THE LOGICAL PROBLEM OF IDENTITY

By

Keith A. Coleman

Submitted to the graduate program in Philosophy,
and to the Graduate Faculty of the University of Kansas
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

_____
chairperson

Committee members   _____

_____

_____

_____

Date defended: _____

The Dissertation Committee for Keith A. Coleman certifies
that this is the approved version of the following dissertation:

THE LOGICAL PROBLEM OF IDENTITY

Committee:

_____

Chairperson

_____

_____

_____

Date approved _____

Abstract

Keith A. Coleman
Department of Philosophy, February 2008
University of Kansas

A traditional problem concerning the meaning or logical content of statements of identity received its modern formulation in Gottlob Frege's "On Sense and Reference." Identity is taken either as a relation between objects or a relation between terms. If identity is interpreted as a relation between objects, then identity statements seem to be of little value since everything is clearly identical to itself. Assertions of identity are thought to convey significant information, but it is hard to see how they can on this interpretation. If identity is instead interpreted as a relation between terms, then identity statements still seem to be of little value since apparently they only convey a linguistic pronouncement to use certain terms interchangeably. Assertions of identity do not appear to be about the use of language, but, on this understanding of identity, they evidently are.

I examine the nature of the problem (and what it would take to solve it) and the advantages and disadvantages of each one of the two approaches to interpreting the content of identity statements. I then investigate two approaches for solving the problem from the perspective of identity as a relation between objects. The first of these represents the account provided by Gottlob Frege, and the second represents the account provided by Saul Kripke. I conclude that neither one of these accounts finally solves the problem of identity in its entirety. I then examine Michael Lockwood's approach to resolving the problem of identity based on the idea of identity as a relation between terms. I discuss and critically evaluate Lockwood's account together with a modified version of that account. After arguing for the inadequacy of the views examined as ultimate solutions to the problem of identity, I end by suggesting a strategy prompted by treating identity as indiscernibility.

Acknowledgements

I wish to thank my previous advisors, James Page and Anthony Genova for their assistance with my project. Without the early encouragement of Professor Genova and his later advice, my project would have been abandoned long ago. Professor Page worked with me on the early drafts of almost all the chapters. His dogged insistence that I make myself clear in the exposition resulted in tremendous improvement in the revisions to the manuscript. I also wish to thank Teresa Robertson for her hard work and her patience with me as she took on the daunting task of getting me to complete the dissertation. Without Professor Robertson's assistance, my project would certainly never have come to fruition. I should also like to thank the other members of my committee, Jack Bricke, Kevan Edwards, and Ed Wiley, who provided very valuable assistance in a number of ways. For the moral and intellectual support they provided over the years, I would also like to thank my friends Richard Botkin, Joe and Suzanne Collins, Christopher Foster, Jian Liu, Pinfei Lu, and Xiufen Lu.

Keith A. Coleman

To my mother,

Mary E. Coleman

And to the memory of my father,

Donald S. Coleman

Contents

Chapter 1: Worries about Identity

1.0   Introduction

Although it would be a brash overstatement to declare, as Bertrand Russell is reputed to have said, that all philosophy is logic or, as Ludwig Wittgenstein stated in the *Tractatus*, that all philosophy is a "critique of language," one of the lessons of analytic philosophy has been that logical matters are often at the heart of philosophical perplexities.[1] Few philosophical problems are reducible to logical problems, but many philosophical problems cannot be adequately resolved until certain key logical problems are solved.[2] If only because any proposed solution of a philosophical difficulty will in general not be recognized as truly resolving an issue unless the arguments offered in defense of the proposed resolution are acknowledged to be sound, the connection between philosophical and logical problems should come as no surprise. Such is the situation, I believe, with the numerous philosophical perplexities surrounding the notion of identity. The various stances that have been taken on what constitutes personal identity (i.e., the various accounts of when person **a** is identical to person **b**) and the identity of objects through time as well as the acceptance or rejection of either the hypothesis that mental events are just brain events or the hypothesis that the mental supervenes on the physical presuppose an adequate prior understanding of the logical nature of the identity predicate. It is to

---

1. The declaration attributed to Russell may an interpretation of a similar remark he made in *Our Knowledge of the External World*.
2. I maintain this in spite of some fairly recent criticism of this sort of thesis by Richard Mason in his book *Before Logic*.

provide that sort of understanding that the following discussion is aimed. Even though I will not specifically discuss the traditional nonlogical or extralogical problems associated with identity (such as the problems just alluded to concerning the identity of persons and of objects over time and the relation of the mental to the physical), I hope to shed some light, at least indirectly, upon the character of these and similar problems by examining the logical problem of identity. Furthermore, the manner in which identity statements are ultimately analyzed may reveal how we should correctly understand not only the logical notion of truth but also the nature of the entire metalanguage-object language distinction. There are some interesting parallels between certain difficulties concerning the ascription of truth and other difficