, we could denote 'Mark Twain' by the somewhat shorter expression:

em^ay^ar^kay^#^tee^double-u^ay^eye^en

Yet,

'Mark Twain'≡em^ay^ar^kay^#^tee^double-u^ay^eye^en would be ill-formed, and for the very same reasons that (16) was ill-formed: the actual expressions said to have the same conceptual content must be exhibited in the ec-sentence (just as they are exhibited in the i-sentence).

We mention this point in particular because of the importance Tarski and Quine attach to the fact that structural-descriptive names can be used in place of quotations. For, if quotations are to be regarded as names, it is claimed, they must be regarded as simple names, i.e., as having no significant logical structure. For example, Tarski says:

Quotation-mark names may be treated like single words of a language, and thus like syntactically simple expressions. The single constituents of these names--the quotation marks and the expressions standing between them--fulfill the same function as the letters and complexes of successive letters in single words. Hence they can possess no independent meaning. Every quotation-mark name is then a constant individual name of a definite expression (the expression enclosed by the quotation marks) and in fact a name of the same nature as the proper name of a man.²⁴

And Quine:

. . . from the standpoint of logical analysis each whole quotation must be regarded as a single word or sign, whose parts count for no more than serifs or syllables. A quotation is not a description, but a hieroglyph; it designates its object not by describing it in terms of other objects, but by picturing it. The meaning of the whole does not depend upon the meaning of the constituent words. The personal name buried within the first word of the statement
'Cicero' has six letters,
e.g., is logically no more germane to the statement than is the verb 'let' which is buried within the last word.²⁵

And, the fact that the word inside the single quotes appears to be of logical significance is claimed to be an incidental feature of the notation which would be avoided had we adopted structural-descriptive names instead.

Quine argues,

The quotational context ''9>5'' of the statement '9>5' has, perhaps, unlike the context 'cattle' of 'cat', a deceptively systematic air which tempts us to think of its parts as somehow logically germane. Insofar as this temptation exists, it is salutary to paraphrase quotations by the following expedient. We may adopt names for each of our letters and other characters, and Tarski's '⌢' to express concatenation. Then, instead of naming a notational form by

putting that notational form itself bodily between quotation marks, we can name it by spelling it.²⁶

The shift from quotation to spelling has an independent advantage . . . , but incidentally it is instructive as stressing that any non-referential occurrences caused by quotation are surface appearances, dispelled by an easy change in notation.²⁷

But, turning the argument around, if we cannot replace quotations by structural-descriptive names, then it is not true that the incorporation of an expression within single quotes is merely an incidental feature of our use of quotations; and so it is now doubtful that the occurrence of 'Mark Twain' in ''Mark Twain'' is of no logical significance; and if it is a condition of our regarding ''Mark Twain'' as a name that the embedded constituent 'Mark Twain' be of no logical significance, then it is doubtful whether ''Mark Twain'' can be regarded as a name.

This brings us, now, to the heart of the matter: the proper logical analysis of ec-sentences.

Frege, no doubt, construed an ec-sentence to be a relational sentence, i.e., as admitting the analysis $(R(\alpha, \beta))$; for ec is said to be a relation which holds between expressions. The most natural analysis, then, of an ec-sentence, ''a'≡'b'', would be

(26) R: _____ ≡ _____

α: 'a'

β: 'b'

However, the remarks at the end of the previous section

would seem to indicate that the quote marks go along with the ec-symbol, so that the relational analysis of $'a' \equiv 'b'$ would be:

(27) R: ' _____ ' \equiv ' _____ '

α : a

β : b

Analyzed in the manner of (26), $'a' \equiv 'b'$ results by completing the function-expression $'(.) \equiv ()'$ by the argument-expressions $'a'$ and $'b'$. Analyzed in the manner of (27), on the other hand, $'a' \equiv 'b'$ results by completing the function-expression $'() \equiv ()'$ by the argument-expressions $'a'$ and $'b'$. Which of these analyses are we to choose?

Frege's discussion of function and argument in Bg is a bit thin, and it is, moreover, flawed by (a) failure to distinguish carefully between functions and function-expressions (and, similarly, arguments and argument-expressions), and by (b) failure to separate functions and arguments (i.e., to distinguish between concept and object). One fact, however, which clearly emerges from the text is that there might be alternative analyses for a given sentence which are equally correct, i.e., that there need not be a unique logical analysis for a given sentence. (Frege seems anxious to make this point in order to underscore the greater flexibility of his own function-argument

analysis than is allowed by the rival subject-predicate analysis.) For example, he says,

Let us suppose that there is expressed in our formalized language the circumstance of hydrogen's being lighter than carbon dioxide. In place of the symbol for hydrogen we may insert the symbol for oxygen or nitrogen. This changes the sense in such a way that 'oxygen' or 'nitrogen' enters into the relations that 'hydrogen' stood in before. If an expression is thought of as variable in this way, it is split up into a constant part representing the totality of these relations and a symbol, imagined as replaceable by others, that stands for the object related by the relations. I call the one part a function, the other an argument. This distinction has nothing to do with the conceptual content; it concerns only our way of looking at it. In the manner of treatment just indicated, 'hydrogen' was the argument and 'being lighter than carbon dioxide' the function; but we can equally look at the same conceptual content in such a way that 'carbon dioxide' is the argument and 'being heavier than hydrogen' is the function. We need in this case merely to imagine 'carbon dioxide' as replaceable by other ideas like 'hydrochloric acid gas' or 'ammonia'. 28

I take it that Frege is arguing that the sentence

(28) hydrogen is lighter than carbon dioxide

can be given the analysis $(F(\alpha))$ in either of the following two ways:

(29) F: ____ is lighter than carbon dioxide

α : hydrogen

or

(30) F: hydrogen is lighter than ____

α : carbon dioxide

Moreover, he claims, either analysis is correct. Analyzed in the manner of (29), (28) results from the completion of the function-expression '() is lighter than carbon dioxide'

by the argument-expression 'hydrogen'. Analyzed in the manner of (30), (28) results from the completion of the function-expression 'hydrogen is lighter than ()' by the argument-expression 'carbon dioxide'.

Hence, it might turn out that the choice between analyzing "'a'≡'b'" in the manner of (26) or in the manner of (27) is entirely superficial, for they might both turn out to be equally good analyses. But what constraints are there on something's counting as a correct analysis? Frege enunciates two constraints which are of particular importance for us.

First, the mode of analysis chosen must have no effect on the conceptual content of the given sentence. Repeating what he says in the above-quoted passage, "This distinction [between function and argument] has nothing to do with the conceptual content; it concerns only our way of looking at it." Again, he says,

We attach no importance to the various ways that the same conceptual content may be regarded as a function of this or that argument, so long as function and argument are completely determinate.²⁹

Second, in order to regard an expression in a given sentence as an argument-expression, that expression must occupy a position open to substitution:

Suppose that a simple or complex symbol occurs in one or more places in an expression (whose content need not be a possible content of judgment). If we imagine this symbol as replaceable by another (the same one each time) at one or more of its occurrences, then

the part of the expression that shows itself invariant under such replacement is called the function; and the replaceable part, the argument of the function.³⁰

Taking these two considerations into account, we now wish to determine which, if either, of the two analyses, (26) and (27), is a correct logical analysis of the ec-sentence 'a'≡'b'.

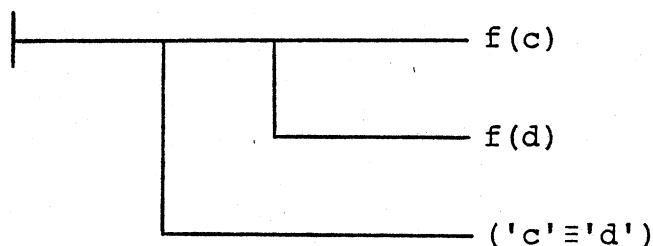
Let us consider (26) first. 'a'≡'b' cannot be correctly analyzed in the manner of (26) because (26) fails to satisfy the substitution condition. The single quotes go along with the ec-symbol. As we have mentioned a number of times, any expression occurring on either end of an ec-symbol must be embedded in single quotes; hence the single quote marks are an invariant part of an ec-sentence. As Frege explained the ec-symbol, whereas signs elsewhere stand for their contents, kehren sie plötzlich ihr eignes Selbst hervor, sobald sie durch das Zeichen der Inhalts-gleichheit verbunden werden³¹ (Bauer-Mengelberg's translation of this clause is more dramatic in placing the burden of autonomous denotation on the context rather than on the expressions than is Geach's quoted on p. 19:

"they suddenly display their own selves when they are combined by means of the sign for identity of content."³²

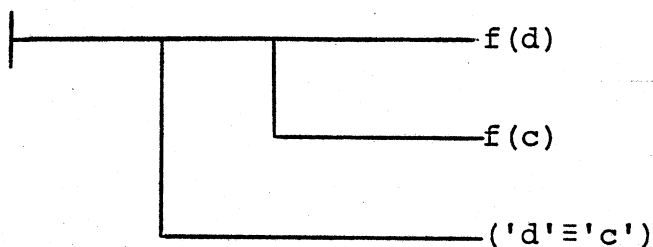
The very understanding of the ec-symbol is that what appears on either side is mentioned; and our single quotes serve merely to mark this fact. By separating the single quotes

from the ec-symbol, as is done in (26), it becomes possible for the blanks to be filled by signs which do not stand for themselves, and so, it becomes possible to substitutionally derive ill-formed nonsense strings like (16).

Indeed, Frege's own practice in Bg is to regard $'a' \equiv 'b'$ in the manner rather of (27). For example, he derives his theorem 57 (with single quotes inserted appropriately)



from



by replacing 'd' by 'c' and 'c' by 'd'.³³ That is, he regards the quoted expression as occupying a position open to substitution. Had he carved up ec-sentences in the manner of (26), he would have had to supplement his substitution table to include replacement of 'c' by 'd' and 'd' by 'c'.

Having disposed of (26), we turn to (27). (27) seems to accord with the substitution constraint. But, if

we are to regard an ec-sentence, $'a \equiv b'$, as built up from 'a', 'b', and $'() \equiv ()'$, we will find ourselves in conflict with the other constraint, namely, that the logical analysis of a given sentence into function-expression and argument-expression(s) have no effect whatsoever on the conceptual content of the sentence. To see this point, turn back to our discussion in section 3. In order to show that (14') and (15') were both true and yet differed in conceptual content--and did not violate (II)--we had to assume that our use of single quote marks entailed that (15') resulted from (14') by replacement of 'b' by 'a' and not by replacement of 'b' by 'a'. Such replacement would be justified, however, only if we were to analyze ec-sentences in the manner of (26), where quotations are regarded as replaceable elements of an ec-sentence. Analyzed in the manner of (27), on the other hand, quotations cannot be so regarded; rather, (15') is viewed as resulting from (14') by replacement of 'b' by 'a'. And, in that case, the argument we had constructed fails.

Hence, neither (26) nor (27) provides an adequate model for a relational analysis of ec-sentences. Since these two are the only plausible candidates, it does not seem possible to assign an ec-sentence the logical form $(R(\alpha, \beta))$. Indeed, there seems to be no way at all of carving an ec-sentence up into function-expression and argument-

expression(s); and so, from the logical point of view, an ec-sentence must be regarded as an unanalyzable whole--a simple symbol possessing no significant logical structure.

6. Formal Difficulties with Equality of Content

The importance of having to treat an ec-sentence as an unanalyzable whole ought not be underestimated.

Consider the argument

$$a \equiv b$$

$$\underline{b \equiv c}$$

$$a \equiv c$$

(In this section, we suppress single quote marks where no confusion will result in order to facilitate exposition.) This argument is clearly valid, because ec is transitive: if 'a' has the same conceptual content as 'b' and 'b' has the same conceptual content as 'c', then it must be the case that 'a' has the same conceptual content as 'c'. One should expect that such an obvious truth as the transitivity of ec be expressible in Bg, i.e., that the argument turn out to be formally valid. (To be sure, it can be shown to be formally valid in Bg, but illegitimately, I think; for the axioms governing ' \equiv ' in Bg are axioms for identity, not ec.) But, if an ec-sentence has no significant logical structure, then we cannot justify the validity of the argument by appealing, e.g., to the fact that the item said to be equivalent in content to 'a' in the first

premise is the same as the item said to be equivalent in content to 'c' in the second premise; for this would appeal to the internal structure of the premises. Perhaps ec is transitive, i.e., perhaps ec is a transitive relation; but, as noted, we cannot logically regard ec -sentences as relational sentences, so we have no formal means of representing the transitivity of ec . If an ec -sentence is a simple symbol possessing no significant logical structure, then we shall have to represent it as such. The most perspicuous logical representation of the argument we can devise, then, is the following:

p
q
r

And this argument, as we can plainly see, is not formally valid.

There are two axioms in Bg governing ec :

(31) $\vdash c \equiv c$

and

(32) $\vdash \begin{array}{l} f(d) \\ f(c) \\ c \equiv d \end{array}$

Although (31) and (32) together would, if ' \equiv ' were assigned the usual interpretation of identity, provide a complete axiomatization of identity within a first order theory,

they cannot consistently carry the interpretation Frege desired in Bg. For example,

$$\text{————— } b \equiv b$$

would seem to be sanctioned by substitution in (31). This cannot be, however, for if our analysis of ec-sentences is correct, 'c≡c' has no replaceable, i.e., substitutable, parts. The difficulty with ec is most dramatic when we consider the interpretation of variables in ec-sentences. Quine has, on many occasions, drawn our attention to the intimate connection between reference and quantification;³⁴ and his observations in other contexts are easily transferred to our own.

By a rule of inference given in section 11 of Bg--essentially, Universal Generalization--we can infer

$$(33) \quad \text{————— } a \text{————— } a \equiv a$$

i.e., in modern notation,

$$(34) \quad (x) (x \equiv x)$$

from (31). How are we to understand (33) (or (34))? The standard rendering of (33) (or (34)) would be

(35) Everything is such that it has the same conceptual content as itself;

just as the standard rendering of

$$(36) \quad \text{————— } a \text{————— } a \text{ is a man,}$$

i.e., '(x)(x is a man)', would be

(37) Everything is such that it is a man.

However, there is a significant difference between (35) and (37). The property of being a man is a property of objects, not names of objects, and so, in (36), the Gothic letter which marks the argument place in '() is a man' is supposed to be replaced by a name which stands not for itself, but for whatever it is customarily taken to stand for. Ec, on the other hand, relates names of objects. Hence, the Gothic letters in (33) are supposed to be replaced by names which, in this context, stand for themselves. So, whereas in (37) the indefinite pronoun 'everything' is the antecedent of 'it' in (35), 'everything' cannot be taken to be the antecedent of the pronouns 'it' and 'itself'. Were we to combine (35) and (37), the result,

(38) Everything is such that it has the same conceptual content as itself and is a man,

would be either false or nonsense: that which is said to be a man cannot be the same thing as that which is said to have the same conceptual content as itself.

There appears to be no way of expressing any general properties of ec by means of quantifiers and variables.

For, consider

(39) $x \equiv y$,

where 'x' and 'y' are supposed to be variables, not constants. If the expressions flanking the ec-symbol are understood to stand for themselves, then (39) is not, as one would expect, an open sentence, but a closed sentence (whose truth value

and conceptual content escape me). For (39) says that the expression 'x' has the same conceptual content as the expression 'y'. This becomes explicit if the single quote marks we dropped earlier are reinserted. Fully spelled out, (39) should be

(40) 'x' ≡ 'y'.

For the same reason, (34) cannot properly be rendered as (35): the expressions occurring on either end of the ec-symbol stand for themselves, and so they are not bound by the universal quantifier. The full English rendering of (34) with single quotes reinserted would be,

(41) Everything is such that 'it' has the same conceptual content as 'itself'.

Nor, can the problem be solved by shifting the universe of discourse and having the variables range over expressions rather than the objects these expression denote. For, whether we construe (34) as

(42) Every object is such that 'it' has the same conceptual content as itself;

or as

(43) Every expression is such that 'it' has the same conceptual content as 'itself',

again with single quote marks reinserted, we find that we cannot express the generalization.. Choice of a range for

the variables is pointless because we have no variables to range over anything.

Frege failed to observe this difficulty in his theory partly, I think, because he failed to clearly separate the objectual from the substitutional interpretation of the quantifiers (although Frege obviously intended the quantifiers to be understood objectually). This is not surprising. For, objectual/substitutional confusion in the interpretation of quantifiers is the direct analogue for quantification of use/mention confusion in the interpretation of singular terms. Here, for example, is the passage in which Frege introduced the quantifier-variable notation in Bg:

In the expression of a judgment we can always regard the combination of signs to the right of \vdash as a function of one of the signs occurring in it. If we replace this argument by a German letter and if in the content stroke we introduce a concavity with this German letter in it, as in

$$\vdash \underbrace{\alpha}_{\phi\alpha}$$

this stands for the judgment that, whatever we may take for its argument, the function is a fact. Since a letter used as a sign for a function, such as ϕ in $\phi(A)$, can itself be regarded as the argument of a function, its place can be taken, in the manner just specified, by a German letter.³⁵

On the substitution interpretation of the universal quantifier,

(x)Fx

reads as

Every substitution instance of 'Fx' is true,

rather than as

Everything is such that it is F,
 the more familiar objectual reading. The main virtue of substitutional quantification is that it succeeds in blurring use/mention distinctions that are prominent in objectual quantification; the substituends of substitutionally interpreted variables need not be names, i.e., they need not stand for or denote anything. Of particular interest to us is the fact that quantification into contexts of quotation is not problematic on the substitutional interpretation. As Dunn and Belnap observe,

. . . the substituends of variables need not be restricted to be names (nouns). Variables may occupy the place of expressions of other grammatical categories, for example, sentences. With this convention firmly in mind, we can say things like

(11) (P)(Q) if "P" and "Q" are sentences, then "P&Q" is a sentence, and "P" and "Q" are its conjuncts.

And we can do so without running the risk of making a mistake about use and mention through inadvertently referring to the sixteenth letter of the English alphabet by using "'P'".

Indeed, if we also have the convention that signs of the object language are used autonomously as names of themselves, then we may drop the quotes without having to introduce any quasi-quotes or other ugly notational conventions:

(12) (P)(Q) if P and Q are sentences, then P&Q is a sentence, and P and Q are its conjuncts.

The substitution interpretation thus gives logicians a way of preaching what they practice.³⁶

But, were we to adopt substitutional quantification, we should still not be able to save the Bg theory. For the

substitutional interpretation has the important feature of shaking loose names from their role as names, i.e., as expressions which stand for or denote something. And it is clear that Frege wished to construe the expressions flanking the ec-symbol as names, indeed, as names of themselves, because Frege wished to maintain that the name occurring in (44) $\text{Mark Twain} \equiv \text{Mark Twain}$, is the same name as that which occurs in (45) $\text{Mark Twain wrote } \underline{\text{Innocents}} \underline{\text{Abroad}}$, even though it names itself in (44) and it names the man in (45). This cannot be consistently maintained on the substitution interpretation. For, as Binkley has recently shown, although we can make sense of

$$(x)F'x'$$

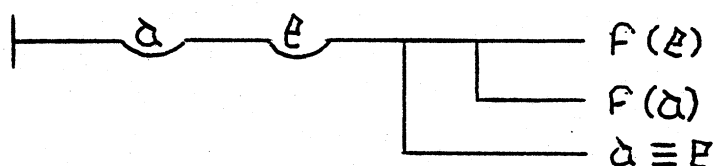
on the substitution interpretation, we cannot make sense of

$$(x) (F'x' \ \& \ Gx)$$

where we have, as Binkley calls it, a mixing of levels, i.e., where a single quantifier binds occurrences of a variable, some of which occur within quotation marks, some of which do not.³⁷

The problems we have run into in this section bear out Furth's remark quoted earlier that it is impossible to incorporate the general theory of identity into Frege's Begriffsschrift. We simply cannot give any coherent

interpretation to a formula like



(i.e., the universal closure of (32)) where we bind two occurrences of a variable and only one of these is an occurrence at either end of an ec-symbol.

7. The Failure of the Begriffsschrift Theory of Identity

At the beginning of the chapter, we noted that if a sentence, $S\alpha$, is about $d(\alpha)$, then it ought not matter which of the many expressions which could stand for the object is actually used in the sentence. Of course, there was some question about what does not matter, i.e., about what remains invariant under substitution of codenotational expressions, whether it is truth value or conceptual content. But, for the sake of argument, we accepted Frege's suggestion that conceptual content is the invariant property. Unlike Frege, however, we kept the substitution principle, (II), firmly in mind when it came to considering ec. In section 3, we took seriously Frege's claim that ec relates expressions, and we supposed that if ec is a relation which holds between expressions, then it ought not matter which of the many names which could stand for a given expression actually serves in an ec-sentence. But it did matter: it turned out that, on the Bg theory, the

expressions flanking the ec-symbol must be understood to stand for themselves.

Frege's claim that an ec-sentence expresses a relation which holds between the expressions flanking the ec-symbol certainly violates the spirit of the substitution principle. Indeed, it cleaves to the letter of the principle only trivially. For, if an expression in an ec-sentence stands for itself, then there is no distinct expression which could stand for the same thing in that context. That is, there are no two distinct singular terms, α and β , such that if each is understood to stand for itself, then each stands for the same thing. The substitution principle is, then, satisfied simply because there is no codenotational substitute for an expression that occurs in an ec-sentence. But why should we suppose that the expressions stand for anything? After all, the substitution principle would be satisfied in the same way--i.e., vacuously--were we to suppose that a sentence (any sentence) is about each and every expression occurring in the sentence. This is surely absurd, for there seems to be a distinction worth preserving here, namely, between saying that an expression stands for itself and saying that an expression does not stand for anything at all.

On closer inspection, we uncovered a further problem with ec-sentences. Previously, we had found that there could not be a distinct codenotational substitute

for an expression flanking the ec-symbol. Then in section 5, we found that we could not even locate the expressions flanking the ec-symbol. They had become welded to the whole ec-sentence in a way that prevented us from logically parsing the sentence into function-expression and argument-expressions.

Finally, we considered how ec-sentences fare in arguments. An ec-sentence is supposed to express a relation which holds between the singular terms occurring at either end of the ec-symbol: each singular term stands for or denotes itself, and the sentence as a whole expresses the circumstance that the one expression has the same conceptual content as the other. The reason Frege understood the singular terms to be standing for or denoting themselves was that the conceptual content of an ec-sentence is affected by the expressions occurring in it. Now, the conceptual content of a sentence, recall, is that part of the content of a sentence which counts for inference. Hence, one would expect that the inferential relations an ec-sentence enters into will be determined by the conceptual content of the sentence, and so, by the singular terms which occur in the sentence. However, the only way the singular terms contributed to our logical treatment of the ec-sentence was in determining whether one ec-sentence was the same as another (in the sense of same sequence of expressions). We could not formally capture those inferences which depended upon

our intuitive understanding of ec.

The evidence, then, is overwhelming that the Bg theory of identity is a failure. However, our criticisms of the theory have turned on technical considerations, and it is, therefore, not easy to understand where the theory went wrong. In this final section, I should like to step back and attempt to explain the error, and I shall do so by examining a suggestion that has been advanced by some commentators, namely, that the Bg account of identity is circular or involves a vicious regress.

Recall the first clause of Frege's definition of ' $A=B$ ', which was, with single quotes inserted appropriately,

(46) The symbol 'A' has the same conceptual content as the symbol 'B'.

In The Principles of Mathematics, Russell remarked that, if taken as a definition of identity, (46), "verbally at least, suffers from circularity."³⁸ Russell's brief, and slightly hedged, criticism of Frege's Bg theory of identity was fleshed out recently by David Wiggins, who argues that (46) generates a vicious infinite regress:

Asking for the sense of ' $a=b$ ', I am told 'a' and 'b' have the same content, or designate only one thing. Unless something is said to justify calling a halt here, the explanation generates a new statement of the same form as the original explicandum--'The content or designatum of "a" = the content or designatum of "b".' Applying the same explanation to this we get 'The content or designatum of "the content or designatum of 'a'" = the content or designatum of "the content or designatum of 'b'".' But evidently

we can never reach in this way what seems to be needed to carry the explanation through, a statement only about signs.³⁹

However, it is not at all clear that Wiggins has demonstrated that (46) generates a vicious infinite regress. For it appears that he has misinterpreted *ec* as a relation which holds between contents, not signs. In Grundlagen, Frege noted:

. . . instead of "the segments are identical in length", we can say "the length of the segments is identical" . . . , and instead of "the surfaces are identical in color", "the color of the surfaces is identical".⁴⁰

Following Frege's example, we could rephrase (46) to read,

(47) The conceptual content of the symbol 'A' is the same as the conceptual content of the symbol 'B'.

The crucial step, however, would be to take (47) as

(48) The conceptual content of the symbol 'A' = the conceptual content of the symbol 'B',

for (48) immediately reduces to

(49) $A = B$,

via the obvious identities,

(50) The conceptual content of the symbol 'A' = A,

and

(51) The conceptual content of the symbol 'B' = B.

So, were to interpret (46) as (48), *ec* turns out to be none other than old-fashioned identity dressed in strange clothes. Now, it appears that this is the way Wiggins has taken (46) in the quoted paragraph, and, if so, then not only has he

shot wide of his mark, but he has also failed to generate the regress, because each sentence he comes up with is a mere reformulation of the original. Following the examples of (50) and (51), we find that 'the content or designatum of "the content or designatum of 'a'"' denotes the same thing as 'the content or designatum of "a"', which in turn, denotes the same thing as 'a'. Applying this same procedure to the expression 'the content or designatum of "the content or designatum of 'b'"', Wiggins' convoluted sentence, 'The content or designatum of "the content or designatum of 'a'" = the content or designatum of "the content or designatum of 'b'"' just reduces to 'a=b'.

Although Wiggins' argument is inconclusive, I think that he and Russell are on the right track. There is something odd about (46), and this hunch is strengthened by the fact we unearthed in section 5, namely, that ec-sentences must be treated as unanalyzable wholes, thereby preventing any general account of ec. The apprehension about (46) centers around the phrase 'the same conceptual content' and takes the following rough form. (46) says that the symbol 'A' and symbol 'B' have one and the same conceptual content--i.e., that they have numerically the same conceptual content. After all, had Frege meant that they have qualitatively the same conceptual content, he would have been explicating similarity or resemblance, but

not identity. However, if, as Frege believed, identity is inapplicable to contents, but only to signs, then the phrase 'the same conceptual content' must be eliminable in favor of some locution which expresses a relation which holds between the signs standing for this conceptual content. Yet (46) is Frege's attempt to do precisely that job, i.e., to express a relation which holds between the signs 'A' and 'B'. Hence, the air of circularity hovering about (46). But we have to make these intuitions more precise, and to do so, we again take our cue from Grundlagen:

The judgment "line a is parallel to line b", or, using symbols

$a//b$

can be taken as an identity. If we do this, we obtain the concept of direction, and say: "the direction of line a is identical with the direction of line b". Thus we replace the symbol $//$ by the more generic symbol $=$, through removing what is specific to the content of the former and dividing it between a and b. We carve up the content in a way different from the original way, and this yields us a new concept.⁴¹

Frege's remarks here coupled with those in the passage quoted earlier suggest the following idea. The sentence, (52) The direction of line a is identical with the direction of line b, might be logically analyzed in two different ways, each of which assigns it the form $(R(\alpha, \beta))$. We might, on the one hand, understand (52) to say that the relation of

identity holds between line-directions, in which case,

(52) would receive the following analysis:

(53) R: _____ is the same as _____

α : the direction of line a

β : the direction of line b

Or, on the other hand, we might understand (52) to say that the relation of sameness-of-direction holds between lines, in which case it would receive the following analysis:

(54) R: the direction of _____ is the same as the direction of _____

α : line a

β : line b

There might, of course, be other reasons for preferring one analysis to the other, but at this state, both are prima facie plausible.

Parallel to this treatment of (52) there would be two logical analyses of (47), each assigning it the form

$(R(\alpha, \beta))$:

(55) R: _____ is the same as _____

α : the conceptual content of the symbol 'A'

β : the conceptual content of the symbol 'B'

and

(56) R: The conceptual content of _____ is the same as the conceptual content of _____

α : the symbol 'A'

β : the symbol 'B'

If we analyze (47) in the manner of (55), then (47) is understood to express a relation which holds between the conceptual content of the symbol 'A' and the conceptual content of the symbol 'B', i.e., between A and B, namely, identity. If, on the other hand, we analyze (47) in the manner of (56), then (47) is understood to express a relation which holds between the symbol 'A' and the symbol 'B', namely, the relation of having the same conceptual content. Earlier, we had taken (47) in the manner of (55), but, to be fair to Frege, we shall have to understand (47) (or (46)) to be analyzed in the manner of (56), i.e., to express an equivalence relation which holds between the symbols 'A' and 'B'.

We are all familiar with equivalence relations. Mathematically, an equivalence relation is a binary relation which is reflexive, symmetric, and transitive. An equivalence relation defined over a set, S, partitions the set into a family of pairwise disjoint subsets. Identity, ordinarily understood, is also an equivalence relation, one which an object bears only to itself, and so each subset belonging to the family into which identity partitions S is a singleton. Clearly, the equivalence relation having the same conceptual content cannot be identity, for a given singular term might bear this relation to a singular term other than itself. That is, different, i.e., non-identical,

singular terms might stand to one another in the relation having the same conceptual content. However, this sameness or difference of singular terms must be understood in terms of ec, and it is in this fact that we can locate the circularity of (46).

Frege's claim is that identity must be construed not as a relation which holds between contents, but as a relation holding between the terms that specify these contents. However, in order to express the relation, we must specify the terms themselves, and so, an ec-sentence expresses an equivalence relation holding between the items denoted by the singular terms flanking the ec-symbol.

For example, the ec-sentence,

(46) The symbol 'A' has the same conceptual content as
the symbol 'B',

expresses the circumstance that an equivalence relation holds between the items denoted by "the symbol 'A'" and "the symbol 'B'". I would suppose that a minimal condition for grasping the conceptual content of a given ec-sentence is that one know which expressions occur in the sentence, and, in particular, whether the expressions flanking the ec-symbol are the same or different. Since the terms flanking the ec-symbol stand for themselves, however, one must therefore know whether the terms flanking the ec-symbol stand for the same item or not. Hence, in order to

grasp the conceptual content of a given ec-sentence, we must be able to determine whether the singular terms flanking the ec-symbol stand for the same--i.e., identical--item or not. In the case of (46), we must be able to determine whether the items denoted by the singular terms "the symbol 'A'" and "the symbol 'B'" are the same or not, and hence we must be able to determine the truth-value of,

(57) The symbol 'A' = the symbol 'B'.

Identity, however, does not relate contents, but the expressions that stand for them, and so, our determining whether the singular terms flanking the ec-symbol stand for the same item or not means that we shall have to determine whether the singular terms themselves have the same conceptual content. That is, (57) must be understood to express an equivalence relation which holds between the terms that denote the symbol 'A' and the symbol 'B'. But, since in this context the terms are understood to stand for themselves, the ec-sentence corresponding to (57) is none other than

(46) The symbol 'A' has the same conceptual content as the symbol 'B'.

Hence, in order to grasp the conceptual content of a given ec-sentence we must already know it, and this is a very small circle indeed.

It is for this reason, then, that we had to treat an ec-sentence as an unanalyzable whole. We do not come to

grasp the conceptual content of an ec-sentence by exercising our knowledge of the conceptual content of the parts of the sentence and the way in which the conceptual contents of these parts combine to form the conceptual content of the whole. Rather, we come to grasp the conceptual content of a given ec-sentence in one fell swoop. Each ec-sentence is, so to speak, sui generis.

We can sum up the results of our analysis of the Bg theory of identity as follows. Sentences of the form $(\alpha=\beta)$ are, by and large, informative, whereas sentences of the form $(\alpha=\alpha)$ never are. The reason seems to be that $(\alpha=\alpha)$ is a logical truth, whereas $(\alpha=\beta)$ need not be a logical truth. Were the Bg theory faithful to these facts, sentences of the form $(\alpha\equiv\beta)$ should be, by and large, informative, whereas sentences of the form $(\alpha\equiv\alpha)$ should not. Unfortunately, it turns out that no ec-sentence is trivial, because the important distinction between those ec-sentences which are logically true and those which are not breaks down.

Footnotes to Chapter 2

¹Gottlob Frege, Begriffsschrift, eine der Arithmetischen Nachgebildete Formelsprache des Reinen Denkens (Halle, 1879); reissued in Begriffsschrift und Andere Aufsätze, ed. Ignacio Angelelli (Hildesheim: Olms, 1964). Portions of Chapter I of Begriffsschrift have been translated by Peter Geach and appear in Translations from the Philosophical Writings of Gottlob Frege, ed. Peter Geach and Max Black (Oxford: Blackwell, 1960). A complete translation by Stefan Bauer-Mengelberg appears in From Frege to Gödel: A Sourcebook in Mathematical Logic, 1879-1931, ed. Jean van Heijenhort (Cambridge: Harvard University Press, 1967), pp. 1-82.

²Gottlob Frege, "On Sense and Reference," trans. Max Black, in Translations from the Philosophical Writings of Gottlob Frege, ed. Peter Geach and Max Black (Oxford: Blackwell, 1960), p. 56.

³Frege, Begriffsschrift (trans. Geach), op. cit., p. 12.

⁴Ibid., p. 10

⁵Leonard Linsky, Referring (New York: Humanities Press, 1967), p. 22.

⁶William Kneale and Martha Kneale, The Development of Logic (Oxford: Oxford University Press, 1962), p. 494.

⁷Bertrand Russell, The Principles of Mathematics, 2nd ed. (New York: Norton, 1938), p. 502.

⁸David Wiggins, "Identity Statements," in Analytic Philosophy, Second Series, ed. R. J. Butler (Oxford: Blackwell, 1965), p. 51.

⁹Kneale and Kneale, op. cit., pp. 494-495.

¹⁰Alonzo Church, "A Formulation of the Logic of Sense and Denotation," in Structure, Method, and Meaning: Essays in Honor of Henry M. Sheffer, ed. P. Henle, H. M. Kallen, and S. K. Langer (New York: Liberal Arts Press, 1951), p. 3.

¹¹Montgomery Furth, trans. and ed., The Basic Laws of Arithmetic: Exposition of the System, by Gottlob Frege (Berkeley and Los Angeles: University of California Press, 1967), p. xix. This is a translation of portions of Frege's Grundgesetze der Arithmetik, I (1893), II (1903) (rpt. Hildesheim: Olms, 1962).

¹²Largely on the basis of the fact that, in articles written to explain his Begriffsschrift, Frege makes no mention of any special treatment of identity. See, e.g., "On the Purpose of the Begriffsschrift," trans. H. V. Dudman, Australasian Journal of Philosophy, 46 (1968), 89-97.

"Über den Zweck der Begriffsschrift" originally appeared in the Jenaische Zeitschrift für Naturwissenschaft, Supplementar-Heft, 16 (1882-83), 1-10, and it has been reprinted in Angelelli, op. cit..

¹³See, e.g., his comments in Furth, op. cit., pp. 6-7.

¹⁴W. V. O. Quine, Methods of Logic, 3rd ed. (New York: Holt, Rinehart and Winston, 1972), p. 221.

¹⁵Ibid., p. 222.

¹⁶Ibid., p. 222.

¹⁷Gottlob Frege, "On the Scientific Justification of a Concept-Script," trans. James M. Bartlett, Mind, LXXIII (1964), 156. "Über die Wissenschaftliche Berechtigung einer Begriffsschrift" originally appeared in the Zeitschrift für Philosophie und philosophische Kritik, 81 (1882), 48-56, and it has been reprinted in Angelelli, op. cit., pp. 106-114.

¹⁸Frege, Begriffsschrift (trans. Geach), op. cit., p. 3.

¹⁹Frege, "On the Scientific Justification of a Concept-Script," pp. 156-157.

²⁰Ibid., p. 157.

²¹Frege, Begriffsschrift (trans. Geach), op. cit., pp. 2-3.

²²Peter Geach, Mental Acts (London: Routledge & Kegan Paul, 1957), p. 83.

²³A. Tarski, "The Concept of Truth in Formalized Languages," rpt. in Logic, Semantics, Metamathematics (Oxford: Oxford University Press, 1956), pp. 156 ff.

²⁴Ibid., p. 159

²⁵W. V. O. Quine, Mathematical Logic, rev. ed. (Cambridge: Harvard University Press, 1951), p. 26.

²⁶W. V. O. Quine, "Three Grades of Modal Logic," rpt. in The Ways of Paradox and Other Essays (New York: Random House, 1966), p. 159.

²⁷W. V. O. Quine, Word and Object (Cambridge: MIT Press, 1960), pp. 143-144.

²⁸Frege, Begriffsschrift (trans. Geach), op. cit., p. 12.

²⁹Ibid., p. 14.

³⁰Ibid., p. 13.

³¹Frege, Begriffsschrift, rpt. Angelelli, op. cit., pp. 13-14.

³²Frege, Begriffsschrift (trans. Bauer-Mengelberg), op. cit., p. 20.

³³Ibid., p. 34.

³⁴See especially, "Reference and Modality" in From a Logical Point of View, 2nd ed. rev. (Cambridge: Harvard University Press, 1961), pp. 139-159.

³⁵Frege, Begriffsschrift (trans. Bauer-Mengelberg), op. cit., p. 24.

³⁶M. Dunn and N. Belnap, "The Substitution Interpretation of the Quantifiers," Noûs, II (1968), 185.

³⁷R. Binkley, "Quantifying, Quotation, and a Paradox," Noûs, IV (1971), 271-278.

³⁸Russell, op. cit., p. 502.

³⁹Wiggins, op. cit., p. 51.

⁴⁰Gottlob Frege, The Foundations of Arithmetic, 2nd ed. rev., trans. J. L. Austin (Oxford: Blackwell, 1953), p. 75. This is a translation of Die Grundlagen der Arithmetik, eine logisch-mathematische Untersuchung über den Begriff der Zahl (Breslau, 1884).

⁴¹Ibid., pp. 74-75.

CHAPTER 3

FREGE'S SENSE/REFERENCE THEORY

1. Introduction

Frege's elegant SR solution to the Paradox of Identity can be seen as proceeding in two parts. First, he shows the argument itself to be unsound: substitution of codenotational singular terms preserves truth value, not, as had been assumed in the argument, cognitive value. And from this correct substitution principle, together with the assumption that "identity relates objects," it does not follow that $(\alpha=\alpha)$ and true $(\alpha=\beta)$ cannot differ in cognitive value. Second, Frege offers an explanation for this difference in cognitive value between $(\alpha=\alpha)$ and true $(\alpha=\beta)$. There is associated with each singular term, he says, besides its reference (if it has one), also a sense, i.e., a particular way of picking out, presenting, or determining the reference; and just as it is the reference of a singular term that is relevant for determining the truth value [Warheitswerth] of a sentence containing that term, so it is the sense of a singular term that is relevant for determining the cognitive value [Erkenntniswerth] (literally, knowledge value) of a sentence containing that term. In so far, then, as α and β are singular terms that have the same reference but different sense,

$(\alpha=\beta)$ will, like $(\alpha=\alpha)$, be true, but $(\alpha=\beta)$ will be informative in a way $(\alpha=\alpha)$ is not.

When we found 'a=a' and 'a=b' to have different cognitive values, the explanation is that for the purpose of knowledge, the sense of the sentence, viz., the thought expressed by it, is no less relevant than its reference, i.e. its truth value. If now a=b, then indeed the reference of 'b' is the same as that of 'a,' and hence the truth value of 'a=b' is the same as that of 'a=a.' In spite of this, the sense of 'b' may differ from that of 'a,' and thereby the thought expressed in 'a=b' differs from that of 'a=a.' In that case the two sentences do not have the same cognitive value.¹

Here, by way of illustration, is Frege's own famous example. The true but trivial sentence,

(1) The evening star = the evening star,

can be obtained from the true, but far from trivial, sentence,

(2) The evening star = the morning star,

by substituting 'the evening star' for 'the morning star'. These two singular terms have the same reference, the planet Venus, and, as is required by the substitution principle for reference, (1) and (2) have the same truth value. But (1) and (2) do not have the same cognitive value: (1) is an instance of the Law of Identity, while (2) expresses a significant astronomical discovery. This difference in cognitive value between (1) and (2) is attributed to the fact that, while 'the evening star' and 'the morning star' have the same reference, they do not have the same sense: 'the evening star' picks out the planet Venus as that bright object in the heavens

appearing just over the horizon at dusk, while 'the morning star' picks out the planet Venus as that bright object in the heavens appearing just over the horizon at dawn. That the bright object in the heavens appearing just over the horizon at dusk is the same as the bright object in the heavens appearing just over the horizon at dawn is, of course, just what constitutes the great astronomical discovery; and, according to Frege, that the bright object in the heavens appearing just over the horizon at dusk is the same as the bright object in the heavens appearing just over the horizon at dawn is just what constitutes the cognitive content of (2), i.e., the thought it expresses.

Frege's sharp distinction in SR between an expression, its sense, and its reference, successfully eliminates the logical difficulties of his Bg theory, and it does not appear to have generated any similar logical difficulties of its own. To be sure, a number of philosophers have raised important questions about a criterion for determining when two expressions have the same sense; but the notorious difficulties in finding such a criterion indicate a lack of clarity about the notion of sense, not, as yet, any deep-rooted logical incoherence. The only philosopher who has presented a significant argument to show that the sense/reference

distinction is logically incoherent is Russell. "The whole distinction of meaning and denotation has been wrongly conceived,"² he claimed, for "we cannot succeed in both preserving the connexion between meaning and denotation and preventing them from being one and the same."³ But the "inextricable tangle" Russell claimed to find in the sense/reference distinction appears to have been of his own making: as has been amply demonstrated by Church⁴ and Searle,⁵ Russell's argument is riddled with confusion between use and mention, and nobody has been able to present a comprehensible reconstruction of his reasoning.

The likelihood of finding logical difficulties with the SR theory comparable to those we had found with Frege's Bg theory, then, is rather small; and so we immediately move to the question of the adequacy of his explanation of the difference in cognitive value between $\langle \alpha = \alpha \rangle$ and true $\langle \alpha = \beta \rangle$. Here, I think it would be fair to say that Frege's SR explanation is widely believed to be the correct one, save for one very important exception, namely, when α and β are proper names.

Briefly, the problem is this. According to Frege, the sense of a singular term determines its reference. When our singular term is a definite description, the way in which sense determines reference is readily grasped: the referent of, say, 'the 37th president of the United

States' is Richard Nixon, for he is the unique individual possessing the properties specified in the description. But when our singular term is a proper name, it is not so obvious how the sense determines the reference. Frege almost invariably gives as the sense of a proper name some definite description, and, although it is debatable whether this was in fact his view, Frege has been widely believed to have held, like Russell, that proper names are simply abbreviations for definite descriptions. Now, if proper names are disguised descriptions, then, of course, the way in which sense determines reference for proper names will be essentially the same as the way in which sense determines reference for definite descriptions. But the view that proper names are simply disguised descriptions is not particularly convincing; and, in so far as this view of proper names appears to be required by his proposed solution to the paradox, many philosophers have to this extent been unhappy with Frege's solution.

There are two sorts of criticism of Frege's view of proper names. On the one hand, philosophers like Searle⁶ and Dummett⁷ believe Frege's notion of sense for proper names to be too rigid to capture adequately the use of proper names in ordinary language, and they have proposed that instead of supposing a proper name to be associated with a single criterion for identifying the reference, that one view the name rather as being

associated with a cluster of such criteria. This criticism of Frege, however, is seen, both by Frege's critics as well as his defenders, as not essentially involving any philosophically important change in the view of proper names; it is taken, rather, in the spirit of an emendation. On the other hand, we have those philosophers in the tradition of Mill,⁸ who mark a sharp distinction between proper names and definite descriptions; on this view, proper names have no meaning (in the sense of connotation), but serve simply as labels or tags for the things to which they have been assigned. Mill's view has been out of fashion for many years, for two basic reasons: (1) on Mill's view, unlike Frege's, there is no explanation of how the reference of a proper name is determined; and (2), on Mill's view, unlike Frege's, there is no explanation of the difference in cognitive value between ' $\alpha=\alpha$ ' and true ' $\alpha=\beta$ ', when α and β are proper names. Recently, however, Mill's view of proper names has been championed in a new and interesting manner by Saul Kripke.⁹ Kripke has addressed himself directly to both of the points just mentioned; I will consider the first point in the present chapter, and the second point in the next chapter.

I will proceed in this chapter as follows. First, I will trace the connections between the Bg theory and the SR theory, showing how the latter evolved from the former.

Second, I will outline some of the main points of the SR theory. And lastly, I will consider Kripke's attack on what he calls the "Frege-Russell description theory of proper names" as well as Dummett's defense of the theory. I will argue in this last section that Dummett has failed to defend Frege adequately against Kripke's criticisms.

2. Bestimmungsweisen and Frege's Criticism of the Bg Theory

There remains one very important element of the Bg theory yet to consider: the Bestimmungsweise, i.e., the mode of determination, of an object.

As we reconstructed the Bg theory in the last chapter, Frege held that a sentence stands for or denotes a thought [content] and that the parts of the sentence, in turn, stand for or denote corresponding parts of the thought. We noted, however, that Frege deviated from this scheme when it came to singular terms, for he held that a singular term denotes not, say, something like Carnap's individual concept, but rather the object itself; and this deviation led to difficulty when he attempted to account for the informativeness of i-sentences. He could not regard identity in the usual manner, namely, as a relation holding between objects, because the substitution principle required that replacement of one singular term by another codenotational singular term preserve conceptual content,

and so, where α and β are distinct singular terms, the true i-sentence ' $\alpha=\beta$ ' would have the same conceptual content as the logically true ' $\alpha=\alpha$ '. However, Frege struck on the idea of taking identity to be a relation holding between expressions. Thus, if α and β are distinct singular terms, ' $\alpha\equiv\alpha$ ' and true ' $\alpha\equiv\beta$ ' would not have the same conceptual content, for where the content denoted by ' $\alpha\equiv\beta$ ' would have β as a constituent, the content denoted by ' $\alpha\equiv\alpha$ ' would have α as a constituent. On the other hand, since singular terms stand for their ordinary contents in all other contexts, the substitution principle would appear to be satisfied because if ' $\alpha\equiv\beta$ ' is true, then α and β have the same conceptual content.

This is how we presented the Bg theory in the last chapter. We then went on to show that the superficial appearance of success was deceptive and that the proposed solution of the Paradox of Identity failed.

However, the Bg theory was a bit more complicated than we had indicated, and thus our reconstruction was somewhat inaccurate. We said that Frege had attributed the informativeness of an i-sentence to the difference in the singular terms flanking the identity sign, whereas he had actually attributed the informativeness to the difference in the way each of the terms determines its content. Hence, although he construed identity as ec,

i.e., as relating expressions, it was the Bestimmungsweisen associated with expressions he was really after.

Much of the bulk of Frege's discussion of ec was devoted to the notion of Bestimmungsweise, and it was aimed precisely at dispelling the impression that ec was solely concerned with expressions. Thus, immediately after introducing the ec-symbol and explaining the unusual convention that the terms at either end of the ec-symbol were to stand for themselves, Frege noted:

At first sight this makes it appear as though it were here a matter of something pertaining only to expression, not to thought; as though we had no need of two symbols for the same content, and therefore no need of a symbol for equality of content either.¹⁰

Frege was anxious, then, lest the reader believe that ec was only incidentally, if at all, connected with thought; for, were this so, there would seem to be little need for the ec-symbol in what was supposed to be a formula language of pure thought.

Frege had good reason to be apprehensive on this score, for a strong prima facie case can be developed against including the ec-symbol in his Begriffsschrift. Consider: if α and β have the same conceptual content, then $S\alpha$ and $S\alpha/\beta$ also have the same conceptual content, excepting, of course, when $S\alpha$ is an ec-sentence. It would seem, then, to be of no logical consequence whether we use the one term or the other in a given sentence, excepting,

again, an ec-sentence. But $\langle \alpha \equiv \alpha \rangle$ and $\langle \alpha \equiv \beta \rangle$ differ in conceptual content (if α and β are distinct signs) only because of what now appears to be the purely ad hoc device of having α and β stand for themselves. The information thus obtained is of very limited applicability--limited, that is, to ec-sentences: were we to excise the ec-symbol, the remainder of the logical symbolism would be indifferent to α and codenotational β . A comparison here with Frege's treatment of the active/passive distinction in Bg is quite telling. Since each one of an active/passive pair of sentences has the same conceptual content as the other, Frege ignored this grammatical distinction in his Begriffsschrift and symbolized each in the same way. Why didn't he adhere to this practice for sentences which differ only in that where the one contains α the other contains codenotational β ? The difference in singular terms fails to reflect a difference in conceptual content, so the indicated course would be to ignore the trivial difference in formulation, symbolize each in the same way, and thus eliminate the need for the ec-symbol altogether.

The assumption apparently underlying this objection is that names are meaningless marks, arbitrarily chosen labels or tags which simply stand for objects but otherwise carry no meaning; and it is this assumption Frege rejected in Bg. Instead, he urged:

. . . different names for the same content are not always just a trivial matter of formulation; if they go along with different ways of determining the content, they are relevant to the essential nature of the case.¹¹

That is, Frege denied that the whole significance of a name is captured by specifying its content: a name does not merely stand for its content, it also goes along with a particular way of determining this content. Frege illustrated this as follows. Fix a point, A, lying on the circumference of a given circle and pass a straight line through A, extending the line so that it intersects with the circle. This point of intersection, which we will call 'B'. obviously depends upon the position of the straight line, so that as the line is rotated about A, B varies accordingly.

We may now ask: What point corresponds to the position of the straight line in which it is perpendicular to the diameter? The answer will be: The point A. The name B thus has in this case the same content as the name A; and yet we could not antecedently use just one name, for only the answer to the question justified our doing so. The same point is determined in a double way:

- (1) It is directly given in experience [Anschauung];
- (2) It is given as the point B corresponding to the straight line's being perpendicular to the diameter.¹²

One moral of this example is that, unlike the active/passive case where sameness of conceptual content is immediate, it is not always obvious whether two singular terms, α and β , happen to have the same conceptual content

(and so, it is not always obvious whether $S\alpha$ and $S\alpha/\beta$ have the same conceptual content. But this is only part of the story. For, since we can generate additional names for a given object as whim dictates, these names differing only trivially in formulation, then the non-obviousness might be attributed to the unpredictable creativity of the language-using community. The main point Frege wishes to establish, I take it, is that a singular term carries with it something, a Bestimmungsweise, which enables us to reason out whether, given a particular object, the term stands for it, and thus to reason out whether the term stands for the same object as does some other singular term. This, then, provides the link between names and thought. In the example, the same point is determined in two different ways, and the nontrivial fact that the same point is determined in each of these two different ways is, as I understand it, adequately expressed by the true ec-sentence ' $A \equiv B$ '.

Finally, now, Frege sums up the connection between names, Bestimmungsweisen, and ec:

The need of a symbol for equality of content thus rests on the following fact: The same content can be fully determined in different ways; and that, in a particular case, the same content actually is given by two ways of determining it, is the content of a judgment. Before this judgment is made, we must supply, corresponding to the two ways of determination, two different names for the thing thus determined. The judgment needs to be expressed by means of a symbol for equality of content, joining the two names together.¹³

(Another reason Frege offered for including the ec-symbol in his Begriffsschrift was that it would be needed to introduce definitions, but this was only a "superficial reason.")

Those who are already familiar with the SR theory will immediately recognize the resemblance between the Bg notion of Bestimmungsweise and the SR notion of an Art des Gegebenseins or Darstellungsweise, i.e., way in which an object is given or presented, this latter being the active ingredient of an expression's sense [Sinn]. Indeed, Frege's example in SR of an object's being presented in different ways could easily have served in Bg as an example of an object's being determined in different ways:

Let a, b, c be the lines connecting the vertices of a triangle with the midpoints of the opposite sides. The point of intersection of a and b is then the same as the point of intersection of b and c. So we have different designations for the same point, and these names ('point of intersection of a and b,' 'point of intersection of b and c') likewise indicate the mode of presentation; and hence the statement contains actual knowledge.¹⁴

This similarity with the SR theory can be pushed still further. Thus, as in SR, where Frege distinguished between that which a term stands for, its reference [Bedeutung], and that which the term expresses [Ausdrucken], its sense, so in Bg Frege distinguished between that which a term stands for, its content, and that which the term goes along with [Entsprechenden], a Bestimmungsweise. And, again, as in SR, where Frege distinguished between that

which identity relates and that wherein the information conveyed by an i-sentence resides, so in Bg, it now appears that although identity is to relate the terms flanking the identity sign, the information is to be that the same content is given by two ways of determining it.

I readily concede that the status of Bestimmungsweisen within the Bg theory is problematic-- in fact, I shall argue below that they cannot carry the load Frege assigned to them--and I also admit to some exaggeration in seeking parallels between the Bg and SR theories. But I take this license because it has gone largely unnoticed that Frege had such a notion in Bg, and thus a widespread misconception has developed about the Bg theory and about Frege's reasons for rejecting it. It is commonly believed that Frege had attributed the informativeness of an i-sentence in Bg to the difference in the terms flanking the identity sign; that he became disenchanted with the Bg account because on close analysis it turned out that the information thus conveyed could only be of the arbitrary conventions of the language-using community; that he subsequently discovered (some years after Bg) that a term has, besides its denotation, a sense, whose connection with the denotation is not a matter of convention; and that this discovery allowed him, in SR, to locate the "actual" or "proper" knowledge conveyed by an i-sentence in the difference in the senses of the terms

flanking the identity sign.

Obviously, this story cannot be correct. For, as we have just seen, (1) Frege already had a notion in Bg, that of a Bestimmungsweise, corresponding roughly to the SR notion of sense; (2) he had attributed the informativeness of an i-sentence in Bg to the difference in the Bestimmungsweisen associated with the expressions flanking the identity sign, not merely to the difference in the expressions alone; and (3), he had introduced Bestimmungsweisen precisely to counter the charge that he had trivialized i-sentences by construing identity as ec.

Clearly, then, our omission of the notion of Bestimmungsweise in the last chapter is no reflection of the weight Frege attached to it, either in Bg or in later writings. Our decision to leave the matter until now was based largely on the following reason. We wanted to show that the Bg idea of taking identity to be a relation holding between expressions was simply incoherent, and we did not want the argument cluttered up by considerations that would inevitably arise as a result of injecting the notion of Bestimmungsweise into the discussion. This was particularly important because received opinion faults the Bg theory for making identity too dependent upon linguistic convention, and we wanted to make it clear that the grounds for rejecting the Bg theory are much closer to hand.

Typical examples of this received view are

[Linsky] Nor is Frege able to accept the other of the two alternatives, that identity is a relation between names or signs of objects. Then 'a=b' would just say that the name 'a' and the name 'b' are names for the same thing. This analysis cannot be correct, Frege argues, because the fact that 'a' is a name for a and that 'b' is also a name for a results from a purely arbitrary agreement concerning the use of these marks (or sounds). Furthermore, when I say that Venus is the morning star I am conveying information about the heavens, not about our arbitrary use of signs.¹⁵

[Kneale] As we have seen, Frege suggested that a statement of identity must really be about the expressions appearing on the two sides of the identity sign, and he tried to make this clear by saying that '=' was to be understood as a symbol for identity of content between expressions. But he came to see later that this was not a satisfactory solution of the puzzle. For he realized that if the original statement [that the morning star is identical with the evening star] was not really about the planet Venus, but about the contents of certain phrases, it would belong to philology rather than to astronomy, which is obviously not the case, since the discovery of the identity of the morning star and the evening star was made by observation and calculation, not by reflection on the use of words.¹⁶

This criticism of the Bg theory goes, roughly: if a sentence expresses a relation holding between expressions rather than what those expressions stand for, then the sentence conveys no information about what those expressions stand for, but only about the expressions themselves, and since 'Venus = the morning star' conveys information about Venus and the morning star, '=' cannot relate the expressions 'Venus' and 'the morning star'. Yet, surely we are sufficiently sophisticated in the

techniques of modern philosophy, e.g., shifting between material and formal mode, semantic ascent, etc., to doubt the premise of the criticism. The relation denotes the same thing as clearly relates expressions, not what expressions stand for, and yet the sentence 'Venus' denotes the same thing as 'the morning star' serves, in certain circumstances, to convey information about the heavens, and in other circumstances, to convey information about words. Again, Linsky's claim that "the fact that 'a' is a name for a and that 'b' is also a name for a results from a purely arbitrary agreement concerning the use of these marks" is false. It might be arbitrarily agreed that 'a' names a and arbitrarily agreed that 'b' names b, yet although 'a' and 'b' name the same thing, it would be a flagrant intensional fallacy to conclude that it was arbitrarily agreed that 'a' and 'b' name the same thing. Linsky saddles Frege with this argument, but I find it hard to believe that Frege would commit just this kind of error when spelling out the SR theory.

The Linsky-Kneale criticism, however, is doubly wrong: not only is the argument weak, but the Bg theory is misrepresented in the way noted earlier. Linsky and Kneale both give the impression that Frege had supposed in Bg that the informativeness of an i-sentence resides simply in the difference in the terms flanking the identity sign; and this, we have seen, is not so. Their

free use of 'about' blurs the distinction Frege had attempted to draw in Bg, viz. that although identity is to be taken to be a relation holding between the terms flanking the identity sign, so that, if you like, an i-sentence is, in a sense, about the terms flanking the identity sign, the judgment expressed by an informative i-sentence is to be that the same content is given by two ways of determining it, so that, if you like, an i-sentence is, in another sense, about the ways of determining this content. This distinction is, I admit, not very clearly drawn in Bg, and it might be that Frege could not carry it through; but this requires argument--argument which Linsky and Kneale are incapable of supplying just because they fail to acknowledge the distinction Frege was groping for.

There is, then, little to recommend this criticism of the Bg theory, and I speculate that many, recognizing the weakness of this criticism, have been inclined to believe that the Bg theory was salvageable. I hope I have disinclined them. But, to allay any lingering doubt, we must show that our ignoring of Bestimmungsweisen in the last chapter in no way affects our case there against the Bg theory. So, we now argue that although Frege wished to attribute the informativeness of i-sentences to Bestimmungsweisen, these Bestimmungsweisen cannot be integrated into the general semantic structure of Bg and thus they are only a frill.

If, as Frege maintained in Bg, the informative identities were those in which the terms corresponded to different Bestimmungsweisen, why didn't he take identity to relate Bestimmungsweisen? All he need do, it would seem, is stipulate that the expressions flanking the identity sign stand for their Bestimmungsweisen--not for their ordinary contents and not for themselves. Then, identity would be construed as an equivalence relation holding between Bestimmungsweisen, so that, e.g., the i-sentence

Mark Twain is identical with Samuel Clemens
would be understood as

The Bestimmungsweise associated with 'Mark Twain' determines the same content as the Bestimmungsweise associated with 'Samuel Clemens'; and this would conform to his desire that an ec-sentence express the judgment that "the same content actually is given by two ways of determining it." The fact is, however, that Frege did not do this, and he did not explain why, though my guess is that he did not think of Bestimmungsweisen as parts of conceptual contents, but rather as intermediaries between signs and their contents (a consequence, I think, of the confusion in the notion of Inhalt: the most likely place for Bestimmungsweisen in conceptual contents had already been reserved for the objects denoted). And so, instead of this direct approach to Bestimmungsweisen, Frege construed identity as ec, i.e.,

as a relation holding between expressions, hoping thereby to get at the Bestimmungsweisen indirectly via the signs they go along with. But it is not clear that he succeeded. For, the conceptual content of ' $\alpha \equiv \beta$ ' differs from the conceptual content of ' $\alpha \equiv \alpha$ ' if α and β are distinct terms, whatever the Bestimmungsweisen. That is, an ec-sentence turns out to be informative so long as the expressions differ, and there is, thus, no distinction between those cases where α and β correspond to the same Bestimmungsweise and those cases where they do not.

Conceptual content is all Frege deems of importance in Bg, and since Bestimmungsweisen are not parts of conceptual contents, there is just no room for them in the Bg semantic structure. Thus, we have just seen that Bestimmungsweisen are irrelevant to the semantic interpretation of ec-sentences. And, if we now look at sentences that do not contain the ec symbol, we find that Frege is faced with a dilemma. Where $d(\alpha) = d(\beta)$, but the Bestimmungsweise associated with α is different from the Bestimmungsweise associated with β , either $S\alpha$ and $S\alpha/\beta$ have the same conceptual content or they do not. If $S\alpha$ and $S\alpha/\beta$ have the same conceptual content, then, as far as inference is concerned, and this is, recall, the sole interest of Bg, it makes no difference whether the Bestimmungsweisen α and β correspond to are the same or different. On the other hand, if $S\alpha$ and $S\alpha/\beta$ differ in

conceptual content, then the substitution principle would be violated.

Summing up, now, we have so far in this section attempted to establish the following points: first, that Frege had a notion of Bestimmungsweise in Bg which corresponds roughly to the SR notion of Darstellungsweise and to which Frege attached great importance; second, that, unfortunately, the notion of Bestimmungsweise turned out to be an idle excrescence of the theory; and third, that the fact that Frege had such a notion in Bg indicates much greater continuity between the Bg and SR theories than has hitherto been recognized by Frege's commentators. In the remainder of this section, we shall develop this last point and propose a reconstruction of how the SR theory grew out of the shortcomings of the Bg theory.

Frege had discarded the Bg theory of identity as early as Grundlagen. Given the view he defended in Grundlagen, namely, that numbers are logical objects, and given the key role identity plays in his definition of Cardinal Number, it is evident that he had to regard identity as a relation holding between objects. Had he persisted with the Bg idea of regarding identity as a relation holding between expressions (e.g., between numerals rather than numbers), he would, for example, have had a very difficult time distinguishing his own view from the naive formalism of Heine and Thomae which he attacked

so vigorously. Indeed, I think that it was Frege's rejection of formalism in mathematics which led him to appreciate the distinction between sign and thing signified, and thus to the rejection of the Bg theory of identity. Yes, oddly, Frege does not criticize, let alone mention, the Bg theory in Grundlagen. He doesn't even acknowledge any tension between the view that identity relates objects and the fact that i-sentences can be informative, and this despite his obviously having been aware of the tension, and, moreover, despite his having rested some of the fundamental theses of Grundlagen, e.g., the rejection of contextual definition and the explanation of how "the empty forms of logic come to disgorge so rich a content" in arithmetic, on the fact that i-sentences can be informative. Also, Frege gives only the barest account of this informativeness:

Why is it, after all, that we are able to make use of identities with such significant results in such diverse fields? Surely it is rather because we are able to recognize something as the same again even although it is given in a different way.¹⁷

We are, by now, acquainted with the family of notions to which this being given in a particular way belongs, but Frege says little about it in Grundlagen, and it is doubtful whether we can identify this being given in a particular way either with the Bg Bestimmungsweise or the SR Darstellungsweise. For, Frege was in between theories

in Grundlagen. He had overcome the use/mention confusion of Bg sufficiently to convince himself that the Bg account of identity was incorrect, but he had not yet penetrated the confusion in the Bg notion of Inhalt to formulate the sense/reference distinction (and to this extent, his understanding of the use/mention distinction was impaired). Thus, Grundlagen is rife with confusion between meaning and reference, as we see in this typical passage in which it is impossible to determine whether the alternative to a symbol's being empty is that it have a reference, that it have a sense, or both:

Everyone who uses words or mathematical symbols makes the claim that they mean something [sie etwas bedeuten], and no one will expect any sense [etwas Sinnvolles] to emerge from empty symbols. But it is possible for a mathematician to perform quite lengthy calculations without understanding by his symbols anything intuitable, or with which we could be sensibly acquainted. And that does not mean that the symbols have no sense [sinnlos]; we still distinguish between the symbols themselves and their content [Inhalt], even though it may be that the content [Inhalt] can only be grasped by their aid. We realize perfectly that other symbols might have been arranged to stand for the same things.¹⁸

Still, Frege's failure to come to grips with the Bg theory of identity in Grundlagen remains something of a mystery to me. The only plausible explanation I have been able to come up with is this. Though he was convinced that it could be consistently maintained both that i-sentences can be informative and that identity relates objects, he was,

due to his confusion about Inhalt, unable to provide a detailed alternative to the Bg theory, so he eschewed all theoretical discussion about the matter in Grundlagen and simply asserted what he was sure was true, secure that the theoretical underpinnings would be forthcoming after further analysis.

At any rate, it is not until SR, by which time Frege had honed his function/argument analysis and distilled out of the crude Bg Inhalt a sense and a reference, that he confronts the Bg theory of identity. Here is the famous opening passage from SR:

Equality gives rise to challenging questions which are not altogether easy to answer. Is it a relation? A relation between objects, or between names or signs of objects? In my Begriffsschrift I assumed the latter. The reasons which seem to favour this are the following. [Here follows Frege's version of the Paradox of Identity quoted in the last chapter.] What is intended to be said by a=b seems to be that the signs or names 'a' and 'b' designate the same thing, so that those signs themselves would be under discussion; a relation between them would be asserted. But this relation would hold between the names or signs only in so far as they named or designated something. It would be mediated by the connexion of each of the two signs with the same designated thing. But this is arbitrary. Nobody can be forbidden to use any arbitrarily producible event or object as a sign for something. In that case the sentence a=b would no longer refer to the subject matter, but only to its mode of designation; we would express no proper knowledge by its means. But in many cases this is just what we want to do. If the sign 'a' is distinguished from the sign 'b' only as object (here, by means of its shape), not as sign (i.e. not by the manner in which it designates something), the cognitive value of a=a becomes essentially equal to that of a=b, provided a=b is true. A

difference can arise only if the difference between the signs corresponds to a difference in the mode of presentation of that which is designated.¹⁹

This passage is, I recognize, very important, but, regrettably, my discussion will be somewhat disappointing. For, I find this passage very perplexing. It certainly looks as though Frege is criticizing the Bg theory here, but I am not clear what this criticism is--if it is criticism at all, for at times I see in this passage merely a rehearsal of the reasons Frege required Bestimmungsweisen to complete the Bg account of identity. This latter reading is, no doubt, incorrect, but it is due, in large measure, to Frege's misrepresenting of the Bg theory here. He gives the impression that in Bg he had located the informativeness of an i-sentence solely in the difference in the terms flanking the identity sign, and he then goes on to criticize this account for failing to recognize that the informativeness is to be attributed rather to the difference in the way the terms present their denotata. But, as we are now well aware, Frege had gone over much the same territory in Bg when he urged that the interesting and informative identities were those in which the different terms coincided with different Bestimmungsweisen. This is not to say, by the way, that I approve, or even understand, Frege's arguments for this claim; but only that one point Frege clearly wishes to

establish here, viz. that the informativeness of an i-sentence resides in the difference in the way the objects are presented or determined and not simply in the difference in the terms, is a point Frege had not only granted but explicitly argued for in Bg. So, if this passage does contain Frege's criticism of the Bg theory, something more must be going on. And, indeed, there is a second issue which occupies Frege through most of the passage (but which he does not clearly separate), namely: Can the interpretation of identity, whereby it is taken to be a relation holding between expressions, adequately capture differences in the way in which an object is presented or determined? We, supposing that Frege is criticizing the Bg theory here, expect a forthright negative answer to this question. Unhappily, the text seems inconclusive; the first half of the passage indicates that the answer is, "No," but the second half appears to leave open the door to such an interpretation if the expressions are treated as signs rather than as shapes. Thus, my perplexity.

Yet, I should like to think that Frege is criticizing the Bg theory in this passage, and I shall hazard a suggestion that his intention here is to establish the following two points: first, that the informativeness of an i-sentence derives from the difference in the way the terms flanking the identity sign present their denotata;

and second, that these modes of presentation must be dealt with directly rather than, as in Bg, indirectly through intermediary expressions. There is no question that Frege is arguing for the first point, and so only the second requires comment. My reasons for thinking that Frege is arguing for this latter point are: (1) it indicates recognition of the failure of the Bg theory to adequately incorporate Bestimmungsweisen into the semantic analysis of ec-sentences and thus, given the first point, cause to reject the Bg theory; (2) it marks the crucial difference in the way the Bg theory handles Bestimmungsweisen and the SR theory handles Darstellungsweisen; and (3), it makes sense out of the sign/shape distinction Frege introduces at the end of the passage. We saw earlier in this section that Bestimmungsweisen was irrelevant to the semantic analysis of ec-sentences because ' $\alpha \equiv \alpha$ ' and true ' $\alpha \equiv \beta$ ' would differ in conceptual content so long as α and β are distinct terms, whether they determine the object in the same way or in different ways. So, for example, ' $1 \equiv 1$ ' and ' $1 \equiv I$ ' would differ in conceptual content because of the trivial notational difference between the Roman 'I' and the Arabic '1'. What is desired, presumably, is that such trivial differences in shape be ignored so that ' $\alpha \equiv \alpha$ ' and true ' $\alpha \equiv \beta$ ' differ in conceptual content only when α and β determine the object differently. But were we to ignore these differences in shape, i.e.,

were we to regard α and β as signs, it is doubtful whether we could any longer regard ec as an equivalence relation holding between expressions. For, if ' $l \equiv l$ ' and ' $l \equiv I$ ' are to have the same conceptual content, ' l ' and ' I ' would have to stand for the same thing in this context; and, since ' l ' and ' I ' are distinct expressions, they could not stand for themselves but only for what they have in common, namely, the way in which they determine the object. Ec would thus turn out to be an equivalence relation holding between Bestimmungsweisen rather than expressions. So, the original motivation for taking identity to relate expressions rather than objects has been undercut.

Of course, our discussion here has been an attempt only at determining what Frege is arguing for in this passage. The arguments themselves have not been touched, and they will remain that way. I find it very difficult to reproduce Frege's arguments here, for a number of major issues are run together and the reasoning is so sketchy. And, again, I have the example of Linsky and Kneale before me, who, in rushing in, have made Frege look so foolish. Any adequate discussion of this passage would entail an examination of Frege's arguments against formalism,²⁰ and also a discussion of the fundamental question about what is and what is not to count as a purely linguistic matter. I should like to avoid this issue here. By and large, my stance in the present chapter will be to evaluate the SR

theory assuming that Frege has made this distinction.

Returning, now, to the main theme, my suggestion is that Frege was aware of the difficulties in integrating Bestimmungsweisen and conceptual contents in the Bg theory and that the SR solution was, essentially, to split up the Bg conceptual content into a thought and a truth value. Starting with the Bg conceptual content denoted by $S\alpha$, Frege removed the object denoted by α and placed in its stead the Bestimmungsweise going along with α , or, more accurately, its descendent, the sense of α , and the result of this transplant is the sense of $S\alpha$, a thought [Gedanke]. He then extended the relation between a singular term and its Bestimmungsweise to hold generally between any expression and its sense, so that a sentence expresses--not denotes--a thought, and the parts of the sentence, in turn, express parts of the thought. The object denoted by α , on the other hand, i.e., the reference of α , is not a part of the thought expressed by $S\alpha$, but rather combines with the reference of the remainder of $S\alpha$ (i.e., the sentence fragment resulting from deletion of α in $S\alpha$) to result in the reference of $S\alpha$, which Frege took to be the truth value of $S\alpha$. And, whereas we had only one substitution principle in Bg, we now have two. In general, substitution of coexpressional terms, i.e., terms expressing the same sense, preserves sense; substitution of codenotational terms preserves reference. In particular,

substitution of coexpressional singular terms in a sentence leaves the thought unchanged, i.e., preserves cognitive value; and substitution of codenotational singular terms in a sentence preserves truth value.

On the SR account, then, an i-sentence, $\langle \alpha = \beta \rangle$, receives the following semantic analysis. '=' denotes a relation, being identical with, which holds between $d(\alpha)$ and $d(\beta)$ if and only if $\langle \alpha = \beta \rangle$ is true. Thus, identity relates objects, and, according to the substitution principle for reference, $\langle \alpha = \alpha \rangle$ and true $\langle \alpha = \beta \rangle$ have the same reference, namely, the truth value, True. The cognitive value of $\langle \alpha = \beta \rangle$, however, is to be found in the thought expressed. Thus, α and β each express a sense, and '=' expresses an equivalence relation (a sense-function, as Jackson calls it²¹) holding between the sense of α and the sense of β , viz. being a sense of the same object as. So, although $\langle \alpha = \alpha \rangle$ and true $\langle \alpha = \beta \rangle$ have the same reference, they would express different thoughts if α and β expressed different senses. Thus, in SR Frege is at last able to get the equivalence relation holding between modes of presentation, which he could not manage in Bg.

3. The Sense/Reference Theory

Frege's Sense/Reference theory is well known, and so I shall restrict myself in this section only to its most significant features.

Whereas in Bg, Frege had responded to the Paradox of Identity by rejecting the view that an i-sentence, ' $\alpha=\beta$ ', is about $d(\alpha)$ and $d(\beta)$, in SR Frege rejects the substitution principle which takes the informational or cognitive value of a sentence as invariant under substitution of codenotational singular terms. Instead, Frege now takes only the truth value of the sentence as the invariant property. We shall formulate the new substitution principle provisionally as follows:

(III) If $d(\alpha)=d(\beta)$, then S_α and $S_{\alpha/\beta}$ have the same truth value.

Underlying this substitution principle is a fundamental metaphysical law: if an object x is identical with an object y , then any property of x is equally a property of y , i.e.,

(IV) If $x=y$, then every property of x is equally a property of y .

(III) and (IV) are both commonly referred to in the literature as Leibniz' Law. We shall not follow this practice here, however, because, aside from the questionable attribution to Leibniz, it obscures the fact that (III) and (IV) are distinct principles. (IV) cannot plausibly be denied. For, intuitively, to suppose that x is one and the same thing as y ,

and yet to suppose further, that x has a property y lacks (or conversely), would be, in effect, to suppose that one and the same thing both has and lacks the given property.

(III), on the other hand, is not so firmly grounded. Here is Quine's well known counterexample.

(3) Cicero = Tully,

and

(4) 'Cicero' has six letters,

are both true. Hence, $d('Cicero')=d('Tully')$, so by (III), replacement of 'Cicero' by 'Tully' in (4) ought to result in a true sentence. But

(5) 'Tully' has six letters,

is false; 'Tully' only has five letters.

Quine's case is a counterexample to (III) only, not to (IV). We would have a counterexample here to (IV) only if (4) and (5) both ascribed the same property to the same object, one truly and the other falsely. But, on the most natural interpretation, (4) and (5) ascribe the same property to different objects. For, (3) says that the man Cicero is identical with the man Tully, while (4) says that the name 'Cicero' has six letters, and (5), that the name 'Tully' has six letters. Now, a man is not, in general, identical with his name (or names); and, indeed, although (3) is true,

(6) 'Cicero' = 'Tully',

is false. Were (6) true, i.e., were the name 'Cicero' identical with the name 'Tully', then, according to (IV), any property of the one would have to be a property of the

other. But, (6) is not true, and certainly, (IV) does not require that distinct objects need have all their properties in common. So, since the name 'Cicero' is not the same object as the name 'Tully', there is no conflict with (IV) arising from the fact that 'Cicero' has a property, viz. having six letters, which 'Tully' lacks.

Of course, although we have supplied an interpretation on which Quine's case is seen to be consistent with (IV), we have not thereby shown that there is no interpretation on which it would conflict with (IV); and so, further assurance might be requested. But what reason remains for persisting in this direction? Only, as far as can be determined, the occurrence of the name 'Cicero' in (4) and the name 'Tully' in (5); for one might be inclined to suppose that (4) must therefore be about Cicero and (5) about Tully. The operative assumption here is that if a sentence contains a singular term, then that sentence must be about that which the singular term stands for, either ascribing a property to the object or expressing that it is one of the relata in some n-ary relation; in short, the assumption is,

(V) $S\alpha$ is of $d(\alpha)$.

Assuming (V), (4) would receive the following analysis: the name 'Cicero' denotes the man Cicero, and the remainder of the sentence, viz.

(7) ' ' has six letters,
stands for the property ascribed to Cicero. And
similarly for (5): the property denoted by (7) is ascribed
to the man Tully. So, since (3) is true, we would have
that (4) and (5) both ascribed the same property to the
same object.

Whether we have a counterexample to (IV), however,
is still unclear, for what has yet to be shown is that
(4) and (5) differ in truth value on this interpretation.
We had originally agreed that (4) was true and (5) false,
but that was based on the original interpretation and it
cannot, without argument, be assumed here. Until we are
told what property (7) is alleged to stand for, then, we
must suspend judgment about truth value. But this is a
minor point. The real interest in the case attaches to
(V), because (V) is the link between (III) and (IV): if
(IV) and (V) are both true, then (III) must also be true.
And, since we have established that (III) is false, one
(at least) of (IV) or (V) must therefore be false. Now,
what has been suggested in the last paragraph is that
(V) is true, and thus (IV) is false. But (IV), as we
saw earlier, is among the most fundamental of logical
principles; hence, if we are to reject (IV) in favor of
(V), it would seem that (V) would have to be at least as
basic. Yet it is doubtful that (V) is even true. The
notation for natural language is not so uniform that a

given sequence of letters must always serve the same function in every context in which it occurs--indeed, that the different functions served by a given sequence of letters need even be related one to the other. Thus, the sequence of letters which forms an initial segment of the term 'cicerone' serves in (4), e.g., to stand for the Roman; but that is accidental to its occurrence in 'cicerone'. Similarly, the name 'Cicero' surely occurs in (4), but it does not stand for the man in that context; rather, it exhibits itself, as the single quote marks indicate, and (4) is thus about the name, not the man. The only reasonable course, then, is to reject (V) and the purported counterexample to (IV).

We see from this example that a singular term occurring in a sentence need not be serving in that sentence simply to stand for what it ordinarily denotes; and in such a case we ought not expect that the sentence is about that which the singular term ordinarily denotes. And, if $S\alpha$ is not about $d(\alpha)$, then we cannot expect the truth value of $S\alpha$ to depend simply on $d(\alpha)$; and so we cannot expect $S\alpha/\beta$ to have the same truth value as $S\alpha$ even though $d(\alpha)=d(\beta)$. This analysis of the purported counterexample to (IV) does, however, have a constructive side. For, in exposing the assumption (V), the nature of the connection between (III) and (IV) has been revealed, and thus, the needed correction for (III) becomes

apparent. If $S\alpha$ is about $d(\alpha)$ then it would seem that $S\alpha$ and $S\alpha/\beta$ must agree in truth value whenever $d(\alpha)=d(\beta)$; as Quine remarks, "whatever can be affirmed about [an] object remains true when we refer to the object by any other name."²² So, we replace (III) by

(VI) If $S\alpha$ is about $d(\alpha)$, then

if $d(\alpha)=d(\beta)$, then $S\alpha$ and $S\alpha/\beta$ have the same truth value.

(Actually, (VI) is still not quite right; it does not handle, for example, (again from Quine) "Giorgione was so-called because of his size." But the principle behind the repair has already been explained, and since the refinements would take us too far afield there is no need to pursue the matter further.)

This analysis of quotations is derived from Frege's own analysis of what have come to be called "opaque contexts." I shall return to these later. The point I should like to emphasize here is the connection Frege has drawn between a singular term denoting a given item and a sentence containing that term being about that item, for this is essentially the role denotation plays in Frege's theory: the denotation of a singular term is that which we use the term to speak about. The paradigm of reference, as Dummett correctly points out,²³ is the name/bearer relation, and in the case of proper names, the two coincide: the referent of 'Aristotle' is just the bearer of that name, Aristotle, the famous Greek.

philosopher. However, a term need not be assigned to an object, in the way a proper name is assigned to its bearer, in order that it refer to the object. A definite description, like 'the teacher of Alexander the Great', has a referent, namely, again, Aristotle; the description is not assigned to him, but he is the referent because he is the one uniquely satisfying the description. Moreover, just as in

(8) Aristotle invented symbolic logic,

we use the name 'Aristotle' to speak about Aristotle, so in

(9) The teacher of Alexander the Great invented symbolic logic,

we use the description 'the teacher of Alexander the Great' to stand for that which the sentence is about, namely, again, Aristotle.

This part of Frege's notion of reference seems to me to be unquestionably correct: the substitution principle, wherein the connection is made between the denotation of a singular term and the truth value of a sentence containing it, and the recognition that the denotation of a singular term is that which we use the term to speak about. As is well known, there was a great deal more to Frege's notion of reference: in particular, he also supposed predicate expressions to refer--the predicate 'horse', e.g., denoted the concept horse--and he even supposed declarative sentences to refer, namely,

to truth values. These other features of Frege's notion of reference, however, are extremely problematic: it is highly doubtful whether they are true, or can even be made coherent. But we need not get into these issues; what I have presented so far of Frege's notion of reference is all that is essential in that notion for his solution to the paradox.

So much for the notion of reference. Let us now turn to the notion of sense.

The referent of a singular term, as we have seen, is the object for which it stands, either by having been assigned to that object or by uniquely describing it. Each of 'the Stagirite', 'Aristotle', 'Plato's greatest pupil', and 'the teacher of Alexander the Great' refer to the same object, namely, Aristotle. But they do not all have the same sense. The sense of a singular term is that "wherein the mode of presentation is contained," and it thus carries the burden of introducing, presenting, or picking out the referent. 'The Stagirite' for example, picks out Aristotle as having been Stagira's most famous native son, while 'Plato's greatest pupil' picks him out as having studied with Plato and as having been his finest student. Both of these terms refer to Aristotle, but since they pick Aristotle out in different ways, they do not have the same sense. To take another example, compare 'the number which is obtained by adding 2 four times' with

'the number which is obtained by adding 4 twice'. Both terms have the same referent, the number 8, but the first instructs us to take 2 four times, i.e.,

X	X
X	X
X	X
X	X

while the second instructs us to take 4 twice, i.e.,

X	X	X	X
X	X	X	X

and so they do not have the same sense.

The sense of an expression is that which is communicated or conveyed by the expression, the information it contains. The sense of an expression is not material, nor is it perceptible, but it is an objective entity nonetheless which exists independent of any individual's consciousness. "For one can hardly deny that mankind has a common store of thoughts which is transmitted from one generation to another";²⁴ but "if every thought requires a bearer, to the contents of whose consciousness it belongs, then it would be a thought of this bearer only and there would be no science common to many, on which many could work."²⁵ Frege is here especially anxious to distinguish senses, which are objective, from ideas [Vorstellungen], which are private to each individual:

[A] man never has somebody else's mental image, but only his own; and nobody even knows how far his image (say) of red agrees with somebody else's; for the peculiar character of the image I connect with the word 'red' is something that I cannot convey.²⁶

When a person grasps [Fassen] a sense, "something in his consciousness must be aimed at [it]";²⁷ nevertheless, "one does not create [the sense] but only comes in a certain relation, which is different from seeing a thing or having an idea, to what already existed before hand."²⁸ Senses, then, belong neither in the outer world of material entities nor in the many private inner worlds, but in a specially designated third realm.

The sense/reference distinction is partly an ontological distinction, but it is also a distinction between the ways in which entities are related to signs: an entity might be referred to by a sign, or it might be expressed by the sign, or it might be associated with the sign. "Idea" and "sense" serve to label particular types of entities, and it appears that ideas are the only sort of things that can be associated with an expression and that senses are the only sorts of things that can be expressed by an expression. "Reference," however, carries no such implications: the referent of an expression is simply that which we use the expression to talk about, and insofar as we talk about ideas and senses as well as anything else, we use expressions that refer to them. A's idea of the moon is not the referent of the expression 'the moon', but presumably that which A associates with the expression; on the other hand, A's idea of the moon is the referent of the expression 'A's

idea of the moon', and with this expression, yet another idea might be associated. Again, the sense of the expression 'the morning star' is the sense of 'the morning star'; but it is the referent of 'the sense of 'the morning star'', and with this expression yet another sense would be expressed. However, although Frege does not explicitly say so, it is reasonable to suppose that the sense, the referent, and the associated idea of a given expression must be distinct; hence, no expression can both refer to and express one and the same entity.²⁹

A sequence of noises or of marks on paper, if it is to be a word or phrase--what Frege calls a "sign"--must have a sense; but it need not follow that any reference corresponds to it. For example, 'the celestial body most distant from the Earth' and 'the least rapidly converging series', Frege says, both have a definite sense, but it is doubtful whether the first has any reference, and it is demonstrable that the second lacks one. Many of the names found in fiction and myth fall here: 'Medusa', 'Odysseus', 'Santa Claus'. Thus Frege answers the question that vexed Russell: How is it that a name that refers to nothing at all can still be a meaningful sign? Nondesignating singular terms have a sense but no reference: corresponding to such an expression is a criterion for recognizing whether a given object is the referent, and although there is no object

satisfying the conditions laid down, the "way of recognizing" gives the term the stability in discourse necessary for communication.

As we mentioned earlier, Frege's distinction between sense and reference is intended to disambiguate the Bg notion of content; and just as in Bg Frege held both sentences and parts of sentences to have content, so in SR he extends the sense/reference distinction to sentences as well as to singular terms. In Bg, recall, Frege construed a sentence as a complex name constituted of a function-expression and argument-expressions in such a manner that the contents of the parts of the sentence were parts of the content of the whole sentence. The same idea is at work in SR, except that Frege now takes the reference of the sentence to be composed out of the reference of the parts of the sentence; and he takes the sense of the sentence to be composed out of the senses of the parts of the sentence. By distinguishing sense from reference, Frege now avoids the awkward feature of the Bg theory that had gotten him into trouble, namely, the fact that objects themselves were parts of contents. For, now, if a sentence contains a singular term, then the reference of that singular term, i.e., the object for which it stands, is part of the reference of the sentence. It is not a part of the thought expressed by the sentence, because the thought, which Frege identifies with the sense of the sentence, contains the sense of the

singular term as a part, not its reference.

But it is unfortunate that Frege should express the contribution of the sense of a part of a sentence to the sense of the whole sentence in terms of parts and wholes of senses. We must distinguish between saying, on the one hand, that the reference (sense) of a complex name is a function of the reference (sense) of the parts of the name, and saying, on the other hand, that the reference (sense) of a complex name contains the reference (sense) of the parts of the name. Frege noted that we must be careful about transferring our talk of parts and wholes with regard to linguistic expressions into the realm of sense and reference; and eventually he had to abandon the claim that the reference of a part of a complex name is part of the reference of the complex name. For, 'the Queen of England' contains the name 'England' as a proper part, but the Queen of England does not contain the country, England, as a proper part. On the other hand, it is interesting that Frege never dropped the part/whole metaphor for senses.³⁰

The sense of a declarative sentence is the thought it expresses, and the reference of a declarative sentence is its truth value; and, as was the case with singular terms, a sentence might have a sense but lack a reference. A sentence is, for Frege, a complex name, and the relation between the reference of a sentence and the reference of the

parts of the sentence is given by the substitution principle for reference. But what about the relation between the sense of a sentence and the sense of the parts of the sentence? That we must recognize the sense of the sentence, the thought it expresses, as consisting of parts, is given in this very famous passage from "Compound Thoughts":

It is astonishing what language can do. With a few syllables it can express an incalculable number of thoughts, so that even a thought grasped by a human being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new. This would be impossible, were we not able to distinguish parts in the thought corresponding to the parts of a sentence, so that the structure of the sentence serves as an image of the structure of the thought. . . . If, then, we look upon thoughts as composed of simple parts, and take these, in turn, to correspond to the simple parts of sentences, we can understand how a few parts of sentences can go to make up a great multitude of sentences, to which, in turn, there correspond a great multitude of thoughts.³¹

And, in SR, he notes:

If it were a question only of the sense of the sentence, the thought, it would be unnecessary to bother with the reference of a part of the sentence; only the sense, not the reference, of the part is relevant to the sense of the whole sentence.³²

This suggests a substitution principle for sense analogous to that for reference. If the sense, and only the sense of a given name contributes to the sense of a sentence containing that name, then were we to replace that name by any other having the very same sense, the resultant

sentence ought to express the same thought as the original. Frege does not actually state such a principle in his published writings, but he often comes close, as, for example, in this passage from Grundgesetze:

The names, whether simple or themselves composite, of which the name of a truth-value consists, contribute to the expression of the thought, and this contribution of the individual [component] is its sense. If a name is part of the name of a truth-value, then the sense of the former name is part of the thought expressed by the latter name.³³

We thus have a substitution principle for sense comparable to the substitution principle for reference discussed earlier. Where η is an expression, let $\langle s(\eta) \rangle$ abbreviate 'the sense of η '. Then, we have as our substitution principle for sense:

- (VII) If $S\alpha$ is about $d(\alpha)$, then
 if $s(\alpha) = s(\beta)$, then $S\alpha$ and $S\alpha/\beta$ have the same
 cognitive value.

Finally, I should like to mention Frege's treatment of what are called "opaque contexts," i.e., contexts in which substitution of codenotational singular terms fails to preserve truth value. From,

(1) the evening star = the morning star,
 and

(10) John knows that the evening star is a body
 illuminated by the sun,

we cannot validly conclude

- (11) John knows that the morning star is a body
illuminated by the sun.

Far from being a counterexample to the substitution principle (V), however, Frege regards the invalidity of this inference as evidence that (10) is not about the evening star, and so, that 'the evening star' is not serving in (9) to stand for the evening star. In so far as it is serving to refer to something in (10), however, it will, as we mentioned earlier in discussing the notion of reference, denote that which one intends to speak about in (10). And Frege says that what we are speaking about in this context is the sense of the expression. Thus Frege distinguishes between the customary reference of an expression and its indirect reference. In (10), 'the evening star' does not have its customary reference, the evening star, but its indirect reference; and since what one is using the term in that context to speak about is the sense of the expression, that is the indirect reference. Similarly, Frege distinguishes between the customary sense of an expression and its indirect sense.

Frege's treatment of these cases is controversial.³⁴ What is important and correct however, is his recognition that they do not stand as counterexamples to the principle (III), but rather show that the terms in those contexts do not serve purely to stand for what they

ordinarily denote.

4. Frege and his Critics

Michael Dummett claims that "there is no alternative theory of proper names that can be opposed to Frege's."³⁵ In simplified form, according to Dummett, Frege's theory is that "a proper name possesses a unique and specific sense, common to all users of the name, which determines its reference."³⁶ The proper name, "if it is to be considered as having a determinate sense, must have associated with it a specific criterion for recognizing a given object as the referent of the name; the referent of the name, if any, is whatever object satisfies this criterion."³⁷ This simplified theory, Dummett claims, "requires qualification in many respects before it becomes a realistic picture of our actual employment of names."³⁸ One such qualification Dummett would place on this simplified theory is Wittgenstein's replacement of a "single sharp criterion for identifying the referent by a cluster of alternative ones."³⁹

For Wittgenstein, the sense of a proper name is given, not by a single specific criterion of identification, but by a cluster of such criteria: for an object to be the referent of the name, it is not necessary that it satisfy all these criteria, but only that it satisfy most of them, or, perhaps, merely a suitable number of them. Thus, for Wittgenstein, the sense of a proper name overdetermines its reference, and is, at the same time, elastic, in that we are prepared in advance to drop some of our criteria of identification if

they are discovered not to converge with the others.⁴⁰

But Dummett urges that the complexities thus introduced in order to provide a more realistic account of our actual use of proper names ought not becloud the essential correctness of Frege's theory:

What must be resisted is the temptation to think that the need for these multiple qualifications of the simple account in order to do justice to the complexity of actual linguistic practice destroys the utility of the original model, whereby a name has reference in virtue of its sense; for that model displays the only mechanism by which a name could acquire reference, even though the actual working has been simplified for the sake of perspicuity. It is not a choice between Frege's theory and some alternative theory: there is no other theory.⁴¹

Dummett concedes that his remarks might "appear preposterous"; he does not, however, admit that they are preposterous. And they most certainly are: unless 'Frege's Theory' is an utter platitude--in which case it is not a theory at all--it is implausible to claim that there can be no alternative theory of proper names. It is not clear to me what Dummett is actually denying here, but it certainly seems to be the case that a number of other theories have been put forward which are at least thought to be alternative theories. Indeed, Dummett has been prompted to defend Frege in this very forceful manner by the attack Saul Kripke has launched in "Naming and Necessity" against what he takes to be Frege's account of proper names. Kripke is very much alive to the power of

Frege's account of proper names: "it is hard," he says, "to see . . . how the Frege-Russell view, or some suitable variant, can fail to be the case."⁴² "Nevertheless," Kripke claims, "I think it pretty certain that the view of Frege and Russell is wrong."⁴³ While it is true that Kripke has not offered an alternative theory to Frege's-- he believes that all philosophical theories are false-- and so, to this extent, agrees with Dummett that there can be no alternative theory to Frege's, he does claim Frege's account to be wrong, and he does offer in its place, not a theory, but, as he says, "a better picture" of the way in which proper names are used.

Dummett and Kripke are proponents of the two basic competing theories of proper names. Dummett, on the one hand, defends a Frege-type theory, on which the significance of a proper name is not completely exhausted by the name's standing for a given object; Kripke, on the other hand, is a proponent of a Mill-type theory, on which names are simply labels or tags for the things to which they are assigned. The Kripke-Dummett interchange affords us an opportunity in this section to discuss the two views, and to present the strengths and the weakness of each.

I will begin with Mill.

Mill's semantic theory is a theory of what he calls 'names', those expressions which the schoolmen termed 'categorematic', viz., "[a] word which could be used either

as the subject or predicate of a proposition without being accompanied by any other word."⁴⁴ Noun phrases (i.e., many-worded names) and adjectives are names; adverbs, prepositions, and inflected noun phrases in accusative and dative cases, on the other hand, are syncategorematic expressions. Of names, Mill draws three interlocking distinctions: (1) between singular and general terms, (2) between abstract and concrete terms, and (3) between connotative and nonconnotative terms. The singular/general distinction goes as follows:

A general name is, familiarly defined, a name which is capable of being truly affirmed, in the same sense, of each of an indefinite number of things. An individual or singular name is a name which is only capable of being truly affirmed, in the same sense, of one thing.⁴⁵

'Man', is a general name, since it is capable of being truly affirmed of any number of individuals, and it does so, in the same sense each time. 'John', on the other hand, is a singular term, for "though there are many persons who bear that name, it is not conferred upon them to indicate any qualities or any thing which belongs to them in common, and cannot be said to be affirmed of them in any sense at all, consequently not in the same sense."⁴⁶ Definite descriptions like 'the king who succeeded William the Conqueror' count also as singular terms, it being "implied in the meaning of the words" that there cannot be more than one. Expressions like 'the king' and 'this

stone', though capable of being applied to many individuals in the same sense, also count as singular terms on Mill's view since, in any given context of use, only one such individual is the subject of which the predicate is affirmed. A general term is the name of a class, but it is predicated of each of the members of the class; a singular term, on the other hand, is predicated of the individual it names.

"Virtuous", Mill says, applies to all the virtuous things "in consequence of an attribute which they are supposed to possess in common, the attribute which has received the name of virtue."⁴⁷ The crucial feature of general terms like 'virtuous' and 'man' is that they are truly applied to each of a number of items in virtue of the fact that these items all possess an attribute or attributes in common. This Mill calls the connotation of the term, and the term is said to connote the attribute or attributes. The denotation, on the other hand, is the thing or things to which the term truly applies; the term is said to denote each of these items to which it truly applies. The terminology, then, goes as follows:

The name, therefore, is said to signify the subjects directly, the attributes indirectly; it denotes the subjects, and implies, or involves, or indicates, or, as we shall say henceforth, connotes, the attributes.⁴⁸

Connotation and denotation are not, so to speak, two parts of the meaning of an expression; rather, Mill identifies

the meaning of a term with its connotation:

. . . whenever the names given to objects convey any information--that is, whenever they have properly any meaning--the meaning resides not in what they denote but in what they connote.⁴⁹

And, as has already been indicated, Mill holds that some names lack connotation, i.e., meaning, altogether. This leads us into "one of the most important distinctions which we shall have occasion to point out and one of those which go deepest into the nature of language,"⁵⁰ namely, the distinction between connotative and non-connotative terms:

A nonconnotative term is one which signifies a subject only, or an attribute only. A connotative term is one which denotes a subject and implies an attribute.⁵¹

The paradigm of nonconnotative term is a proper name:

Proper names are not connotative; they denote the individuals who are called by them, but they do not indicate or imply any attributes as belonging to those individuals. When we name a child by the name Paul or a dog by the name Caesar, these names are simply marks used to enable those individuals to be made subjects of discourse. It may be said, indeed, that we must have had some reason for giving them those names rather than any others, and this is true, but the name, once given, is independent of the reason. . . . Proper names are attached to the objects themselves and are not dependent on the continuance of any attribute of the object.⁵²

A proper name, then, simply stands for an object because the object is called by that name, and not because the object happens to possess certain attributes by means of

which we determine that the name truly applies to it. General terms, on the other hand, are paradigms of non-connotative terms. The term 'man', for example, truly applies to Aron Nimzovich because he happens to possess those attributes in virtue of which we call something a man, namely, corporeity, animal life, rationality, and human form. However, the connotative/nonconnotative distinction does not coincide with the singular/general distinction, because definite descriptions, which are singular terms, are connotative. They "may be significant of some attribute or some union of attributes which, being possessed by no object but one, determines the name exclusively to that individual."⁵³ Unlike a proper name which simply stands for an object, a definite description denotes an object because that object happens to fit the description uniquely.

Finally, Mill distinguishes between terms like 'white' and 'virtuous', on the one hand, and 'whiteness' and 'virtue', on the other. The former he calls concrete, the latter abstract, according to the formula:

A concrete name is a name which stands for a thing; an abstract name is a name which stands for an attribute of a thing.⁵⁴

'Virtue' is an abstract term because it denotes the attribute, not the things which possess the attribute, those, that is, the virtuous things, being denoted by the concrete general term 'virtuous'. 'Virtue' is also a

singular term; it denotes but a single attribute, and it is nonconnotative--it is simply a mark for the attribute and does not imply any attributes which an attribute must possess in order to be called 'virtue'. Hence, 'virtue' is exactly like the proper name 'John', except that 'virtue' is abstract while 'John' is concrete.

If a term is connotative, then the connotation of the term enables us to determine whether a given item is truly called by that term. All we need do is check to see whether the item in fact possesses the attributes which the term connotes. But, if a term is nonconnotative, there remains a question about how we determine whether a given item is properly called by it. Although it does not appear in the usual accounts of Mill's theory, Mill actually did anticipate this question, and he provided the following answer:

When we impose a proper name, . . . we put a mark, not, indeed, upon the object itself but, so to speak, upon the idea of the object. A proper name is but an unmeaning mark which we connect in our minds with the idea of the object, in order that, whenever the mark meets our eyes or occurs to our thoughts, we may think of that individual object. Not being attached to the thing itself, it does not . . . enable us to distinguish the object when we see it, but it enables us to distinguish it when it is spoken of, either in the records of our own experience or in the discourse of others, to know that what we find asserted in any proposition of which it is the subject is asserted of the individual thing with which we were previously acquainted.⁵⁵

However, it is clear why Mill's answer has been by and large ignored: this little story about our putting a mark next to our idea of a given object is picturesque, but one cannot take it in any literal sense. Moreover, if this idea of a given object is something that more than one person can possess, then, what he is associating the name with is something objective, a concept, and thus something very much like Frege's sense.

To sum up, now, let us see how Mill fares on the issues with which we are specifically interested. On the one hand, Mill does seem to be correct in contrasting proper names with definite descriptions: we can say that a definite description stands for a given object because the description is true of the object (uniquely), but we cannot say that a proper name stands for a given object because it is true of the object--it has been assigned to the object.⁵⁶ On the other hand, Mill fails to give an adequate answer to the question of how we determine the reference of a nonconnotative term; and also, since he identifies information with meaning, and meaning with connotation, then if proper names lack connotation, Mill provides no explanation for the difference in cognitive value between ' $\alpha=\alpha$ ' and true ' $\alpha=\beta$ ', when α and β are proper names--indeed, it is difficult to see how, on his view, they can even differ in cognitive value.

Russell is a bridge between Mill and Frege: on the one hand, he operates with the Mill assumption that a proper name--a genuine proper name--has denotation but no connotation; on the other hand, he believes most ordinary proper names not to be genuine proper names at all, but rather disguised descriptions, and thus as having meaning.

In a detective story propositions about "the man who did the deed" are accumulated, in the hope that ultimately they will suffice to demonstrate that it was A who did the deed. We may even go so far as to say that, in all such knowledge as can be expressed in words--with the exception of "this" and "that" and a few other words of which the meaning varies on different occasions--no names, in the strict sense, occur, but what seem like names are really descriptions. We may inquire significantly whether Homer existed, which we could not do if "Homer" were a name. The proposition "the so-and-so exists" is significant, whether true or false; but if a is the so-and-so (where "a" is a name), the words "a exists" are meaningless. It is only of descriptions--definite or indefinite--that existence can be significantly asserted; for, if "a" is a name, it must name something: what does not name anything is not a name, and therefore, if intended to be a name, is a symbol devoid of meaning, whereas a description, like "the present King of France," does not become incapable of occurring significantly merely on the ground that it describes nothing, the reason being that it is a complex symbol, of which the meaning is derived from that of its constituent symbols. And so, when we ask whether Homer existed, we are using the word "Homer" as an abbreviated description: we may replace it by (say) "the author of the Iliad and the Odyssey." The same considerations apply to almost all uses of what look like proper names.⁵⁷

As is evident in this passage, Russell had been seeking to explain how a name which lacked denotation could be meaningful, and so how any sentence containing such a

name could also be meaningful. In his early work, The Principles of Mathematics,⁵⁸ Russell had solved this problem simply by enlarging Reality: Pegasus, although not an existent, was or had being, and so there was something answering to the name 'Pegasus'. However, in "On Denoting,"⁵⁹ he instead denied that what appeared on the surface to be names were names at all. Most ordinary proper names were, on this view, not genuine or logically proper names, but rather disguised or truncated definite descriptions; and he provided an analysis of sentences containing definite descriptions on which they were seen to "lack meaning in isolation." A sentence of the form 'the F is G', where 'the F' is a definite description, would not be given the logical form $(G\alpha)$; instead it would be taken to be equivalent to "there is at least one thing that is F, and there is no more than one thing that is F, and whatever is F is G"--in symbols $(\exists x)((y)(Fy \supset x=y).Gx)$ --in which there is no singular term standing for the F.

The details of Russell's theory of descriptions are by and large well known, but there is one point I would like to mention briefly, since it will play an important role later, and that is his scope distinction for descriptions.

When a sentence containing a definite description occurs as part of a larger construction, an ambiguity may arise, depending upon whether, as Russell put it, the

description has primary occurrence or secondary occurrence.

For example, the sentence,

(12) The present King of France is not bald,
might be understood in two ways. It might be taken as,

(13) It is not the case that the present King of France
is bald,

i.e., as the denial of the statement that the present King
of France is bald; or, it might be understood as,

(14) The present King of France is such that he is not
bald.

Whereas in (13) one is denying a statement, namely,

(15) The present King of France is bald,

in (14), one is affirming (so Russell would put it) that
one and only one thing is presently King of France, while
denying that that thing is bald. If we let ' $F\xi$ '

abbreviate ' ξ is presently King of France', and if we let
' $B\xi$ ' abbreviate ' ξ is bald', the difference in the
symbolization would be as follows. (13) would be
symbolized as,

(16) $\neg(\exists x)((y)(Fx \equiv y=y) \cdot Bx)$,

while (14) would be symbolized as

(17) $(\exists x)((y)(Fx \equiv x=y) \cdot \neg Bx)$.

There being nothing which is presently King of France,

(16) is trivially true and (17) trivially false.

Russell says that in (16) the description has
secondary occurrence while in (17), the description has

primary occurrence

. . . writing ' $(\exists x)(\phi x)$ ' for 'the term x which satisfies ϕx ', $\psi(\exists x)(\phi x)$ is to mean $(\exists b): \phi x. \exists_x. x=b: \psi b$. This, however, is not yet quite adequate as a definition, for when $(\exists x)(\phi x)$ occurs in a proposition which is part of a larger proposition, there is doubt whether the smaller or the larger proposition is to be taken as the " $\psi(\exists x)(\phi x)$." . . . In order to avoid ambiguities as to scope, we shall indicate the scope by writing " $[(\exists x)(\phi x)]$ " at the beginning of the scope, followed by enough dots to extend to the end of the scope. . . . Thus we arrive at the following definition:

*14.01 $[(\exists x)(\phi x)]. \psi(\exists x)(\phi x). = :(\exists b): \phi x. \exists_x. x=b: \psi b$ Df⁶⁰

Using Russell's notation, (16) would be expressed as

(18) $\neg[(\exists x)(Fx)]. B(\exists x)(Fx)$,

and (17) would be expressed as

(19) $[(\exists x)(Fx)]. \neg B(\exists x)(Fx)$.

Whenever the F exists, however, i.e., whenever

$(\exists y)((x)Fx \equiv x=y)$, " $(\exists x)(\phi x)$ behaves, formally, like an

ordinary argument to any function in which it may occur":⁶¹

. . . provided $(\exists x)(\phi x)$ exists, it has (speaking formally) all the logical properties of symbols which directly represent objects. Hence when $(\exists x)(\phi x)$ exists, the fact that it is an incomplete symbol becomes irrelevant to the truth-values of logical propositions in which it occurs.⁶²

Restricting ourselves, then, to "logical propositions,"

i.e., to propositions which are either atomic or have been

constructed only by truth-functional or extensional

functions, we have

. . . when $\exists!(\exists x)(\phi x)$, the scope of $(\exists x)(\phi x)$ does not matter to the truth-value of any proposition in which $(\exists x)(\phi x)$ occurs.⁶³

That is, provided that F exists, and also that 'G . . .' is a context of the requisite kind,

$$(20) \quad [(1x)Fx]G(1x)Fx \equiv G[(1x)Fx] (1x)Fx.$$

Note, however, that this is not to say that when the F exists, 'the F is G' will express the very same proposition as 'Ga', supposing 'a' to be a proper name for the F. For, on Russell's view, the object itself enters into the proposition expressed by 'Ga'; the proposition expressed by 'the F is G' will not contain the object, however, but the denoting complex. And, although Russell does not say so, I am inclined to believe that when the F exists, even though (20) is true, it need not be the case that '[(1x)Fx] G (1x)Fx' expresses the same proposition as 'G [(1x)Fx] (1x)Fx'.

Russell had presented his theory of descriptions explicitly as an alternative to Frege's sense/reference treatment of singular terms which, as I mentioned earlier, Russell believed to be logically incoherent. Despite this, and despite the fact that Russell was working within the semantic framework of Mill--genuine proper names were, for Russell, simply labels or tags--Russell might actually be viewed as an ally of Frege, elucidating Frege's view that ordinary proper names have sense: to say that a proper name has a sense is to say that a proper name is an abbreviation for a definite description which uniquely specifies the object. The name stands for the object

because that object has the property specified by the description. Kripke believes this to be so:

Frege and Russell both thought, and seemed to arrive at these conclusions independently of each other, that Mill was wrong in a very strong sense: really a proper name, properly used, simply was a definite description abbreviated or disguised. Frege specifically said that such a description gave the sense of the name.⁶⁴

Dummett, on the other hand, has denied this identification of the two views.

[Kripke] attributes to Frege an express declaration that the sense of a proper name is always the same as that of some one definite description, though noting that Frege allowed that a proper name in common use in natural language may have many different such senses associated with it by different speakers. He proceeds to launch an attack on this theory. In fact, Frege made no explicit statement to this effect, and it is extremely dubious that he supposed such a thing. It is true that, in giving examples of possible senses that may be associated with a proper name, Frege expresses these by means of definite descriptions; but this should be considered as merely a device for a brief characterization of a sense, rather than as a means of conveying the thesis which Kripke ascribes to Frege. What is important about Frege's theory is that a proper name, if it is to be considered as having a determinate sense, must have associated with it a specific criterion for recognizing a given object as the referent of the name; the referent of the name, if any, is whatever object satisfies the criterion. Sometimes the criterion may be capable of being conveyed by means of a definite description, in other cases not.⁶⁵

Now, it is not clear to me what Dummett's disagreement with Kripke comes to: in particular, it is not clear how a "criterion for recognizing an object" differs from a description that uniquely specifies it.

If, for example, the criterion takes the form 'the unique object satisfying conditions C_1, \dots, C_n ', then this is just so far as I can tell, to make the sense of a proper name the same as some definite description. It might be that Dummett regards a criterion as a feature of the object itself, in which case the disagreement would be that whereas Dummett takes a criterion to be associated with a name, Kripke takes the specification of a criterion to be associated with the name. It is not clear to me that this is what Dummett has in mind, however, for he does, for example, consider the description theory to be narrower than Frege's "sense theory," as he calls it, and to this extent, on the same level with it. In any event, it would certainly not be Frege's view, for to take the associated criterion as a feature of the object would make of the sense of a name a perceptible, and so, nonconceptual item, contrary to Frege's intention. Is the issue, then what sort of description a name might abbreviate? Kripke, when attacking the Frege-Russell description theory of proper names says:

It seems to me to be wrong to think that we give ourselves some properties which somehow qualitatively uniquely pick out an object and determine our reference in that manner.⁶⁶

And, this theme, namely that the Frege-Russell view requires that we have qualitative properties by means of which we pick out objects, recurs throughout his lectures.

Dummett clearly rejects this reading of Frege:

No one supposes that the reference of a proper name is determined solely by qualitative features of the individual either at the time at which the name is used or at the time that is being spoken of, that is, by features whose presence or absence can be determined by examination of the individual at the relevant time. If, to take a crude example, the reference of 'Manhattan Island' is fixed by 'the long narrow island off the Atlantic coast of the United States with all those skyscrapers, just by the Statue of Liberty,' that does not preclude the use of the name at, or of, a time when the island did not yet have, or no longer has, all those features. That is precisely why it is essential to the understanding of a proper name that one knows the criterion of identity associated with it; the sense of a personal proper name requires that its bearer be identified as the same man as the one who, at an appropriate period, fitted some description (if that is how the reference of the name has been fixed). For this reason, if the description by which we attempt to fix the reference of a name is later found not to have a unique application, we do not insist that anyone fitting the description is a referent of the name, but feel bound to change the sense of the name.⁶⁷

But this 'criterion of identity' associated with a name, on Dummett's view, is clearly a description. If the sense of a proper name for a person is to be of the form 'the same man who satisfies C_1, \dots, C_n at T_m ', then the sense of the proper name will just be the same as that of some description.

Now, I find the positions becoming a bit confusing here, and although I cannot be sure, I am inclined to believe that one of the central issues separating Kripke and Dummett is right here, with the kind of description we use to determine reference. For Dummett seems, each

time he explains his "criterion of identification," to give, as the sense of a proper name some definite description. And Kripke acknowledges in a footnote that this is what his dispute with the description theorist might eventually reduce to.⁶⁸

The main textual support for believing Frege to have held that the sense of a proper name is the same as that of some definite description is the following famous passage from SR:

In the case of an actual proper name such as 'Aristotle' opinions as to the sense may differ. It might, for instance, be taken to be the following: the pupil of Plato and teacher of Alexander the Great. Anybody who does this will attach another sense to the sentence 'Aristotle was born in Stagira' than will a man who takes as the sense of the name: the teacher of Alexander the Great who was born in Stagira. So long as the reference remains the same, such variations of sense may be tolerated, although they are to be avoided in the theoretical structure of a demonstrative science and ought not to occur in a perfect language.⁶⁹

This surely provides considerable support for Kripke's claim that Frege had "specifically said that such a description gave the sense of the name," and I am unaware of any other passage in which Frege had indicated otherwise. In his examples, Frege almost invariably gives, as the sense of a proper name, some definite description. Moreover, Russell's theory of descriptions seems to give us just what Frege wanted in ascribing sense to proper names, namely, a way of determining the reference of the name and the machinery for accommodating

the fact that different names for the same item might be of significance in inference. Dummett's response to Kripke, namely, that Frege's examples were meant to be a "brief characterization" of the sense of a name and no more, is certainly a reasonable response; but, when Dummett himself attempts to fill out this brief characterization, then, as we have seen, he ends up with the sense just being that of a definite description. So, taking into consideration both Frege's examples and Dummett's inability to provide an account on which the sense of a proper name is not a definite description, it would appear that on Frege's view, the sense of a proper name just is that of some definite description.

Happily, however, we need not decide the issue here. Kripke, as we have seen, certainly believes Frege and Russell to have held roughly the same view about proper names; and Dummett, in his reply, sees the substance of his dispute with Kripke as being in no way affected by this interpretation. Since the argument is cast by both as though Frege had held proper names to be disguised descriptions, we shall continue as if this were so.

Kripke's attack on Frege's theory has two prongs: first, he charges Frege with having been confused about the notion of sense for proper names; and second, he argues that, whichever way one takes the notion of sense,

Frege is wrong in supposing that sense determines reference. Dummett's response also has two parts: first, he argues that Frege was not confused about the notion of sense, although Kripke is, and second, that sense does determine reference.

. . . we may simply say that Kripke agrees with Frege that a proper name is first introduced into the language by associating with it a criterion of identification, but that he differs from him in holding that such a criterion serves merely to fix the reference of the name and not to give its meaning. In fact, Kripke accuses Frege precisely of confusing these two things: he says . . . that Frege uses 'sense' in two senses, both for the way the reference of a term is determined and for its meaning. We shall see later what Kripke understands by 'meaning.'⁷⁰

What we shall see later is that Dummett has completely misunderstood Kripke's notion of meaning. But the notion of meaning does seem to be one of the central issues separating the two. Kripke, on the one hand, says:

Frege should be criticized for using the term 'sense' in two senses. For he takes the sense of a designator to be its meaning; and he also takes it to be the way its reference is determined. Identifying the two, he supposes that both are given by definite descriptions. Ultimately, I will reject this second supposition too; but even were it right, I reject the first. A description may be used as synonymous with a designator, or it may be used to fix its reference. The two Fregean senses of 'sense' correspond to two senses of 'definition' in ordinary parlance. They should carefully be distinguished.⁷¹

Dummett, on the other hand, replies: "Kripke's meaning is not a credible representation of the knowledge that someone has when he understands the expression."⁷² And

again:

[Kripke's account] cannot be an account of what endows a proper name with reference, as Frege understands the term 'reference.' For reference, as it figures in Frege's theory of meaning, has an essential connection with sense: the sense--which is what a speaker knows when he understands the word--must be capable of being exhibited as a means of determining the reference. The notion of reference ought not, that is, to be idle within the theory of meaning. When someone knows the sense of a sentence, what he knows is how the truth-value of the sentence is to be recognized, whenever we are in a position to do so; and if a reference is to be significantly attributed to a word, then, for at least some sentences containing the word, the account of the process of recognizing that such sentences have one or other truth-value must involve the recognition of something as the referent of that word. If this is not so, then, however clearly we may be able to explain the attribution of something which we choose to call 'reference' to that word, the notion of reference we are employing has no role to play within the theory of meaning, as far as it relates to that word. Kripke's account, however, does not describe anything which could be involved in anyone's recognition of an object as the referent of a proper name: hence, though it may succeed in stipulating a sense in which we might, for some purpose or other, choose to say that, by using the name, someone had referred to the object, the sense of 'reference' so stipulated can have no part in any theory of meaning, that is, in a theory of what the use of proper names consists in so far as a mastery of a language requires a mastery of that use.⁷³

For the remainder of this section I should like to look carefully at Dummett's defense of what he calls Frege's notion of meaning.

I find Dummett's characterization of Frege's view to be very much more like something someone under the influence of the later Wittgenstein would hold than

anything Frege himself held. In particular, I find Dummett's expression "mastery of the use of proper names" to be utterly mystifying. Dummett speaks of a person's having a "mastery of a name" or of having "grasped the use of a name," and he even ponders whether, in a given circumstance, one person might have a better grasp of the use of a name than another. Here are some examples, in which Dummett speaks of the difficulty in stating precisely the point at which one could be said to grasp the sense of a proper name.

If a person knows of Milan only that it is a city somewhere on the continent of Europe, we should hardly ascribe to him a complete grasp of the name 'Milan.' But how much exactly should he know in order to be said to have such a grasp?⁷⁴

Suppose that someone, largely ignorant of science, knows of General Relativity that it is that branch of physics in which his nephew specialized at the university. In one respect, he attaches a definite sense to the name 'General Relativity Theory': he has a reasonably precise criterion for identifying some branch of physics as being what the name denotes. Should we say that he has a better grasp of the name than does someone who has no such criterion, but can give a sketchy and inadequate account of the theory?⁷⁵

But, surely, if two students are enrolled in a course on General Relativity Theory, and one received an A in the course and the other a C, it would be absurd to say that the first student had a better grasp of the use of the name 'General Relativity Theory' than the second. No; the first had a better grasp of General Relativity

Theory, not 'General Relativity Theory'. In the very same vein, I would hardly suppose that the Mayor of Milan has a complete grasp of the use of the name 'Milan', even a better grasp of the use of the name 'Milan' than I have. He might know a great deal more about the city than I do; but this hardly counts as having a better grasp of the use of the name 'Milan'.

This notion of "grasping the use of a name" is not to be found in Frege.

On the other hand, there does seem to me to be a sense in which one can speak of someone's having grasped the use of a name, but this would be the sense in which one might speak of a child's having learned to use a language: to grasp the use of a name would be to know, in a general sense, how proper names function in the language. If, for example, the correct view about proper names were that proper names are labels or tags, then to have grasped the use of a proper name would be to know that proper names are labels or tags and to use them as such. And, it seems to me that it is Kripke, and not Dummett, who has attempted to provide an account of what it is to grasp the use of proper names in this sense. For, Kripke is very much concerned with the fact that proper names are different from definite descriptions and that they have a very special role in language, and he has attempted to say what this difference amounts to; while

Dummett, on the other hand, has not acknowledged any such difference, and, as such, he has failed to acknowledge that proper names do have a special role in the language. To this extent, then, it is Kripke, and not Dummett, who has attempted to capture "our grasp of the use of proper names."

Now, Dummett appears to be way off base in his characterization of Frege's notion of sense; and I am inclined to believe that this is connected with Dummett's supposing that for Frege the description fixes the reference of a name. Let me then say a word about Kripke's distinction between using a description to give the meaning of a name and using it to fix the reference. For a description to give the meaning of a name, the two would be synonymous: if, for example, 'Cicero' meant the same as 'the Roman orator who denounced Catiline' then, since the two expressions would be synonymous, it could not be the case that anything be Cicero and not have denounced Catiline. For a description to fix the reference of a name, the two need only in fact have the same reference: if, for example, 'the Roman orator who denounced Catiline' is used to fix the reference of 'Cicero', then although Cicero is the one who denounced Catiline, he need not have been. If 'the Roman orator who denounced Catiline' were used to fix the reference of the term 'Cicero,' then the description would serve to

tell someone who Cicero is, not what the name 'Cicero' means. Kripke believes that insofar as a proper name is introduced by a definite description, as is envisaged by both Russell and Frege, the description serves to fix the reference, not to give the meaning. For if, he says, 'the Roman orator who denounced Catiline' were equivalent in meaning with 'Cicero', then

Cicero = the Roman orator who denounced Catiline, would express a proposition which was necessarily true. And this, he argues is surely wrong: the proposition that Cicero was the Roman orator who denounced Catiline is contingent; it is only a contingent fact about Cicero that he denounced Catiline, i.e., he might not have done so.

Now, Dummett, as we have seen, takes Frege's notion of sense to be such that the description fixes the reference of the name. The sense of a name is that which one knows if one knows the use of the name, and anything one knows about the referent, apparently, will fall into this category, there being no one favored fact--the meaning--which serves as our ultimate criterion for determining reference. Hence, we find Dummett forced into the position of saying that the more one knows about the referent, the greater the mastery or grasp of the name one has.

The first part of Dummett's defense of Frege, then, seems to me to be just wild; the second part, in

which he attempts to show how, on Frege's view, sense determines reference, is also unconvincing.

On the description theory of proper names, the referent of a proper name is to be whoever it is that satisfies the associated description. Kripke, however, drives a wedge between the name and the associated description by presenting cases in which we would be inclined to take the referent of a proper name as being something other than the object picked out by the associated description. For example:

Suppose that someone says that no prophet ever was swallowed by a big fish or a whale. Does it follow, on that basis, that Jonah did not exist? There still seems to be the question whether the Biblical account is a legendary account of no person or a legendary account built on a real person. In the latter case, it's only natural to say that though Jonah did exist, no one did the things commonly related to him. I choose this case because while Biblical scholars generally hold that Jonah did exist, the account not only of his being swallowed by a big fish but even going to Nineveh to preach or anything else that is said in the Biblical story is assumed to be substantially false. But nevertheless there are reasons for thinking this was about a real prophet.⁷⁶

Another of Kripke's examples is this: one might suppose the name 'Aristotle' to be associated with the description 'the teacher of Alexander', but this description serves at best to fix the reference of the name, not to give the meaning. For, Kripke says, it is only a contingent property of Aristotle that he taught Alexander: that very man might not have been the one to have taught Alexander.

And if this is so, then the description cannot be the ultimate appeal for determining the reference of the proper name 'Aristotle'.

Dummett construes these cases as demonstrating the inadequacy of a view on which one supposes a proper name to be associated with a single description:

Not only can [the objection] be met by a modification of a Wittgensteinian type, but consideration shows that only such a modification will account for those cases in which the objection fails. For, in so far as a single definite description supplies someone with the only means he has for determining the reference of a name, he cannot treat as intelligible the suggestion that the description does not apply to the bearer of the name.⁷⁷

Now, it is not clear to me at all what is gained by supposing the name to be associated with a cluster of descriptions instead of a single description; in particular, as McKinsey⁷⁸ has recently shown, when one attempts to make precise the connection supposed to hold between the name and the associated cluster of descriptions, there is no advantage gained. But, furthermore, Dummett's belief that the cluster theory is necessitated by such examples seems to be simply wrong. With regard to another of Kripke's cases, in which Kripke attacks the view that the description 'the man to have first proved the incompleteness of arithmetic' gives the meaning of 'Gödel', Dummett says:

What makes it possible to entertain the possibility that Gödel might be discovered not to have proved, or not to have been the first to prove, the incompleteness of arithmetic is the fact that there exist other generally accepted ways of determining the reference of the name 'Gödel.' This is always the case with any name about whose bearer a good deal is known by at least some who use the name; and it is never the case with a name about whose bearer practically nothing is known save that it satisfies the description which fixes the reference of the name. Hence something like Wittgenstein's modification of Frege's account is not merely adequate to meet this objection, but is actually called for by the facts.⁷⁹

However, it seems to me to be false that there need be other "generally accepted ways of determining the reference of [a] name" in order that it be possible to doubt that the referent satisfies the description at hand. I might find, in an ancient manuscript, the phrase 'John, the son of Harry, . . . ' with no other reference to John: as such, the description 'the son of Harry' is the only description available for determining the reference of 'John', and yet I can still entertain the possibility that John not be the son of Harry. Furthermore, Kripke's causal account is designed specifically to show how the reference of a name is determined independent of there being an available description: namely, there is a baptism of the object, and a historical link between the current use of the name and the original baptism. This is exceedingly vague, as Kripke himself admits, but it does give a different 'picture' of how reference is

determined, and it also shows that the view Dummett advocates is not required by these examples.

Moreover, Dummett's attempt to explain how the reference of a given proper name is determined on Frege's theory is seriously inadequate. Consider, for example, his handling of the following case:

Someone who can pick out Innsbruck from among other cities only by the fact that it was there that his Aunt Rosemarie broke her ankle will nevertheless, in using the name 'Innsbruck,' intend to be taken as referring to the city for which that name is ordinarily understood as standing; so that, if he happens to be mistaken in supposing that it was Innsbruck that Aunt Rosemarie broke her ankle, it will nevertheless have been Innsbruck, and not the city where that accident in fact occurred, that he will have been talking about.⁸⁰

This account is one with which Kripke would have very little to quarrel, at least in essentials. But, it does not seem to have penetrated Dummett that the situation he describes here is not one his "criterion of identity" account can easily handle. If sense determines reference, then, if my sole criterion for determining the referent of 'Innsbruck' is that it is the city in which Aunt Rosemarie broke her ankle, and if Innsbruck is, in fact, not that city, then the sense I attach to the word would, it seems to me, pick out some city other than Innsbruck. If, for example, it should turn out to be the case that someone tells me that Aunt Rosemarie broke her ankle in, say, Vienna, then I would no doubt drop this

description as part of the sense of the term 'Innsbruck'. But, then, all I am left with as a criterion of identity is that 'Innsbruck' is whatever is ordinarily taken to be such.

Dummett is certainly among the ablest of philosophers working within Frege's semantic framework, and his unsatisfactory defense of Frege's theory indicates that Kripke has located some very deep problems with the theory. Whether Kripke's own causal theory of reference fares any better, I cannot say, and I have nothing of significance to say about it. What I think I can contribute is some clarity about Kripke's own view of meaning; both critics and defenders of Kripke seem to me to have missed what he has said.

Footnotes to Chapter 3

¹Gottlob Frege, "On Sense and Reference," trans. Max Black, in Translations from the Philosophical Writings of Gottlob Frege, ed. Peter Geach and Max Black (Oxford: Blackwell, 1960), p. 78.

²Bertrand Russell, "On Denoting," rpt. in Logic and Knowledge, ed. Robert Charles Marsh (London: George Allen & Unwin, 1956), p. 50.

³Ibid., p. 49.

⁴Alonzo Church, "Carnap's Introduction to Semantics," The Philosophical Review, LII (1943), 298-304.

⁵John R. Searle, "Russell's Objection to Frege's Theory of Sense and Reference," Analysis, XVIII (1957-8), 137-143.

⁶John R. Searle, "Proper Names," Mind, 67 (1958), 166-173, and Speech Acts (Cambridge: Cambridge University Press, 1969).

⁷Michael Dummett, Frege, Philosophy of Language (London: Duckworth, 1973).

⁸John Stuart Mill, Philosophy of Scientific Method, ed. Ernest Nagel (New York: Hafner, 1950).

⁹Saul Kripke, "Naming and Necessity," in Semantics of Natural Language, ed. Gilbert Harman and Donald Davidson (Dordrecht: D. Reidel, 1972), pp. 253-355, and "Identity and Necessity," in Identity and

Individuation, ed. Milton K. Munitz (New York: NYU Press, 1971), pp. 135-164.

¹⁰Peter Geach trans., "Begriffsschrift (Chapter I)," in Translations from the Philosophical Writings of Gottlob Frege, ed. Peter Geach and Max Black (Oxford: Blackwell, 1960), p. 11.

¹¹Ibid., p. 12.

¹²Ibid., p. 11.

¹³Ibid., pp. 11-12.

¹⁴Frege, "On Sense and Reference," op. cit., p. 57.

¹⁵Leonard Linsky, Referring (New York: Humanities Press, 1967), p. 22.

¹⁶William Kneale and Martha Kneale, The Development of Logic (Oxford: Oxford University Press, 1962), p. 494.

¹⁷Gottlob Frege, The Foundations of Arithmetic, rev. ed., trans. J. L. Austin (Oxford: Blackwell, 1953), p. 79.

¹⁸Ibid., p. 22.

¹⁹Frege, "On Sense and Reference," op. cit., pp. 56-57.

²⁰See especially "Frege Against the Formalists," trans. Max Black, in Translations from the Philosophical Writings of Gottlob Frege, ed. Peter Geach and Max Black (Oxford: Blackwell, 1960), pp. 182-233.

²¹Howard Jackson, "Frege's Ontology," The

Philosophical Review, LXIX (1960), 394-395.

²²W. V. O. Quine, "Reference and Modality," in From a Logical Point of View, 2nd ed. rev. (Cambridge: Harvard University Press, 1961), p. 140.

²³Dummett, op. cit., pp. 196-203.

²⁴Frege, "On Sense and Reference," op. cit., p. 59.

²⁵Gottlob Frege, "The Thought," trans. A. M. and Marcelle Quinton, rpt. in Philosophical Logic, ed. P. F. Strawson (Oxford: Oxford University Press, 1967), p. 29. "Der Gedanke" originally appeared in Beiträge zur Philosophie des deutschen Idealismus, I (1918-19), 58-77.

²⁶Peter Geach trans., "Illustrative Extracts from Frege's Review of Husserl's Philosophie der Arithmetik," in Translations from the Philosophical Writings of Gottlob Frege, ed. Peter Geach and Max Black (Oxford: Blackwell, 1960), p. 79.

²⁷Frege, "The Thought," op. cit., p. 27.

²⁸Ibid., p. 27.

²⁹This has been called in question recently by Dummett (op. cit., pp. 264-294), who suggests that in opaque contexts the ordinary sense is both the indirect reference and the indirect sense. I am not sure that this can be coherently worked out; nevertheless, it is clear that Dummett's suggestion runs counter to the spirit of Frege's theory.

³⁰Michael Dummett, "Note: Frege on Functions," The Philosophical Review, LXV (1956), 229-230.

³¹Gottlob Frege, "Compound Thoughts," trans. R. H. Stoothoff, Mind, LXXII (1963), p. 1. "Gedankengefüge" originally appeared in Beiträge zur Philosophie des deutschen Idealismus, III (1923), 36-51.

³²Frege, "On Sense and Reference," op. cit., pp. 62-63.

³³Gottlob Frege, The Basic Laws of Arithmetic: Exposition of the System, trans. and ed. Montgomery Furth (Berkeley and Los Angeles: University of California Press, 1964), p. 90.

³⁴For a discussion of the issues, see David Kaplan, "Quantifying In," in Words and Objections, ed. Donald Davidson and Jaakko Hintikka (Dordrecht: D. Reidel, 1969), pp. 178-214.

³⁵Dummett, Frege, Philosophy of Language, p. 146.

³⁶Ibid., p. 142.

³⁷Ibid., p. 110.

³⁸Ibid., p. 142.

³⁹Ibid., p. 135.

⁴⁰Ibid., p. 111.

⁴¹Ibid., p. 143.

⁴²Kripke, "Naming and Necessity," op. cit., p.

⁴³ Ibid., p. 257.

⁴⁴ Mill, op. cit., p. 19.

⁴⁵ Ibid., p. 20.

⁴⁶ Ibid., p. 21.

⁴⁷ Ibid., p. 24.

⁴⁸ Ibid., p. 25.

⁴⁹ Ibid., p. 28.

⁵⁰ Ibid., p. 24.

⁵¹ Ibid., p. 24.

⁵² Ibid., p. 26.

⁵³ Ibid., p. 27.

⁵⁴ Ibid., p. 22.

⁵⁵ Ibid., p. 29.

⁵⁶ This is meant to distinguish proper names from definite descriptions; whether this serves also to distinguish proper names from general terms is open to question.

⁵⁷ Bertrand Russell, Introduction to Mathematical Philosophy (London: George Allen & Unwin, 1919), pp. 178-179.

⁵⁸ Bertrand Russell, The Principles of Mathematics, 2nd ed. (New York: Norton, 1938).

⁵⁹ Bertrand Russell, "On Denoting," op. cit.

⁶⁰ Alfred North Whitehead and Bertrand Russell, Principia Mathematica, 2nd ed., I (Cambridge: Cambridge University Press, 1927), p. 173.

- ⁶¹ Ibid., p. 174.
- ⁶² Ibid., p. 180.
- ⁶³ Ibid., p. 184.
- ⁶⁴ Kripke, "Naming and Necessity," op. cit., p. 255.
- ⁶⁵ Dummett, Frege, Philosophy of Language, p. 110.
- ⁶⁶ Kripke, op. cit., p. 268.
- ⁶⁷ Dummett, op. cit., p. 145.
- ⁶⁸ Kripke, op. cit., p. 349.
- ⁶⁹ Frege, "On Sense and Reference," op. cit., p. 58.
- ⁷⁰ Dummett, op. cit., p. 111.
- ⁷¹ Kripke, op. cit., p. 277.
- ⁷² Dummett, op. cit., p. 133.
- ⁷³ Ibid., p. 147.
- ⁷⁴ Ibid., p. 138.
- ⁷⁵ Ibid., p. 138.
- ⁷⁶ Kripke, op. cit., p. 282.
- ⁷⁷ Dummett, op. cit., p. 136.
- ⁷⁸ Michael McKinsey, "Searle on Proper Names," The Philosophical Review, LXXX (1971), 220-229.
- ⁷⁹ Dummett, op. cit., p. 136.
- ⁸⁰ Ibid., p. 140.

CHAPTER 4

KRIPKE'S RESPONSE TO THE PARADOX

1. Introduction

As we saw in the last chapter, Kripke's attack on Frege's description theory of proper names is extremely damaging. Whether Kripke's own causal account is any more successful at explaining how the reference of a proper name is determined is, as Kripke himself acknowledges, open to question; in any event, the causal theory needs to be made much more precise before any considered judgment of it is possible. However, the most difficult problem facing Kripke, if he is to defend Mill's view of proper names, is this: Kripke must show how, on Mill's view, $(\alpha=\alpha)$ and true $(\alpha=\beta)$ can differ in cognitive value when α and β are proper names. For it is widely believed that the fact that $(\alpha=\alpha)$ and true $(\alpha=\beta)$ can differ in cognitive value when α and β are proper names demands that α and β have different meaning; and, so long as Kripke is unable to show otherwise, there remains the feeling that, despite Kripke's damaging criticism, Frege's description theory--or something very much like it--must be the right one.

The point can be put very simply as follows. If

(1) Hesperus = Hesperus,

and

(2) Hesperus = Phosphorus,

expressed different propositions, then the information conveyed by 'Hesperus' would have to be different from that conveyed by 'Phosphorus'. Where is this difference in information to be located? Not, certainly, in the entities denoted, for 'Hesperus' and 'Phosphorus' denote the very same object. Nor, it is commonly argued, in the words themselves, for though the names differ, there is good reason to deny that the proposition expressed by (2) presupposes (or entails) the existence of a language. It would seem, then, that 'Hesperus' and 'Phosphorus' would have to differ in meaning. Conversely, if the sole significance of a proper name is to stand for an object, then it is difficult to see how (1) and (2) can express different propositions. Indeed, many of those who hold that proper names lack meaning would deny that $(\alpha=\alpha)$ and true $(\alpha=\beta)$ express different propositions when α and β are proper names. Russell,¹ for example, maintained that if α were a genuine proper name, then the object α stands for--not a concept--would be a constituent of the proposition expressed by $S\alpha$; and so, on this view, $(\alpha=\alpha)$ and true $(\alpha=\beta)$ would express the very same proposition when α and β are genuine proper

names. More recently, Alvin Plantinga² has also maintained that $(\alpha=\alpha)$ and true $(\alpha=\beta)$ express the same proposition when α and β are both proper names, although, on Plantinga's view, it is not the object itself which is a constituent of the proposition but an essence of the object.

The view one holds about whether proper names have or lack meaning, then, is intimately connected with the view one holds about whether $(\alpha=\alpha)$ and true $(\alpha=\beta)$ express the same or different propositions when α and β are proper names. Frege's description theory of proper names would appear to go hand-in-hand with the view that, e.g., (1) and (2) express different propositions; Mill's theory of proper names, on the other hand, would appear to go hand-in-hand with the view that (1) and (2) express the same proposition. And the preeminence of Frege's theory of proper names is a consequence of the widely held belief that sentences like (1) and (2) express different propositions.

What reason is there for supposing that (1) and (2) express different propositions? One reason is this: it is widely believed that it is a necessary truth that Hesperus = Hesperus, and only a contingent truth that Hesperus = Phosphorus. However, our modern defenders of Mill's view, Kripke and Plantinga, both deny this: they hold that $(\alpha=\beta)$ is necessarily true, if true at all,

whenever α and β are proper names. Their defense of this view depends heavily on their distinguishing the metaphysical issue of whether a given proposition is necessary or contingent from the epistemological issue of whether the proposition is known a priori or a posteriori. For, another reason for supposing that (1) and (2) express different propositions is that, while it is reasonable to suppose that it is known a priori that Hesperus = Hesperus (if it is reasonable to suppose that any proposition at all is known a priori), it would seem to be an a posteriori truth that Hesperus = Phosphorus. Now, Plantinga and Kripke agree in separating the epistemological status of a given proposition from its metaphysical status: indeed, both view the running together of the metaphysical/epistemological issues as a concomitant to the view that proper names are disguised descriptions. However, Plantinga and Kripke disagree on the epistemological status of the proposition expressed by (2). It is Plantinga's view that if α and β are proper names, then ' $\alpha=\alpha$ ' and true ' $\alpha=\beta$ ' express the very same proposition; so, on Plantinga's view, if it is known a priori that Hesperus = Hesperus, then it is also known a priori that Hesperus = Phosphorus, for, on Plantinga's view, the proposition that Hesperus = Hesperus is the very same proposition as the proposition that Hesperus = Phosphorus. Plantinga, of course, is thus faced with the problem of explaining what it is that constitutes

the great astronomical discovery: it could not be that the Babylonians discovered that Hesperus = Phosphorus, for this, as we have just seen, is something the Babylonians knew a priori. Now, Plantinga does attempt to explain what the great discovery amounts to, and I shall examine it in detail below: I will argue, however, that Plantinga has failed to make plausible his view that (1) and (2) express the very same proposition.

Kripke's view, on the other hand, is that (2) expresses a proposition that is necessary and a posteriori, while (1) expresses a proposition that is necessary and a priori. That is, Kripke holds that (1) and (2) express different propositions even though 'Hesperus' and 'Phosphorus' are proper names, and so lack meaning. I quote at length his statement of the distinction between the metaphysical notions and the epistemological notions:

What do we mean by calling a statement necessary? We simply mean that the statement in question, first, is true, and, second, that it could not have been otherwise. When we say that something is contingently true, we mean that, though it is in fact the case, it could have been the case that things would have been otherwise. If we wish to assign this distinction to a branch of philosophy, we should assign it to metaphysics. To the contrary, there is the notion of an a priori truth. And a priori truth is supposed to be one which can be known to be true independently of all experience. Notice that this does not in and of itself say anything about all possible worlds, unless this is put into the definition. All that it says is that it can be known to be true of the actual world, independently of all experience. It may, by some philosophical argument, follow from our knowing, independently of experience, that something is true of the actual world,

that it has to be known to be true also of all possible worlds. But if this is to be established, it requires some philosophical argument to establish it. Now, this notion, if we were to assign it to a branch of philosophy, belongs, not to metaphysics, but to epistemology. It has to do with the way we can know certain things to be in fact true. Now, it may be the case, of course, that anything which is necessary is something which can be known a priori. . . . I will not have time to explore these notions in full detail here, but one thing we can see from the outset is that these two notions are by no means trivially the same. If they are coextensive, it takes some philosophical argument to establish it. As stated, they belong to different domains of philosophy. One of them has something to do with knowledge, of what can be known in certain ways about the actual world. The other one has to do with metaphysics, how the world could have been; given that it is the way it is, could it have been otherwise, in certain ways? Now I hold, as a matter of fact, that neither class of statements is contained in the other.³

It is Kripke's position, then, that the notions of necessity and a-prioricity are not merely different concepts, but that the predicates ' ξ is necessarily true' and ' ξ is a priori true' are not even coextensive: the class of necessarily true propositions perhaps overlaps the class of a priori true propositions, but neither class is contained in the other. There are, for Kripke, four possibilities: a proposition might be (i) necessary and a priori, (ii) necessary and a posteriori, (iii) contingent and a priori, or (iv) contingent and a posteriori. Cases (i) and (iv) are relatively uncontroversial; it is Kripke's claim that there are propositions of types (ii) and (iii) that is of significance. Kripke argues that, where we stipulate that one meter is to be the length of a given

stick S at t_0 , then,

(3) The length of stick S at t_0 is one meter, expresses a proposition of type (iii), i.e., contingent and a priori; and he argues that (2) expresses a proposition of type (iv), i.e., necessary and a posteriori. Kripke reserves the bulk of his argument to show that (2) expresses a proposition of type (iv), though he of course intends the result to hold good for the general case: "We have concluded," he says, "that an identity statement between names, when true at all, is necessarily true, even though one may not know it a priori."⁴

Kripke thus offers a new and radical response to the paradox. In this chapter, I will examine Kripke's claim that one and the same proposition is both necessary and a posteriori. I am basically interested in establishing the coherence of Kripke's view, and especially in correcting Dummett's⁵ misrepresentations, which have gained currency and contributed to a wide-spread misunderstanding of Kripke's position.

2. A Modal Paradox

It is widely believed that among true identities, some are necessary and others contingent. It would seem, for example, that the proposition that

(4) $9=9$

is a necessary truth: it simply could not be the case that 9 be anything other than 9. On the other hand, the proposition that

(5) The number of planets = 9

would appear to be a contingent truth. It is true that there are exactly 9 planets, but, one is inclined to say, it is only contingently true that this is so; the facts could certainly have been otherwise, e.g., it could have been that there were 10 planets. Indeed, I think it would be fair to say that in so far as the distinction between necessary and contingent truths marks any distinction at all, it would be just that distinction we perceive between (4) and (5). Yet there is a well known argument which appears to threaten the viability of this distinction, for it purports to show that if x is identical with y , then x is necessarily identical with y . I should like now to look at this argument.

Assume the following are both true:

(6) $(x)(y)(x=y \supset (Fx \supset Fy))$,

(7) $(x)\Box(x=x)$.

Substituting ' $\Box(x=\xi)$ ' for ' $F\xi$ ' in (6), we obtain,

(8) $(x)(y)(x=y \supset (\Box(x=x) \supset \Box(x=y)))$;

and since (7) has been assumed true, we can delete ' $\Box(x=x)$ ' from (8) to obtain,

(9) $(x)(y)(x=y \supset \Box(x=y))$.

Once we have derived (9), the rest is trivial. Suppose that we have a true identity,

$$(10) \quad a=b.$$

Applying Universal Instantiation twice to (9), we obtain,

$$(11) \quad a=b \supset \Box(a=b);$$

and by Modus Ponens, finally, we derive,

$$(12) \quad \Box(a=b).$$

Of the two assumptions made in the argument, (6), it is widely agreed, cannot plausibly be denied. And so our suspicions turn to (7).

In considering (7), however, we must be careful to distinguish,

$$(7) \quad (x) \Box(x=x),$$

from

$$(13) \quad \Box(x)(x=x).$$

One might be inclined to suppose that (7) is true because '(x)(x=x)' is a logical truth (of First Order Logic with Identity), and all logical truths are necessary truths. Certainly, on the traditional view of necessity, all logical truths are necessary truths; and this fact is embodied in one of the rules of inference said to characterize normal modal systems, namely,

$$(14) \quad \text{If } \vdash p, \text{ then } \vdash \Box p.$$

But from our logical truth '(x)(x=x)', (14) allows us to infer (13); it does not allow us to derive (7). In order to obtain (7) from (13), we require a principle which

allows us to move the modal operator inside the quantifier; and such principles are notoriously controversial. The one that immediately comes to mind, of course, is the converse Barcan formula,

$$(15) \quad \Box (x)Fx \supset (x)\Box Fx,$$

which would allow us to derive (7) from (13). But the derivability of the Barcan formula in the usual systems of modal logic has been seriously questioned by Kripke;⁷ and, moreover, there are fairly cogent intuitive reasons for supposing the converse Barcan formula to be false, and so for not including it in any of the usual modal systems.

What reason, then, is there for accepting (7)?

The difference between (7) and (13) corresponds to a distinction drawn by the Medievals between necessity de re and necessity de dicto. What is said to be necessary in (13) is the proposition (dictum) that each thing is identical with itself. In (7), on the other hand, it is the property of being identical with itself that is said to be necessarily true of each thing (res). We de re necessity, one is not speaking of a proposition's being necessarily true or contingently true, but of an object's having a property necessarily (essentially) or contingently (accidentally). Now, this distinction between essential and accidental properties is not a particularly clear distinction; but I would think that if there are any essential

properties, being identical with itself would certainly be one. And so, it is this commitment to there being essential properties that leads one to believe (7) to be true.

Anyone who rejects the notion of de re necessity will not be affected by the argument; so long as one restricts one's ascriptions of necessity to propositions, the argument poses no threat to the intuition that (4) expresses a proposition that is necessarily true, and that (5) expresses a proposition that is contingently true. It is the Essentialist who is directly affected by the argument.

The Essentialist can gain some relief from the paradoxical conclusion by adapting Arthur Smullyan's extension of Russell's scope distinction for definite descriptions to modal contexts.⁸ According to Russell, when a sentence containing a definite description is itself embedded in a sentence, an ambiguity arises about the scope of the description. Smullyan shows that this holds true as well when we have a sentence containing a definite description embedded in a sentence containing a modal operator. In the sentence,

(16) It is necessary that the number of planets = 9,
the definite description 'the number of planets' might be taken as having large scope, in which case (16) would be symbolized as

(17) $(\exists x)((y)(x=y \equiv Py) \cdot \Box x=9)$;

or, 'the number of planets' might be given small scope, in which case (16) would be symbolized as

(18) $\Box (\exists x)((y)(x=y \equiv Py)) \cdot x=9$.

When (16) is taken to have the logical structure of (17), it has the de re reading: (16) expresses that that number, which in fact numbers the planets, is such that it is necessarily identical with 9. When, on the other hand, (16) is taken to have the logical structure of (18), it has the de dicto reading: (16) then expresses that the proposition that the number of planets is identical with 9 is necessarily true. Clearly, when (16) is given the de dicto reading, it expresses something false, for, the proposition that the number of planets is identical with 9 is only contingently true. When (16) however, is given the de re interpretation, it is plausible to suppose that it expresses something true: that number which in fact numbers the planets is 9, and surely, that number could not be anything other than 9. Now, Smullyan shows how, consonant with Russell's Principia Mathematica rules governing the iota operator, (17), i.e., the true reading, can be validly derived from (5) and (9), and (18), i.e., the false reading, cannot be validly derived in this manner. In this way, the paradoxical conclusion is somewhat disarmed: (5) expresses a contingent proposition, as we intuitively

believe to be the case, and the de re reading does not controvert this, for, on the de re reading of (5), (5) expresses that a particular number has an essential property, namely, that of being identical with 9.

Smullyan's proposal, however, works only for those i-sentences in which we find definite descriptions. What happens when we have an i-sentence, $(\alpha=\beta)$, where α and β are both proper names? Here, Cartwright's comments are pertinent:

The ambiguity in question has been noticed by others, but some who have seen it have gone on to characterize it inadequately. Arthur Smullyan saw it as having to do with the scope of definite descriptions, and he accordingly proposed to treat it by means of an extended version of Russell's theory. But though this may serve the purposes of disambiguation in some cases, it does not in all. For '9 is necessarily greater than 7' would appear to admit of both de dicto and de re readings. Of course, it is open to Smullyan to argue that '9' is in reality a disguised definite description; but he countenances names that are not, and with these sentences exhibiting the ambiguity will surely be constructible.⁹

I agree with Cartwright that sentences of the form $(\Box F\alpha)$ exhibit a de re/de dicto ambiguity even when α is a proper name; indeed, this holds true even of i-sentences $(\alpha=\beta)$ where α and β are both proper names. The sentence (19) It is necessary that Hesperus = Phosphorus, for example, is ambiguous: it might be read de dicto, expressing that the proposition that Hesperus = Phosphorus is necessarily true, or, on the other hand, it might be read de re, expressing (say) that Hesperus is such that it

(that object) is necessarily identical with Phosphorus. However, I think that Cartwright has overlooked a relatively noncontroversial way in which Smullyan's proposal can be extended to handle cases like (19). Quine has shown how, in his canonical notation, proper names can be eliminated entirely, replacing, for example, 'Socrates', a proper name, by 'the Socratizer', a definite description.¹⁰ This is a purely syntactic trick Quine has used here, and it does not appear, so far as I can tell, to beg any of the important questions about proper names. Taking proper names as definite descriptions in this way, then, Smullyan's scope distinction will be applicable to sentences like (19); and it therefore seems to me that Smullyan's scope distinction does characterize the de re/de dicto distinction adequately.

However, there are many who apparently believe that the de re/de dicto distinction cannot be drawn for sentences of the form $(\exists x Fx)$ where α is a proper name, or more accurately, that the distinction collapses in such cases; and apparently, they are inclined to take this view because they believe that proper names are not, in some substantial way, disguised descriptions. Ruth Marcus,¹¹ Alvin Plantinga,¹² David Kaplan,¹³ and Saul Kripke¹⁴ all reject the Russellian view of proper names and favor instead Mill's view that proper names are simply

labels or tags. These philosophers reject the claim that (19) is false on the de dicto interpretation; they hold (setting aside questions about worlds in which Hesperus does not exist) that if Hesperus is necessarily identical with Phosphorus, then it is necessarily true that Hesperus is identical with Phosphorus.

What stands in the way of these anti-Fregean views is the wide-spread belief that (2) is a contingent proposition. Quine, for example, provided the following response to Marcus' use of the distinction between a tag and a description to buttress her case that (2) was a necessary truth:

[I] think I see trouble anyway in the contrast between proper names and descriptions as Professor Marcus draws it. Her paradigm of the assigning of proper names is tagging. We may tag the planet Venus, some fine evening, with the proper name 'Hesperus'. We may tag the same planet again, some day before sunrise, with the proper name 'Phosphorus'. When at last we discover that we have tagged the same planet twice, our discovery is empirical. And not because the proper names were descriptions.¹⁵

That is, even though 'Hesperus' and 'Phosphorus' are both proper names, and even if we suppose that proper names are tags, the truth of (2) was ascertained by empirical means and it is therefore an a posteriori truth; and as such, it would surely seem to be a contingent truth. Conversely, the view under discussion would seem to be committed either to the view that (2) is not a posteriori, or, alternatively, that there are propositions that are both necessary truths and a posteriori truths.

The first alternative is that maintained by Plantinga; the second, by Kripke. I will consider both of these positions below.

3. Rigid Designators

Kripke believes Mill to have been right, and both Frege and Russell wrong, on this issue of proper names: proper names have no meaning, but are simply labels or tags for the things to which they are assigned. Part of what is involved in their acting in this manner is that they be what Kripke calls rigid designators. This notion of a rigid designator is very important for Kripke, and I should like to explain it carefully in this section.

Here is Kripke's introduction of the notion of a rigid designator:

Let's use some terms quasitechnically. Let's call something a rigid designator if in any possible world it designates the same object, a nonrigid or accidental designator if that is not the case. Of course we don't require that the objects exist in all possible worlds. Certainly Nixon might not have existed if his parents had not gotten married, in the normal course of things. When we think of a property as essential to an object we usually mean that it is true of that object in any case where it would have existed. A rigid designator of a necessary existent can be called strongly rigid.¹⁶

A rigid designator, then, is a term that designates the same object in every possible world in which the object exists; and whether a given expression is a rigid designator is determined as follows:

. . . we can perfectly well talk about rigid and non-rigid designators. Moreover, we have a simple, intuitive test for them. We can say, for example, that the number of planets might have been a different number from the number it in fact is. For example, there might have been only seven planets. We can say that the inventor of bifocals might have been someone other than the man who in fact invented bifocals. We cannot say, though, that the square root of 81 might have been a different number from the number it in fact is, for that number just has to be 9. If we apply this intuitive test to proper names, such as for example 'Richard Nixon', they would seem intuitively to come out to be rigid designators. First, when we talk even about the counterfactual situation in which we suppose Nixon to have done different things, we assume we are still talking about Nixon himself. We say, "If Nixon had bribed a certain Senator, he would have gotten Carswell through," and we assume that by 'Nixon' and 'Carswell' we are still referring to the very same people as in the actual world. And it seems that we cannot say "Nixon might have been a different man from the man he in fact was," unless, of course, we mean it metaphorically: He might have been a different sort of person (if you believe in free will and that people are not inherently corrupt). You might think the statement true in that sense, but Nixon could not have been in the other literal sense a different person from the person he, in fact, is, even though the thirty-seventh President of the United States might have been Humphrey. So, the phrase "the thirty-seventh President" is non-rigid, but "Nixon", it would seem, is rigid.¹⁷

Let us see now this test works. Kripke says that the description 'the number of planets' is not a rigid designator because it is true that

- (20) The number of planets might have been a different number from the number it in fact is.

Clearly, what Kripke has in mind here is not

- (21) It is possible that (the number of planets \neq the number of planets),

for this is trivially false; rather, he must mean

(22) The number of planets is such that it is possible that it \neq the number of planets.

It is (22) Kripke claims to be true, not (21); and the truth of (22) shows that 'the number of planets' is not a rigid designator. Obviously, then, in order to determine whether 'the number of planets' is a rigid designator, we have to separate the de re reading of (20) from the de dicto reading; it is the truth or falsity of the de re reading that determines whether the designator is rigid or nonrigid. And this must be true whether our designator is a definite description or a proper name. For, consider Kripke's claim that 'Nixon' is a rigid designator. 'Nixon', he argues, is a rigid designator because the following is false:

(23) Nixon might not in fact have been Nixon.

What Kripke has in mind here, clearly, is not

(24) It is possible that (Nixon \neq Nixon),

for, although (24) is false, it is trivially false, just as (21) is: the falsity of (24) does not distinguish 'Nixon' from any other singular term in an interesting manner. Hence, Kripke must mean that

(25) Nixon is such that it is possible that he \neq Nixon, is false, and this, of course, is the de re interpretation of (23).

Kripke is therefore committed to the de re/de dicto distinction for modal sentences containing proper names as well as for modal sentences containing definite descriptions.

I had earlier, in considering the paradoxical argument on pp. 171-2 shown how one could make the scope distinction for proper names without adopting the substantial Russellian description theory of proper names: we use Quine's technique of replacing 'Socrates' by, say, 'the Socratizer'. This being a description, the scope distinction will be available to us to capture the de re/de dicto distinction for proper names. In this way, the ambiguity seen in, e.g.,

(26) Nine is necessarily greater than seven,

will be of the same sort as that found in

(27) The number of planets is necessarily greater than seven,

which, as Kripke says, is adequately handled by the scope distinction. Is Kripke's thesis about proper names in any way affected by this syntactic trick of Quine's? Kripke himself indicates that it is not:

When I speak of the Frege-Russell view and its variants, I include only those versions which give a substantive theory of the reference of names. In particular, Quine's proposal that in a 'canonical notation' a name such as 'Socrates' should be replaced by a description 'the Socratizer' (where 'Socratizes' is an invented predicate), and that the description should then be eliminated by Russell's method, was not intended as a theory of reference for names but as a proposed reform of language with certain advantages. The problems discussed here will all apply, mutatis mutandis, to the reformed language; in particular, the question, 'How is the reference of 'Socrates' determined? yields to the question, 'How is the extension of 'Socratizes' determined?' Of course I do not suggest that Quine has ever claimed the contrary.¹⁸

Apparently, then, it is consistent with Kripke's views about proper names that we eliminate them à la Quine; and if this is so, then Smullyan's scope distinction will be available to us.

What is the rigid/nonrigid distinction intended to explain? We have already discussed Smullyan's scope distinction and seen how it can be employed to explain an ambiguity in certain modal statements, and also how it can be employed to avoid the paradoxical consequence of the argument we had considered on pp. 171-2. Kripke believes this solution to be adequate, at least in so far as our singular terms are definite descriptions:

Provided that the notion of modality de re, and thus of quantifying into modal contexts, makes any sense at all, we have quite an adequate solution to the problem of avoiding paradoxes if we substitute descriptions for the universal quantifiers in [(9)] because the only consequence we will draw, for example, in the bifocals case, is that there is a man who both happened to have invented bifocals and happened to have been the first Postmaster General of the United States, and is necessarily self-identical. There is an object x such that x invented bifocals, and as a matter of contingent fact an object y , such that y is the first Postmaster General of the United States, and finally, it is necessary, that x is y . What are x and y here? Here, x and y are both Benjamin Franklin, and it can certainly be necessary that Benjamin Franklin is identical with himself. So, there is no problem in the case of descriptions if we accept Russell's notion of scope. And I just dogmatically want to drop that question here and go on to the question about names. . . .¹⁹

Kripke, however, rejects this solution for the case of proper names:

It would . . . seem that the function of names is simply to refer, and not to describe the objects so named by such properties as "being the inventor of bifocals" or "being the first Postmaster General." It would seem that Leibniz' law and the law [(6)] should not only hold in the universally quantified form, but also in the form "if a=b and Fa, then Fb," wherever 'a' and 'b' stand in place of names and 'F' stands in place of a predicate expressing a genuine property of the object:

$$(a=b.Fa) \supset Fb$$

We can run the same argument through again to obtain the conclusion where 'a' and 'b' replace any names, "if a=b, then necessarily a=b." And so, we could venture this conclusion: that whenever 'a' and 'b' are proper names, if a is b, that it is necessary that a is b. Identity statements between proper names have to be necessary if they are going to be true at all.²⁰

But what is it that Kripke is rejecting in the case of proper names?

We have seen that Kripke believes that Smullyan's scope distinction provides the basis for an adequate response to the paradox when the i-sentence contains definite descriptions. Why, then, does the appeal to the scope distinction constitute an inadequate response to the paradox in the case of proper names? The reason is not that Kripke denies that the scope distinction can be drawn for proper names, for he does no such thing. The reason is that, for Kripke, the identities are necessary if true, whatever scope is given: that is, invoking the scope distinction for proper names would be incorrect if it were attempted thus to show how there can be contingent identities when the expressions used are proper names. A rigid

designator, then, will be, roughly, one for which scope makes no difference.²¹ Note that Kripke's claim that 'Hesperus' is a rigid designator does not require that he regard

(29) Hesperus is such that it is necessarily identical with Phosphorus,

and

(30) It is necessary that Hesperus is identical with Phosphorus,

as expressing the same proposition. In fact, as we have argued, Kripke holds that (29) and (30) express different propositions, the former, that Hesperus has an essential property, the latter, that a given proposition is necessary. That 'Hesperus' is a rigid designator only requires that (29) and (30) have the same truth value.²²

I have spent a considerable amount of time trying to state Kripke's view clearly because Dummett, in his criticism of Kripke's notion of a rigid designator, has gotten matters quite confused. Let me turn, now, to Dummett's criticisms.

Dummett claims that modal statements of the form 'It is necessary (contingent) that α is F' are ambiguous, whether α be a proper name or a definite description:

. . . any theory which represents proper names and definite descriptions as functioning in essentially similar manners has an advantage over one that

widens the difference between them, in that it allows a uniform explanation to be given of what appears to be just the same phenomenon--the occurrence of ambiguity in modal contexts--in the two cases.²³

Dummett believes that the ambiguity in statements of the form 'It is necessary (contingent) that α is F' is adequately and correctly diagnosed as an ambiguity in the scope of the necessity operator; and that the distinction in the scope of the necessity operator corresponds to the distinction between a statement's being necessary and an object's having a property necessarily (or essentially).

Now, Dummett poses this view which I have just outlined as standing in opposition to Kripke's view:

Kripke, on the other hand, wants to give an entirely different explanation of the phenomenon when it relates to proper names. In this case, he acknowledges no role for the notion of scope: and so he explains the ambiguity by saying that we are concerned, under the two interpretations, with different modal notions, different kinds of possibility.²⁴

That is, according to Dummett, Kripke has denied that there is a distinction for proper names, and he has explained the ambiguity arising when we have modal statements containing proper names by invoking a distinction between metaphysical necessity and epistemic necessity. Thus, Dummett says:

Kripke's doctrine that proper names are rigid designators and definite descriptions non-rigid ones thus reduces to the claim that, within a modal context, the scope of a definite description should always be taken to exclude the modal operator, whereas the scope of a proper name should always be taken to include it.²⁵

And, in another passage:

This can only be interpreted as the thesis that, in a modal context, a definite description must always be construed as lying within the scope of the modal operator, while a proper name must always be construed as lying outside its scope. To assign to a term a reference varying from one possible world to another is just to take it as having, in each possible world, the reference which it would bear in that world; conversely, to assign it a constant reference is to take it as having, in each world, just that reference which it has in the real world. But to take a term in the former of these two ways is precisely to treat it as being within the scope of the modal operator, while to take it in the second way is to treat it as falling outside that scope.²⁶

However, it seems to me that Dummett has gotten Kripke all wrong.

First, Dummett supposes that on Kripke's view proper names are rigid designators and definite descriptions are not; and this is just inaccurate. For Kripke, the rigid/nonrigid distinction does not correspond to the proper name/definite description distinction. A rigid designator is one that designates the same object in every possible world (in which the object exists); a designator is nonrigid otherwise. Every proper name is, on Kripke's view, a rigid designator; but the converse does not hold, i.e., not every rigid designator is a proper name. For example, the description 'the square of 2' designates the same object in every possible world, namely, the number 4, and so it is a rigid designator; but it is obviously not a proper name. Kripke's own example is 'the ratio of the circumference of a circle to its diameter', which designates

π in every possible world.²⁷ Hence, Kripke does not seek to distinguish proper names from definite descriptions by their behavior in modal contexts, as Dummett claims.

Second, it is quite clear that Kripke does acknowledge the scope distinction for proper names as well as for definite descriptions, as I have pointed out earlier.

Dummett sees this, but he finds this to be a peculiarity or inconsistency in Kripke's view:

. . . in order to understand the sort of contingency Kripke alleges to exist in these cases, we are compelled after all to invoke just that notion of scope to which Kripke appealed in the case of definite descriptions. . . . It is thus not merely that the uniform explanation, in terms of scope, of the ambiguity that occurs when either definite descriptions or proper names occur in modal contexts is preferable, because more economical, than having in the latter case to introduce the a priori/necessary distinction: it is that, in order to understand the notions of necessity and contingency that Kripke uses, we find ourselves forced to appeal to the notion of scope, for proper names as well as definite descriptions.²⁸

But to admit the ambiguity for proper names, and thus to admit a scope distinction for proper names, is not yet to admit that the two readings need differ in truth value.

Dummett and Kripke thus agree that

(23) Nixon might not in fact have been Nixon,

is ambiguous. It can be taken either as

(24) It is possible that (Nixon \neq Nixon),

or as

(25) Nixon is such that it is possible that he \neq Nixon.

On the other hand, Dummett and Kripke differ on the truth

value of (25): Dummett believes (25) to be true, and Kripke believes (25) to be false. Kripke believes (25) to be false because 'Nixon', on his view, is a rigid designator, one that designates the same thing in every possible world (in which Nixon exists).

It is clear, then, that Dummett and Kripke also disagree on the connection between the scope distinction and the assigning of reference. Let us just recall Dummett's statement:

To assign to a term a reference varying from one possible world to another is just to take it as having, in each possible world, the reference which it would bear in that world; conversely, to assign it a constant reference is to take it as having, in each world, just that reference which it has in the real world.²⁹

So far, Kripke would agree. But, Dummett continues:

But to take a term in the former of these two ways is precisely to treat it as being within the scope of the modal operator, while to take it in the second way is to treat it as falling outside that scope.³⁰

And this, Kripke believes to be just false. (It ought not be surprising, however, that Dummett and Kripke should disagree on this point; for it is here that the connection is made between one's view about proper names and one's view about the notion of necessity.)

Dummett appears to be maintaining a certain picture about possible worlds Kripke explicitly seeks to reject. Kripke describes this picture as follows:

One thinks, in this picture, of a possible world as if it were like a foreign country. One looks upon it as an observer. Maybe Nixon has moved to the other country and maybe he hasn't, but one is given only qualities. One can observe all his qualities, but of course, one doesn't observe that someone is Nixon. One observes that something has red hair (or green or yellow) but not whether something is Nixon. So we had better have a way of telling in terms of properties when we run into the same thing again as we saw before; we had better have a way of telling, when we come across one of these other possible worlds, who was Nixon.³¹

And again:

It is as if a 'possible world' were like a foreign country, or distant planet way out there. It is as if we see dimly through a telescope various actors on this distant planet. Actually David Lewis' view seems the most reasonable if one takes this picture literally. No one far away on another planet can be strictly identical with someone here. But, even if we have some marvelous methods of transportation to take one and the same person from planet to planet, we really need some epistemological criteria of identity to be able to say whether someone on this distant planet is the same person as someone here.³²

Now, Dummett voices some uneasiness about the metaphor of possible worlds--as does Kripke himself--but he does appear to be holding just this picture Kripke describes. How is the reference of, say, 'Nixon', determined in another possible world on Dummett's view? The reference of 'Nixon' will be whoever in that world has the properties specified in the criterion of identity associated with the name 'Nixon': in so far as different objects may, in different worlds, satisfy this criterion, it need not be the case that the object which is Nixon in one world is

one and the same as the object which is Nixon in another. On Dummett's view, then, 'Nixon' need not have a constant denotation. Kripke's view about possible worlds is somewhat different. "Possible worlds are stipulated, not discovered by powerful telescopes,"³³ he says. That is, Kripke takes the domain of a possible world to be given by specifying the objects themselves, not, as David Lewis³⁴ would have it, by saying that something is in a given possible world if it satisfies a certain qualitative description. On Lewis' view, one determines whether a given object is in a possible world by determining whether it has the requisite properties; not so on Kripke's view. On Kripke's view, a proper name is not a disguised description, and so the reference of, say, 'Nixon', in another possible world will be determined not by seeing who, in that world, has such-and-such properties, but by seeing who, in that world, is Nixon. Hence, on Kripke's view, to treat a proper name as having constant reference is not thereby to take the name as falling outside the scope of the modal operator. For Kripke still distinguishes between a proposition's being necessary and an object's having a property necessarily.

It seems fairly clear, then, that Dummett has not read Kripke at all sympathetically. He has simply ignored Kripke's careful attempt to dissociate his own view about

possible worlds from the view of someone like David Lewis, and he has therefore failed to appreciate the distinction between (29) and (30) having the same truth value, and their expressing the same proposition. In the next section, I will explore the consequence of failing to observe this distinction.

4. Plantinga's Position

Do

(1) Hesperus is identical with Hesperus,
and

(2) Hesperus is identical with Phosphorus,

express the same proposition? Alvin Plantinga, who holds a view about proper names which is similar to Kripke's, adopts the view that (1) and (2) do express the same proposition, and having acknowledged this, he squarely faces the issue of accounting for the felt difference in cognitive value between (1) and (2). I should like to turn now to Plantinga and examine his attempt at a solution to the Paradox of Identity.

Here is Plantinga's statement of his position, and of the problem he faces:

If my account is accurate, 'Hesperus' and 'Phosphorus' express essences. This conceded, it is plausible to suppose that they express the very same essence. If so, however, does it not follow that the sentences
(20) Hesperus is identical with Phosphorus

- (21) Hesperus was named 'Phosphorus'
 and
 (22) Hesperus has the property of being identical
 with Phosphorus
 express the very same propositions as
 (20') Phosphorus is identical with Phosphorus
 (21') Phosphorus was named 'Phosphorus'
 and
 (22') Phosphorus has the property of being identical
 with Phosphorus?

I think it does. But what about the Babylonian discovery that Hesperus is identical with Phosphorus? Suppose for simplicity that the Ancient Babylonians spoke English rather than Ancient Babylonian. Prior to their discovery, the Babylonian astronomers would have accepted the primed items as trifling trivialities. But their attitude towards the unprimed items, one supposes, was quite different; here their attitude was one of suspension of belief if not outright denial. If 'Hesperus' and 'Phosphorus' are proper names of Venus, however, then (20) expresses the very same proposition as (20'). And if (20) expresses the very same proposition as (20'), then the Babylonians did believe the former, since they believed the latter. How, then, can we account for their sincere claim to reject (20)? Shall we suppose that they believed (20)--the proposition--but did not know or believe that they believed it? No; for no doubt they knew they believed (20'); and (20) is (20').³⁵

Let us make sure that we understand Plantinga's position.

- It is Plantinga's position, not only that (1) and (2) express the very same proposition, but also that the pair
- (31) Hesperus was named 'Phosphorus',
 and
 (32) Phosphorus was named 'Phosphorus',
 express the very same proposition, and, again, that the pair
- (33) Hesperus has the property of being identical with
 Phosphorus,

and

(34) Phosphorus has the property of being identical with
Phosphorus,

express the very same proposition. Clearly, there is a general principle at work here. Plantinga claims that it follows, in each case, that both sentences express the same proposition from the fact that 'Hesperus' and 'Phosphorus' express the same essence. Now, in each case, we have a pair of sentences which differ one from the other only in that where the first contains the name 'Hesperus', the second contains the name 'Phosphorus'. Obviously, then, Plantinga is assuming some substitution principle to the effect, roughly, that when α and β are singular terms that express the same essence, then $S\alpha$ and $S\alpha/\beta$ are sentences that express the same proposition. I say roughly because Plantinga would surely not wish to hold that we can replace 'Hesperus' by 'Phosphorus' anywhere and preserve propositionhood; Plantinga would, for example, surely wish to eliminate those contexts wherein the names occurred within quotation marks. Of course, such a context is one in which codenotational singular terms would not even (necessarily) preserve truth value, and I suppose that this is the most reasonable qualification to put on the substitution principle. That is, I suppose that the substitution principle Plantinga has in mind here (at least,

the weakest one I can think of) is that whenever substitution of one codenotational term for another preserves truth value, then, whenever the two terms also express the same essence, then substitution of one for the other also preserves propositionhood. And, supposing that the extensionality condition reduces to $S\alpha$'s being about $d(\alpha)$, then we can phrase this substitution principle as follows:

- (35) If $S\alpha$ is about $d(\alpha)$, then if α and β express the same essence, then $S\alpha$ and $S\alpha/\beta$ express the same proposition.

We need not worry about the elaborate technical mechanism Plantinga constructs in his book to explicate the notion of essence. To say that a given singular term expresses an essence, on Plantinga's view, seems to be roughly the same as saying that the term is a rigid designator:

. . . proper names express essences. The proper name 'Aristotle' expresses an essence of Aristotle. It therefore expresses a property P that is instantiated by the same object in every world--in every world in which P is instantiated, of course; for there are worlds in which Aristotle does not exist. What is characteristic of proper names, then, is that the properties they express are instantiated by the same objects in every world.³⁶

But, as Plantinga himself says, his account, if accurate, gives him that the proper names 'Hesperus' and 'Phosphorus' each expresses an essence. However, it does not appear to be necessary that the two expressions express the same essence: Plantinga gives us no argument to this effect

and does not even indicate that there is one, saying only that it is plausible that they express the same essence. And, again, even if we cede that it is plausible that 'Hesperus' and 'Phosphorus' express the same essence, Plantinga still offers no argument to show that (1) and (2) must therefore express the same proposition: he says that he thinks that it follows, and, at the level at which we are interested in this argument, that is sufficient for us. That is, we shall take this passage in the spirit in which Plantinga offers it to us: He believes it to be true that 'Hesperus' and 'Phosphorus' express the same essence, and he also believes it to be true that (1) and (2) (and the other pairs) must therefore express the same proposition. He is not interested, at this point, in arguing for this position, but only in showing how it can account for the difference in cognitive value between (1) and (2). In this section he is explaining how, given his belief in the facts stated but not argued for, this felt difference in cognitive value is to be explained.

As we see, then, it is Plantinga's position that (1) and (2) express the very same proposition; or, put in a slightly different way, Plantinga holds that (36) the proposition that Hesperus is identical with Hesperus, is the very same proposition as (i.e., is identical with)

(37) the proposition that Hesperus is identical with Phosphorus.

And since, on Plantinga's view as well as on Frege's, propositions (as opposed, say, to sentences) are the objects of belief, knowledge, assertion, and the other propositional attitudes, it follows that anyone who believes the proposition that Hesperus is identical with Hesperus to be true thereby believes the proposition that Hesperus is identical with Phosphorus to be true, these being, as we have just mentioned, the very same proposition. In particular, it turns out, as Plantinga explicitly acknowledges, that

(38) The Babylonians discovered that Hesperus is identical with Phosphorus,

must have the same truth value as

(39) The Babylonians discovered that Hesperus is identical with Phosphorus.

Apparently, then, Plantinga deems it to be a valid argument to infer (38) from (39) and

(40) 'Hesperus' expresses the very same essence as 'Phosphorus'.

That is, Plantinga believes that since 'Hesperus' and 'Phosphorus' express the very same essence, we can substitute the former for the latter in (38) and obtain a sentence, (39), that has the same truth value as the original. Now, the substitution principle appealed to here is not

(35); for (35) sanctions substitution of terms in extensional contexts only, and the substitution here is made in a notoriously nonextensional context. Hence, Plantinga is assuming a different substitution principle. We can state this principle as follows. If $S\alpha$ is an extensional construction on singular terms, then if α and β express the very same essence, $\text{that-}S\alpha$ and $\text{that-}S\alpha/\beta$ have the same truth value. That is,

(41) If α and β express the same essence, then $\text{that-}S\alpha$ and $\text{that-}S\alpha/\beta$ have the same truth value.

Clearly, this will allow Plantinga to derive (39) from (38) and (40).

Actually, Plantinga's remarks at the very end of the quoted passage indicate that he holds a stronger substitution principle. Plantinga considers whether it is true that

(42) The Babylonians believed that they believed that Hesperus is identical with Phosphorus,

and he argues that (42) is true because it is true that

(43) The Babylonians believed that they believed that Phosphorus is identical with Phosphorus,

and the proposition that Hesperus is identical with Phosphorus is, as has already been admitted, the very same proposition as the proposition that Phosphorus is identical with Phosphorus. But the inference of (42) from (43) is

not sanctioned by our substitution principle (41). What (41) allows is that, since 'Hesperus' and 'Phosphorus' express the very same essence, then

(44) The Babylonians believed that Hesperus is identical with Phosphorus

has the same truth value as

(45) The Babylonians believed that Hesperus is identical with Hesperus.

What is required for the inference to (43) is not that (44) and (45) have the same truth value, but that they express the same proposition. That is, Plantinga apparently believes, if his inference from (43) to (42) is to make sense--and below we shall question whether he really wants to maintain this--the following substitution principle:

(46) If α and β express the same essence, then that- $S\alpha$ and that- $S\alpha/\beta$ express the same proposition.

Where does this leave us, now, with regard to the Paradox of Identity? It was Frege's view that (1) and (2) did not express the same proposition because they differed in cognitive value: the Babylonians discovered that Hesperus was identical with Phosphorus, not that Hesperus is identical with Hesperus. In so far as they knew that Hesperus existed (and this discovery is not, so far as I know, credited to the Babylonians), and in so far as they knew the basic logical fact that each thing is identical with

itself, then the Babylonians knew--but did not discover-- that Hesperus is identical with Hesperus. Frege therefore maintained that (1) and (2) differed in cognitive value, and therefore, that they did not express the same proposition; and Frege also maintained that 'Hesperus' and 'Phosphorus' differed in sense, because they differed in the contribution they made to the proposition expressed by a sentence containing them. What is Plantinga's view of the paradox?

Plantinga holds that (1) and (2) express the very same proposition; and he therefore holds that anyone who believes the proposition expressed by (1) thereby believes the proposition expressed by (2), these being the very same proposition. Supposing Seth, then, to have predated the great Babylonian discovery, and supposing it to be true that

(47) Seth knew that Hesperus is identical with Hesperus, it would follow, on Plantinga's analysis, that it is also true that

(48) Seth knew that Hesperus is identical with Phosphorus. On Plantinga's view, then, the proposition that Hesperus is identical with Phosphorus was known to be true by Seth, and, interestingly enough, it was known to be true in advance of the great Babylonian discovery. But, then, what was the great Babylonian discovery? It would, I think, be surely incorrect to suppose it to be true that the

Babylonians discovered that Hesperus is identical with Hesperus. But since (1) and (2) express the very same proposition, it would also be incorrect to suppose that the Babylonians discovered that Hesperus is identical with Phosphorus. It would seem, then, that on Plantinga's analysis, (38) is false: the Babylonians did not discover that Hesperus is identical with Phosphorus. Now, I find this result to be utterly implausible: indeed, I would, as would Frege, take this result as a reductio of Plantinga's claim that (1) and (2) expressed the very same proposition.

What, then, according to Plantinga, did the Babylonians discover? Here is one attempt he makes at the end of the relevant section:

It was not that the Babylonian linguistic competence was insufficient, not that they did not grasp or apprehend the proposition expressed by the sentence (20) Hesperus is identical with Phosphorus that is, the proposition that Hesperus is identical with Hesperus. There may have been people who had no grasp of this proposition--people who had never studied astronomy or looked at the night sky, for example. But this was not Babylonian trouble. Their problem was that they did not realize that (20) expresses the proposition that Hesperus is identical with Hesperus; and they were not apprised of this fact because they did not know that Hesperus bore the name 'Phosphorus' as well as the name 'Hesperus!'.³⁷

However, I find Plantinga to be confused here and to have failed to have provided an explanation for the fact that the Babylonians discovered something to be true that had not been known to be true before.

In this passage, Plantinga is attempting to locate what the Babylonians had been ignorant of prior to their great discovery. This ignorance, Plantinga says, did not stem from linguistic incompetence, and in this, Plantinga is surely right. For one would like to say that their ignorance was about the nature of the heavens, and that they were enlightened, by by studying language, but by studying the heavens. What knowledge, then, did they lack? Plantinga says that the Babylonians did not realize that (2) expressed a certain proposition, i.e., the Babylonians did not realize that

(49) 'Hesperus is identical with Phosphorus' expresses that Hesperus is identical with Hesperus.

Now, it is not immediately clear to me what to make of this claim, but the reason Plantinga gives for supposing that the Babylonians did not realize that (49) is true is surely wrong. Plantinga says that the Babylonians did not realize that 'Hesperus is identical with Phosphorus' expressed that Hesperus is identical with Hesperus because, he says, the Babylonians did not know that Hesperus bore the name 'Phosphorus' as well as the name 'Hesperus!'. That is, Plantinga claims (if I understand him correctly) that the Babylonians were ignorant of which proposition (2) expressed because of the following:

(50) The Babylonians did not know both that Hesperus was called 'Hesperus' and that Hesperus was called 'Phosphorus'.

And, I cannot see how Plantinga can maintain that (50) is true. For, surely, it is true that

(51) The Babylonians knew that Hesperus is called 'Hesperus',

and, again, it is true that

(52) The Babylonians knew that Phosphorus is called 'Phosphorus'.

But, it is obvious that on Plantinga's view,

(53) Phosphorus is called 'Phosphorus'

and

(54) Hesperus is called 'Phosphorus'

express the same proposition, since 'Hesperus' and 'Phosphorus' express the same essence and (35) sanctions the substitution. And since (53) and (54) express the very same proposition, then if (54) is true, then, by (41),

(55) The Babylonians knew that Hesperus is called 'Phosphorus'

would also have to be true. But, if (52) and (53) are both true, then the Babylonians knew Hesperus is called 'Phosphorus' and they knew that Hesperus is called 'Hesperus'; and if so, then they knew that Hesperus was called 'Phosphorus' and 'Hesperus'. Hence, it would appear that on Plantinga's own view, (52) would have to be false.

Whatever the merits of Plantinga's proposal, viz., that the Babylonians did not know that (2) expressed a

certain proposition, this much is clear: Plantinga's stated reason for the Babylonians' ignorance will not do. It simply could not be the case, as Plantinga claims, that the Babylonians failed to see that Hesperus was called both 'Hesperus' and 'Phosphorus'--if, as Plantinga also wishes to hold, 'Hesperus' expresses an essence, and the very same essence as 'Phosphorus'. I do not think that this is sufficient to justify rejection of Plantinga's attempted solution to the Paradox, but, at the very least, it is certainly sufficient to cast grave doubts on the explanation he offers: the Babylonians could not have been ignorant for the reason Plantinga gives, and since it is not clear wherein their ignorance might arise, (other than the usual reasons, which Plantinga cannot use), it would seem to be a clearly questionable issue whether they were ignorant in the way Plantinga claims.

Our next question, then, is this: Can Plantinga adequately characterize the Babylonians' ignorance in the way he does in the quoted passage, namely, by holding it to be false that

(56) The Babylonians knew that 'Hesperus is identical with Phosphorus' expresses the proposition that Hesperus is identical with Hesperus?

There seems to me, again, to be a rather serious problem with this explanation. For, surely, Plantinga would want

to hold that

(57) The Babylonians knew that 'Hesperus is identical with Phosphorus' expresses the proposition that Hesperus is identical with Phosphorus.

So, on Plantinga's view, it would seem that (57) is true and (56) is false. But, judging from the way in which he sets up the problem in the first passage we quoted, it seems to me that Plantinga cannot consistently maintain that (56) is false and (57) true. For, returning to our substitution principle (46) above, which Plantinga certainly seems to be committed to, we see that since (56) and (57) differ only in that the one contains the name 'Hesperus' and the other 'Phosphorus', and since these two names express the same essence, (56) and (57) would not only have the same truth value, they would, according to (46), also have to express the same proposition. So, again, I do not see how Plantinga could consistently offer this as his explanation of the ignorance the Babylonians eventually overcame. And this also redeems a remark I made earlier, that it was not clear to me whether Plantinga actually wished to commit himself to (46): for the whole thrust of his attempt at reconciling his view on essences and propositions with the Paradox requires that sameness of proposition not be preserved in nested that-constructions.

Let us see, then, if we can try to capture what Plantinga was trying to get at, and set aside this substitution principle (46). Consider the following explanation Plantinga offers:

The Babylonians believed that an utterance of the sentence 'this is identical with that' expresses a true proposition when the speaker accompanies 'this' with a demonstration (a pointing to) of Venus, and, an instant later, accompanies 'that' the same way. Had the circumstances surrounding the utterance of this sentence been different, however, the Babylonians would have been doubtful. Imagine it thus: pointing to the evening sky, to Venus, we say (very slowly) "This is not identical with (long pause) that" (pointing to the eastern sky, to Venus, some fine morning when Venus is the morning star). Now let us suppose that the sentence expresses the same proposition on these two occasions. Of course the Babylonians were not apprised of this fact. But why not? Not because they had an insufficient grasp of the semantic role of demonstratives such as 'this' and 'that', nor because of any other insufficiency in their command of the language. Their difficulty was rather their failure to realize that the second occurrence of 'that' was accompanied by a demonstration of the same heavenly body as was the second occurrence of 'this' and the first occurrences of 'this' and 'that'. And this defect in their knowledge issued in their failure to realize that the second occurrence of 'this is identical with that' expressed the same proposition as the first occurrence. The truth is they did not really know what proposition was expressed by that second occurrence.

In the same way, the Babylonian astronomers were ignorant of the fact that

(20) Hesperus is identical with Phosphorus
and

(20') Phosphorus is identical with Phosphorus
express the same proposition. They did not really know what proposition was expressed by (20). The latter did indeed express a proposition, and one with which they were acquainted. But they did not know that this proposition was the one (20) expressed. They knew that (20) expressed a proposition, and they knew that the proposition expressed

by (20) was true if and only if the first heavenly body to appear in the evening had the property of lingering longer in the morning than any other heavenly body. Still, they did not know that (20) expresses the proposition that Phosphorus is identical with Phosphorus.³⁸

Plantinga's suggestion in the quoted passage is that the Babylonians were in some way ignorant or mistaken about what proposition a given sentence expresses. In particular, he says that the Babylonians did not 'really' know what proposition was expressed by (2). I'm not sure what it is to really know what proposition a given sentence expresses, but I would certainly think it to be true that

(58) The Babylonians knew that 'Hesperus is identical with Phosphorus' expresses that Hesperus is identical with Phosphorus,

and, to that extent, the Babylonians knew which proposition the sentence expresses. Again, it seems to be clear that the Babylonians were sufficiently knowledgeable to grasp or apprehend the proposition expressed by (2), that is,

(59) The Babylonians grasped or apprehended the proposition that Hesperus is identical with Phosphorus.

Now, I suppose that Plantinga would agree that both are true, and I do not think he would wish to deny that the Babylonians knew which proposition a given sentence expressed in this sense. Well, then, in which sense did he mean it? Plantinga suggests that the Babylonians were

ignorant of the fact that (1) and (2) expressed the same proposition. I take it, then, that Plantinga wishes to hold it to be true that

(60) The Babylonians discovered that the proposition that Hesperus is identical with Hesperus is the very same proposition as the proposition that Hesperus is identical with Phosphorus,

even though, as I understand it, he would hold it to be false that

(61) The Babylonians discovered that the proposition that Hesperus is identical with Hesperus is the same as the proposition that Hesperus is identical with Hesperus.

Again, Plantinga apparently wishes to hold it to be false that

(62) The Babylonians knew that 'Hesperus is identical with Phosphorus' expresses the proposition that Hesperus is identical with Hesperus.

even though, I would think, Plantinga would clearly wish to hold it to be true that

(63) The Babylonians knew that 'Hesperus is identical with Phosphorus' expresses the proposition that Hesperus is identical with Phosphorus.

I think, then, that Plantinga is attempting the following. In ceding that 'Hesperus' and 'Phosphorus' both express the same essence, Plantinga is ceding the

truth of

- (64) (The proposition that Hesperus = Hesperus) = (the proposition that Hesperus = Phosphorus).

However, he apparently wishes to distinguish this from

- (65) The proposition that [(the proposition that Hesperus = Hesperus) = (the proposition that Hesperus = Phosphorus)] = [(the proposition that Hesperus = Hesperus) = (the proposition that Hesperus = Hesperus)].

Again, Plantinga is committed to the truth of (64) and the truth of

- (66) (The proposition that Hesperus = Hesperus) = (the proposition that Hesperus = Hesperus),

but he apparently wishes to deny that he is therefore committed to the fact that (64) and (66) express the very same proposition. In other words, Plantinga has raised the Paradox of Identity to a new level. He apparently believes that when α and β are proper names for the same object, then ' $\alpha = \alpha$ ' and true ' $\alpha = \beta$ ' express the same proposition (i.e., have the same cognitive value) although it need not be the case that the trivially true

('the proposition that $\alpha = \alpha$ ') = ('the proposition that $\alpha = \alpha$ '),

expresses the same proposition as

('the proposition that $\alpha = \alpha$ ') = ('the proposition that $\alpha = \beta$ ').

Now, I am inclined to believe that this particular tactic on Plantinga's part is mistaken: I am inclined to

believe that the same sort of reasons which led him to deny that (1) and (2) expressed different propositions are also operative here, and should also lead him to deny that (64) and (66) express different propositions. More specifically, I think that the reasons which led Plantinga to decide that the proper names 'Hesperus' and 'Phosphorus' express essences are the same as those which should lead him to decide that 'The proposition that Hesperus = Hesperus' and 'The proposition that Hesperus = Phosphorus' also express essences.

An expression like

(67) The proposition that Hesperus = Phosphorus, is not, on the face of it, a proper name, at least not in the way in which 'Hesperus' is clearly a proper name; but, simply because we find 'the' in front, we ought not yet conclude that we have here a definite description. (67) ought clearly to be distinguished from, say,

(68) The proposition that Plantinga says the Babylonians knew to be true on p. 85 of his book.

Here we clearly have a description of the given proposition: it is the proposition said to have the property of being claimed by Plantinga, etc. And it is certainly a contingent property of this proposition that it was claimed to be such-and-such by Plantinga, that is,

(69) The proposition that Hesperus = Phosphorus is such that it might not have been the proposition Plantinga said the Babylonians knew to be true on p. 85 of his book.

But, it seems to me that (67) does not work in this way.

It does not seem to me that (67) has picked out the proposition by specifying a property it has, in the way (69) does;

(67) seems to me as close as one can come to a proper name for propositions. Just as it could not be the case that

(70) Hesperus might not be identical with Hesperus,

it could not be the case that

(71) The proposition that Hesperus = Phosphorus might not be identical with the proposition that Hesperus = Phosphorus.

Again, if this is so, then, clearly, (67) and

(72) the proposition that Hesperus = Hesperus

would express essences, and, I would suppose, the very same essence; and so (64) and (66) would express the same proposition.

If my argument is correct, then Plantinga has failed to explain adequately the nature of the Babylonian discovery. I don't claim to have refuted the view Plantinga was attempting to defend, namely, that (1) and (2) expressed the same proposition, but, (a) the explanation he offered fails, (b) it is not, as he supposes, so intuitively plausible that (1) and (2) express the same proposition.

tion, and (c) I don't see any way in which Plantinga could make it seem so.

5. Kripke's Position

Kripke rejects the common identification of the metaphysical notions of necessity and contingency with the epistemological notions of a priori and a posteriori truth. How something is discovered or known to be true, he says, is a question that belongs to epistemology; whether something might have been or might not have been true, on the other hand, is a question that belongs to metaphysics. And any attempt to identify the a priori/a posteriori distinction with the necessary/contingent distinction takes argument.

The terms 'necessary' and 'a priori', then, as applied to statements are not obvious synonyms. There may be a philosophical argument connecting them, perhaps even identifying them; but an argument is required, not simply the observation that the two terms are clearly interchangeable. (I will argue below that in fact they are not even coextensive--the necessary a posteriori truths, and probably contingent a priori truths, both exist.

I think people have thought that these two things must mean the same for these reasons. First, if something not only happens to be true in the actual world but is also true in all possible worlds, then, of course, just by running through all the possible worlds in our heads, we ought to be able with enough effort to see, if a statement is necessary, that it is necessary, and thus know it a priori. But really this is not so obviously feasible at all.³⁹

Kripke's position can be put roughly as follows: don't assume that the notions are the same, for that leads to

well known difficulties; prise them apart and we will find interesting philosophical answers to traditional problems. Of course, Kripke intends his denial in a stronger way, but at this point, let us take it in the spirit given. It is so novel an idea that a number of philosophers have failed to grasp it.

Plantinga, for example, rejects the identification of necessary truth with a priori truth, but he does so only in order to make plausible the claim that the proposition that Hesperus = Phosphorus is necessarily true. He does not go so far as Kripke, in holding that the very same proposition is known a posteriori. For Plantinga believes that the reasons for supposing that Hesperus = Phosphorus is necessary if true are reasons for supposing that the proposition that Hesperus = Phosphorus is the very same as the proposition that Hesperus = Hesperus, and since the latter is known a priori, so is the former. Of course this undercuts somewhat Plantinga's attempt to distinguish the necessary and the a priori; at any rate, if our argument in the last section is correct, Plantinga's position is untenable. In this section, I should like to consider Kripke's position, namely, that the proposition that Hesperus = Phosphorus is necessary and a posteriori. I am basically interested here in clarifying Kripke's position and establishing that it is coherent.

Let me begin with Kripke's distinction between what might be the case and what might have been the case:

If I say, "Gold might turn out not to be an element," I speak correctly: 'might' here is epistemic and expresses the fact that the evidence does not justify a priori (Cartesian) certainty that gold is an element. I am also strictly correct when I say that the elementhood of gold was discovered a posteriori. If I say, "Gold might have turned out not to be an element," I seem to mean this metaphysically. . . .⁴⁰

It seems to me to be reasonable for Kripke to distinguish what might be the case from what might have been the case; it does seem to capture a distinction we make in ordinary language between expressing some notion of epistemological possibility and some other notion of--following Kripke's terminology--metaphysical possibility. That is, to say that such-and-such could be the case or might be the case seems to mean something like: for all we know, such-and-such is the case. On the other hand, to say that such-and-such could have been the case or might have been the case (except where what we are talking about occurred in the past, in which case it could have been the case could be taken either way) is to mean something different. For example,

(73) The number of planets might be 7,

would seem to mean something like

(74) For all we know, there are 7 planets,

and this, surely, is false: given the epistemological

situation we are in, i.e., given all we know, surely, everybody knows that there are 9 planets. On the other hand,

(75) The number of planets might have been 7, cannot be understood in this way. (75), rather, has the force of

(76) It might have been the case that there be 7 planets, i.e., it is not necessary that there be 9 planets.

Kripke's claim, then, that

(2) Hesperus = Phosphorus,

expresses a proposition that is both necessary and a posteriori, is therefore the claim that

(77) It could not have been the case that Hesperus \neq Phosphorus,

while

(78) it could be that Hesperus \neq Phosphorus.

And Kripke explains how it is possible that (77) and (78) both be true as follows:

Any necessary truth, whether a priori or a posteriori, could not have turned out otherwise. In the case of some necessary a posteriori truths, however, we can say that under appropriate qualitatively identical evidential situations, an appropriate corresponding qualitative statement might have been false. . . . The inaccurate statement that Hesperus might have turned out not to be Phosphorus should be replaced by the true contingency mentioned earlier in these lectures: two distinct bodies might have occupied in the morning and the evening, respectively, the very positions actually occupied by Hesperus-

Phosphorus-Venus. . . . I have not given any general paradigm for the appropriate corresponding qualitative contingent statement. Since we are concerned with how things might have turned out otherwise, our general paradigm is to redescribe both the prior evidence and the statement qualitatively and claim that they are only contingently related. In the case of identities, using two rigid designators, such as the Hesperus-Phosphorus case above, there is a simpler paradigm which is often usable to at least approximately the same effect. Let ' R_1 ' and ' R_2 ' be the two rigid designators which flank the identity sign. Then ' $R_1 = R_2$ ' is necessary if true. The references of ' R_1 ' and ' R_2 ', respectively, may well be fixed by nonrigid designators ' D_1 ' and ' D_2 ' in the Hesperus and Phosphorus cases these have the form 'the heavenly body in such-and-such position in the sky in the evening (morning)'. Then although ' $R_1 = R_2$ ' is necessary, ' $D_1 = D_2$ ' may well be contingent, and this is often what leads to the erroneous view that ' $R_1 = R_2$ ' might have turned out otherwise.⁴¹

It might appear, from the way Kripke puts the matter, that what might be is one thing and what might have been another; that what is necessary is one thing, and what is a posteriori another. What appears to be necessary is that Hesperus = Phosphorus; what appears to be a posteriori, on the other hand, is that a planet having such-and-such characteristics is identical with a planet having such-and-such characteristics. Kripke might appear to be arguing, then, not that one and the same thing is both necessary and a posteriori, but that one thing is necessary, another a posteriori. Again, consider the following passage in which Kripke seeks to explain the sense in which the sentence

(79) This table is not made of ice,

expresses a proposition that is necessary and a posteriori:

. . . if I hold that this table could not have been made of ice, then I must also hold that it could not have turned out to be made of ice; it could have turned out that P entails that P could have been the case. What, then, does the intuition that the table might have turned out to have been made of ice or of anything else, that it might even have turned out not to be made of molecules, amount to? I think that it means simply that there might have been a table looking and feeling just like this one and placed in this very position in the room, which was in fact made of ice. In other words, I (or some conscious being) could have been qualitatively in the same epistemic situation that in fact obtains, I could have the same sensory evidence that I in fact have, about a table which was made of ice. The situation is thus akin to the one which inspired the counterpart theorists; when I speak of the possibility of the table turning out to be made of various things, I am speaking loosely. This table itself could not have had an origin different from the one it in fact had, but in a situation qualitatively identical to this one with respect to all the evidence I had in advance, the room could have contained a table made of ice in place of this one. Something like counterpart theory is thus applicable to the situation, but it applies only because we are not interested in what might have been true of this particular table, but what might or might not be true of a table given certain evidence. It is precisely because it is not true that this table might have been made of ice from the Thames that we must turn here to qualitative descriptions and counterparts. To apply these notions to genuine de re modalities is, from the present standpoint, perverse.⁴²

Now, Dummett believes Kripke to be holding just this view, and so he charges Kripke with holding that the notions of a-prioricity and of necessity apply to different sorts of things.

As we have seen, Kripke does not succeed in disentangling epistemic properties from others as completely as he claims. Properly speaking, Kripke's notions of contingency and necessity are not properties of statements at all, but of facts. His wish to dispense with the notion of sense for proper names leads him to regard a fact as consisting, e.g., in the possession by an object of a certain

property, or in two objects' standing to one another in a certain relation. A fact, so conceived, may be taken as forming the content of a particular statement, but it certainly cannot be identified with the thought expressed by the statement, as Frege conceives of it, and hence cannot properly speaking be said to be an object of knowledge at all. The knowledge which someone expresses by means of an assertion (when it is knowledge) is the knowledge that the thought expressed by the sentence used to make the assertion is true; it cannot, properly speaking, be taken to be the knowledge that that fact obtains (in Kripke's sense of 'fact') which is the content of the assertion. Thus, for instance, Kripke's notion of facts leads straight to the conclusion, willingly drawn by Kripke, that the fact which is the content of a true statement of identity is always a necessary one: for it is just the fact that a certain object bears to itself that relation which every object bears to itself and to no other. By adopting Russell's Theory of Descriptions, it is possible for Kripke to refrain from applying this doctrine to identity-statements involving definite descriptions. This, however, has no real bearing on the tenability of Kripke's view; it merely serves to make it less evident that the fact conveyed by a statement, as understood by Kripke, cannot be equated with its cognitive content, and thus to prepare the trap which Kripke falls into when he speaks of someone's knowing a contingent fact.⁴³

However, this interpretation of Kripke seems to me to be wrong.

Dummett is still working, of course, on the basis of his erroneous reading of Kripke mentioned earlier. Dummett believes that instead of drawing the scope distinction for proper names, Kripke had employed the epistemological/metaphysical distinction to the same purpose: according to Dummett, Kripke took the de re reading to be metaphysical necessity and the de dicto reading to be epistemological necessity. And so, according to Dummett, Kripke had applied these notions to two different things: he has

epistemic necessity applied to a statement, and metaphysical necessity applied not to a statement, but rather to an object's having a property. So, Dummett says, Kripke has not shown that one and the same thing is both necessary and a posteriori; what Kripke has shown if anything, is that the statement that Hesperus = Phosphorus is a posteriori and that Hesperus has the property of being necessarily identical with Phosphorus.

What makes Dummett's interpretation plausible is that Kripke, in explicating the notion of metaphysical necessity, employs a possible world semantics in which individuals can exist in more than one possible world, while for the epistemological notion of necessity, in so far as Kripke employs anything like a possible worlds semantics, it is not the possible worlds semantics for necessity but something which, as Kripke describes, is somewhat closer to Lewis' counterpart theory. Lewis' view is that no object exists in more than one possible world, but that an object in a given world has counterparts in others, related not by identity, but by similarity.⁴⁴ We should note, however, that it is not evident that Kripke assigns more than one interpretation to "Hesperus = Phosphorus". Perhaps he assigns distinct interpretations to

(80) It is (metaphysically) necessary that Hesperus =
Phosphorus,

and

(81) It is (epistemically) necessary that Hesperus =
Phosphorus.

But from this it does not follow that he assigns distinct interpretations to 'Hesperus = Phosphorus'.

But also, I think that Dummett is making a certain assumption here. If 'Hesperus' and 'Phosphorus' are rigid designators, then one can substitute one name for the other inside the scope of the necessity operator and preserve truth value. Given that these are rigid designators, it will follow that

(82) It is necessary that (if Hesperus exists) Hesperus =
Hesperus,

and,

(83) It is necessary that (if Hesperus exists) Hesperus =
Phosphorus

have the same truth value. Now, suppose

(84) It is known a priori that Hesperus = Hesperus.

Does the fact that 'Hesperus' and 'Phosphorus' are rigid designators allow us to infer

(85) It is known a priori that Hesperus = Phosphorus?

The answer is, I think, No: to say they are rigid designators is to say they designate the same object in every possible world in which the object exists. But that, the notion of possible world, has to do with metaphysical necessity, not the epistemological notion of necessity. To say that 'Hesperus' and 'Phosphorus' are rigid designators, then, is to allow us to substitute one for the other within

the context 'It is necessary that'; it does not thereby license substitution within the context 'it is known a priori that'.

It is not recognized how radical this suggestion of Kripke's is. Dummett takes Kripke to be holding the following:

It is rather natural to think that, while the actual reference of an expression relates only to the real world, its sense must be determined by what its reference would be in every possible world. For instance, must not the sense of a predicate both determine and be determined by what objects it would be true of in all possible circumstances? It therefore seems very plausible that we may identify the sense of an expression with what Kripke calls its 'meaning', i.e., the function which maps each possible world on to the reference of that expression within that world. Actually, however, this is an illusion. Kripke's notion of meaning is still a nonepistemic one; that is, it does not give an account of what it is that someone knows when he understands a word, which is precisely what the notion of sense, as introduced by Frege, is required to do. In certain cases, it is at least plausible that there will be a one-one correspondence between Frege's senses and Kripke's meanings: but that will be so only for words and expressions of which we can say that there is no gap between their meanings and the way in which their reference is determined, as we saw might, in general, be said of definite descriptions. Of course, even in such a case, Kripke's meaning is not a credible representation of the knowledge that someone has when he understands the expression: what someone grasps when he understands a predicate is the principle by which we determine whether or not it applies to any given object, not what its actual extension is in each of the infinitely many possible worlds.⁴⁵

Dummett's assumption, then, is that the notion of possible worlds is being used to explicate the notion of meaning. Now, this, classically, is the way in which necessity had been understood. The attempt to characterize the notion of necessary truth is intimately connected with the attempt to characterize the notion of meaning and, therefore,

the notion of proposition. It is usual for modal logicians to take a proposition as a function from sentences to the possible worlds in which the sentences are true, or, what comes to the same thing, to identify a proposition with the set of possible worlds in which it is true. And on this notion of proposition, the proposition expressed by (2) would obviously be identified with the proposition expressed by

(1) Hesperus = Hesperus,

for in every world in which the planet Venus exists, Hesperus would be none other than Hesperus, i.e., Phosphorus. This, I am inclined to believe, lies behind Plantinga's view: it is his view that because 'Hesperus' and 'Phosphorus' designate the same thing in every world (in which Hesperus exists), they therefore express essences, and (1) and (2) therefore express the very same proposition.

This identification of a proposition with the set of possible worlds in which it is true is, however, definitely inadequate: it does not cut fine enough. In every possible world in which the number 25 exists, for example, the number 25 would be identical with 5^2 . Hence,

(86) the proposition that $25 = 25$, and

(87) the proposition that $25 = 5^2$,

would be one and the same proposition; for (86) would be true in exactly those worlds in which (87) is true. But the inadequacy of this identification is even more dramatic when

we recognize that there would be, on this view, exactly one necessary truth: for if a proposition is identical with the set of possible worlds in which it is true, then any two sentences which express necessary propositions and hence, propositions that are true in every possible world, would therefore express the same proposition.

To get around this difficulty, Carnap⁴⁶ had proposed that we require as well that the propositions be intensionally isomorphic, roughly, that they be constructed from the same intensional entities and in the same way. With this added requirement, (86) would turn out to be a different proposition from (87), because '25' and '5²' correspond to entities which are not intensionally isomorphic. On Carnap's view, (1) and (2) would also be distinct propositions because 'Hesperus' and 'Phosphorus' express, as Carnap says, different individual concepts.

But Kripke explicitly rejects this notion of individual concept: proper names have no meaning in this sense. Those who accept possible world semantics and who also deny that proper names have meaning, would thus be forced to hold that (1) and (2) express the same proposition: Plantinga does, as we have seen, and so does David Kaplan.⁴⁷ Kripke, however, holds a different view.

Kripke is denying that possible world semantics exhausts the notions of meaning and of propositionhood:

I think, even in cases where the notion of rigidity versus accidentality of designation cannot be used to make out the difference in question, some things called definitions really intend to fix a reference rather than to give the meaning of a phrase, to give a synonym. Let me give an example. π is supposed to be the ratio of the circumference of a circle to its diameter. Now it's something that I have nothing but a vague intuitive feeling to argue for: it seems to me that here this Greek letter is not being used as short for the phrase 'the ratio of the circumference of a circle to its diameter' nor is it even used as short for a cluster of alternative definitions of π , whatever that might mean. It is used as a name for a real number, which in this case is necessarily the ratio of the circumference of a circle to its diameter. Note that here both ' π ' and 'the ratio of the circumference of a circle to its diameter' are rigid designators, so the arguments given in the metric case are inapplicable.⁴⁸

These are both rigid designators, then, and so they designate the same thing in every possible world; but as he says, he does not believe that the description gives the meaning of the name.

Again, we look to his definition of analytic truth:

. . . let's just make it a matter of stipulation that an analytic statement is in some sense true by virtue of its meaning and true in all possible worlds by virtue of its meaning. Then something which is analytically true will be both necessary and a priori.⁴⁹

In other words, what others have tried to capture about meaning in terms of possible worlds semantics, Kripke is saying is inadequate; meaning is rather an amalgamation of what is going on in possible world semantics as well as whatever would be needed to explicate the notion of epistemological necessity.

However, I am unable to state Kripke's position with any greater clarity. I hope that I have at least made

plausible Kripke's claim that

(1) Hesperus = Hesperus,

and

(2) Hesperus = Phosphorus,

might both express necessary truths even though (1) expresses an a priori truth and (2) an a posteriori truth.

But much remains to be done before Kripke could be said to have provided a solution to Frege's problem.

Since, on Kripke's view, (2) expresses a necessary a posteriori truth, it cannot be analytic, and so 'Hesperus' and 'Phosphorus' cannot be synonyms. Now, there are two ways in which 'Hesperus' and 'Phosphorus' can fail to have the same meaning: either (i), 'Hesperus' and 'Phosphorus' lack meaning altogether, and so fail to have the same meaning, or (ii), 'Hesperus' and 'Phosphorus' both have meaning, though different meaning. This second possibility does not seem consonant with the main thrust of Kripke's account, which is, of course, that Mill was right in supposing proper names to lack meaning. I would think that the first possibility is the obvious one Kripke should take. The problem is that if 'Hesperus' and 'Phosphorus' lack meaning, then it is difficult to see how to account for the contribution each makes to the proposition expressed by (2). The only relevant difference would seem to be in the names; but I do not think that Kripke would be any more comfortable in supposing that the "informational content" of (2) is that

the names name the same thing than Frege was.

Kripke seems to be holding that a proposition encompasses ingredients both of meaning and knowledge. (1) and (2) express different propositions, not because 'Hesperus' and 'Phosphorus' differ in meaning, but because they differ in "cognitive value," i.e., in associated qualitative descriptions. This, however, is precisely where Kripke's account requires filling in. For, if 'Hesperus' and 'Phosphorus' are both rigid designators, i.e., if they both serve simply to refer to objects, then it is difficult to fathom how (1) and (2) could express different propositions; no doubt, 'Hesperus' and 'Phosphorus' differ in "cognitive value," but what has yet to be explained is how this information can be conveyed via expressions whose role is simply to refer to objects.

Kripke has not, I think, provided us with a solution to Frege's problem. He has, however, provided us with the framework for a solution, one which, for its novelty and imaginativeness, for its ability to cast new light on old problems, is certainly worth pursuing.

Footnotes to Chapter 4

¹See, e.g., Bertrand Russell, "The Philosophy of Logical Atomism," rpt. in Logic and Knowledge, ed. Robert Charles Marsh (London: George Allen & Unwin, 1956), pp. 175-282.

²Alvin Plantinga, The Nature of Necessity (Oxford: Oxford University Press, 1974).

³Saul Kripke, "Identity and Necessity," in Identity and Individuation, ed. Milton K. Munitz (New York: NYU Press, 1971), p. 150. Hereafter cited as IN.

⁴Saul Kripke, "Naming and Necessity," in Semantics of Natural Language, ed. Gilbert Harman and Donald Davidson (Dordrecht: D. Reidel, 1972), p. 310. Hereafter cited as NN.

⁵Michael Dummett, Frege, Philosophy of Language (London: Duckworth, 1973), pp. 110-115.

⁶This argument is found in W.V.O. Quine, From a Logical Point of View, 2nd ed. rev. (Cambridge: Harvard University Press, 1961), pp. 155-56.

⁷Saul Kripke, "Semantical Considerations on Modal Logic," Acta Philosophica Fennica, Fasc. XVI (Helsinki, 1963), pp. 83-94.

⁸Arthur F. Smullyan, "Modality and Description," Journal of Symbolic Logic, 13 (1948), 31-37.

⁹Richard L. Cartwright, "Some Remarks on Essentialism," The Journal of Philosophy, LXV (1968), 617-618.

¹⁰See, e.g., W.V.O. Quine, "On What There Is," rpt. op. cit., pp. 1-19.

¹¹Ruth Barcan Marcus, "Modalities and Intensional Language," Boston Studies in the Philosophy of Science, I (New York: Humanities Press, 1963), pp. 71 ff.

¹²Plantinga, op. cit., pp. 72-87.

¹³Kaplan, op. cit., pp. 131-143.

¹⁴IN and NN.

¹⁵W.V.O. Quine, "Reply to Professor Marcus," rpt. in The Ways of Paradox and Other Essays (New York: Random House, 1966), pp. 179-180.

¹⁶NN, pp. 269-270.

¹⁷IN, pp. 148-149.

¹⁸NN, p. 343.

¹⁹IN, pp. 139-140.

²⁰Ibid., p. 140.

²¹Stating this precisely is problematic because it is not clear on Kripke's view (a) whether, e.g., 'Pegasus = Pegasus' expresses a necessary truth (let alone a truth), and (b) whether, e.g., 'Hesperus = Hesperus' expresses a truth in a possible world in which Hesperus does not exist. This precludes us from saying: '(λx)Fx' is a rigid designator iff every substitution instance of

$$\Box [(\lambda x)Fx]G(\lambda x)Fx \equiv [(\lambda x)Fx] \Box G(\lambda x)Fx,$$

where G is a schematic predicate letter, is true. This much, however, seems clear: if '(λx)Fx' is a rigid designator, then

$$(Ex) ((y) (Fy \equiv y=x) \cdot \Box (E!(\iota x)Fx \supset \dots x \dots))$$

and

$$\Box (E!(\iota x)Fx \supset (Ex) (y) ((Fy \equiv y=x) \cdot \dots x \dots))$$

have the same truth value.

²²To be sure, Kripke does symbolize (29) and (30) in the same way, namely, as ' $\Box(a=b)$ ', and this might lead one to suppose that (29) and (30) therefore express the same proposition. But, if Kripke had employed iota notation, he would, no doubt, symbolize the two sentences differently; I suppose he did not want to complicate matters unduly in the text.

²³Dummett, op. cit., p. 115.

²⁴Ibid., p. 115.

²⁵Ibid., p. 128.

²⁶Ibid., pp. 127-128.

²⁷NN, p. 278.

²⁸Dummett, op. cit., p. 125.

²⁹Ibid., pp. 127-128.

³⁰Ibid., p. 128.

³¹NN, pp. 266-267.

³²IN, p. 147.

³³NN, p. 267.

³⁴David K. Lewis, "Counterpart Theory and Quantified Modal Logic," The Journal of Philosophy, 65 (1968), 113-126.

³⁵Plantinga, op. cit., pp. 83-84.

³⁶Ibid., pp. 80-81.

³⁷Ibid., p. 87.

³⁸Ibid., pp. 84-85.

³⁹NN, p. 263.

⁴⁰NN, p. 354.

⁴¹NN, pp. 333-334.

⁴²NN, pp. 332-334.

⁴³Dummett, op. cit., p. 126.

⁴⁴Lewis, op. cit.

⁴⁵Dummett, op. cit., pp. 132-133.

⁴⁶Rudolf Carnap, Meaning and Necessity, 2nd ed.

(Chicago: University of Chicago Press, 1947).

⁴⁷David Kaplan, "Dthat," unpublished ms.

⁴⁸NN, p. 278.

⁴⁹NN, p. 264.

BIBLIOGRAPHY

- Binkley, R. "Quantifying, Quotation, and a Paradox."
Nôus, IV (1971), 271-278.
- Carnap, Rudolf. Meaning and Necessity. 2nd ed. Chicago:
 University of Chicago Press, 1947.
- Cartwright, Richard L. "Some Remarks on Essentialism."
The Journal of Philosophy, LXV (1968), 615-626.
- Church, Alonzo. "Carnap's Introduction to Semantics."
The Philosophical Review, LII (1943), 298-304.
- Church, Alonzo. "A Formulation of the Logic of Sense and
 Denotation." Structure, Method, and Meaning: Essays
 in Honor of Henry M. Sheffer. Ed. P. Henle, H. M.
 Kallen, and S. K. Langer. New York: Liberal Arts
 Press, 1951. 3-24.
- Dummett, Michael. "Note: Frege on Functions." The Philo-
 sophical Review, LXV (1956), 229-230.
- Dummett, Michael. Frege, Philosophy of Language. London:
 Duckworth, 1973.
- Dunn, M. and Belnap, N. "The Substitution Interpretation
 of the Quantifiers." Nôus, II (1968), 177-185.
- Frege, Gottlob. Begriffsschrift, eine der arithmetischen
 nachgebildete Formalsprache des reinen Denkens.
 Halle, 1879.

- Frege, Gottlob. Begriffsschrift und andere Aufsätze. Ed. Ignacio Angelelli. Hildesheim: Olms, 1964.
- Frege, Gottlob. "Begriffsschrift, a formula language, modeled upon that of arithmetic, for pure thought." Trans. Stefan Bauer-Mengelberg. From Frege to Gödel: A Sourcebook in Mathematical Logic, 1879-1931. Ed. Jean van Heijenhort. Cambridge: Harvard University Press, 1967. 1-82.
- Frege, Gottlob. "Begriffsschrift (Chapter I)." Trans. Peter Geach. Translations from the Philosophical Writings of Gottlob Frege. Ed. Peter Geach and Max Black. Oxford: Blackwell, 1960. 1-20.
- Frege, Gottlob. "Über den Zweck der Begriffsschrift." Jenaische Zeitschrift für Naturwissenschaft, Supplementar-Heft, 16 (1882-83), 1-10.
- Frege, Gottlob. "On The Purpose of the Begriffsschrift." Trans. H. V. Dudman. Australasian Journal of Philosophy, 46 (1968), 89-97.
- Frege, Gottlob. "Über die Wissenschaftliche Berechtigung einer Begriffsschrift." Zeitschrift für Philosophie und philosophische Kritik, 81 (1882), 106-114.
- Frege, Gottlob. "On the Scientific Justification of a Concept-Script." Trans. James M. Bartlett. Mind, LXXIII, 155-160.
- Frege, Gottlob. "Über Sinn und Bedeutung." Zeitschrift für Philosophie und philosophische Kritik, 100 (1892), 25-50.

- Frege, Gottlob. "On Sense and Reference." Trans. Max Black. Translations from the Philosophical Writings of Gottlob Frege. Ed. Peter Geach and Max Black. Oxford: Blackwell, 1960. 56-78.
- Frege, Gottlob. Funktion, Begriff, Bedeutung: Fünf Logische Studien. Ed. Günther Patzig. Göttingen: Vandenhoeck und Ruprecht, 1966.
- Frege, Gottlob. Translations from the Philosophical Writings of Gottlob Frege. Ed. Peter Geach and Max Black. Oxford: Blackwell, 1960.
- Frege, Gottlob. Die Grundlagen der Arithmetik, eine logisch-mathematische Untersuchung über den Begriff der Zahl. Breslau, 1884.
- Frege, Gottlob. The Foundations of Arithmetic. 2nd ed. rev. Trans. J. L. Austin. Oxford: Blackwell, 1953.
- Frege, Gottlob. "Illustrative Extracts from Frege's Review of Husserl's Philosophie der Arithmetik." Trans. Peter Geach. Translations from the Philosophical Writings of Gottlob Frege. Ed. Peter Geach and Max Black. Oxford: Blackwell, 1960. 79-85.
- Frege, Gottlob. Grundgesetze der Arithmetik. I (1893). II (1903). Rpt. Hildesheim: Olms, 1964.
- Frege, Gottlob. The Basic Laws of Arithmetic: Exposition of the System. Trans. and ed. Montgomery Furth. Berkeley and Los Angeles: University of California Press, 1967.

- Frege, Gottlob. "Frege Against the Formalists." Trans. Max Black. Translations from the Philosophical Writings of Gottlob Frege. Ed. Peter Geach and Max Black. Oxford: Blackwell, 1960). 182-233.
- Frege, Gottlob. "Der Gedanke. Eine Logische Untersuchung." Beiträge zur Philosophie des deutschen Idealismus, I (1918-19), 58-77.
- Frege, Gottlob. "The Thought." Trans. A. M. and Marcelle Quinton. Philosophical Logic. Ed. P. F. Strawson. Oxford: Oxford University Press, 1967. 17-38.
- Frege, Gottlob. "Logische Untersuchungen. Dritter Teil: Gedankengefüge." Beiträge zur Philosophie des deutschen Idealismus, III (1923), 36-51.
- Frege, Gottlob. "Compound Thoughts." Trans. R. H. Stoothoff. Mind, LXXII (1963), 1-17.
- Geach, Peter. Mental Acts. London: Routledge & Kegan Paul, 1957.
- Jackson, Howard. "Frege's Ontology." The Philosophical Review, LXIX (1960), 394-395.
- Kaplan, David. "Quantifying In." Words and Objections. Ed. Donald Davidson and Jaakko Hintikka. Dordrecht: D. Reidel, 1969. 178-214.
- Kaplan, David. "Dthat." Unpublished ms.
- Kneale, William and Kneale, Martha. The Development of Logic. Oxford: Oxford University Press, 1962.

- Kripke, Saul. "Semantical Considerations on Modal Logic." Acta Philosophica Fennica, Fasc. XVI (1963), 83-94.
- Kripke, Saul. "Identity and Necessity." Identity and Individuation. Ed. Milton K. Munitz. New York: NYU Press, 1971. 135-164.
- Kripke, Saul. "Naming and Necessity." Semantics of Natural Language. Ed. Gilbert Harman and Donald Davidson. Dordrecht: D. Reidel, 1972. 253-355.
- Lewis, David K. "Counterpart Theory and Quantified Modal Logic." The Journal of Philosophy, 65 (1968), 113-126.
- Linsky, Leonard. Referring. New York: Humanities Press, 1967.
- Marcus, Ruth Barcan. "Modalities and Intensional Language." Boston Studies in the Philosophy of Science, I. New York: Humanities Press, 1963.
- McKinsey, Michael. "Searle on Proper Names." The Philosophical Review, LXXX (1971), 220-229.
- Mill, John Stuart. Philosophy of Scientific Method. Ed. Ernest Nagel. New York: Hafner Publishing Co., 1950.
- Plantinga, Alvin. The Nature of Necessity. Oxford: Oxford University Press, 1974.
- Quine, W.V.O. From a Logical Point of View. 2nd ed. rev. Cambridge: Harvard University Press, 1961.
- Quine, W.V.O. Word and Object. Cambridge: MIT Press, 1960.

- Quine, W.V.O. The Ways of Paradox and Other Essays. New York: Random House, 1966.
- Russell, Bertrand. The Principles of Mathematics. 2nd ed. New York: Notron, 1938.
- Russell, Bertrand. "On Denoting." Logic and Knowledge. Ed. Robert Charles Marsh. London: George Allen & Unwin, 1956. 39-56.
- Russell, Bertrand. "The Philosophy of Logical Atomism." Logic and Knowledge. Ed. Robert Charles Marsh. London: George Allen & Unwin, 1956. 175-282.
- Russell, Bertrand. Introduction to Mathematical Philosophy. London: George Allen & Unwin, 1919.
- Searle, John R. "Russell's Objection to Frege's Theory of Sense and Reference." Analysis, XVIII (1957-58), 137-143.
- Searle, John R. "Proper Names." Mind, 67 (1958), 166-173.
- Searle, John R. Speech Acts. Cambridge: Cambridge University Press, 1969.
- Smullyan, Arthur F. "Modality and Description." Journal of Symbolic Logic, 13 (1948), 31-37.
- Tarski, A. "The Concept of Truth in Formalized Languages." Logic, Semantics, Metamathematics. Ed. J. H. Woodger. Oxford: Oxford University Press, 1956. 152-278.
- Whitehead, Alfred North and Russell, Bertrand. Principia Mathematica. 2nd ed. I. Cambridge: Cambridge University Press, 1927.

BIOGRAPHICAL NOTE

Richard Lloyd Mendelsohn was born March 15, 1944 in New York City. He attended Cornell University, receiving his B.A. degree cum Laude in Philosophy in June 1964. His graduate work was done at the Massachusetts Institute of Technology, where he was in residence from September 1964 to June 1968, and from September 1976 to February 1977. Since September 1968, he has been a Lecturer in the Department of Philosophy at Herbert H. Lehman College of The City University of New York.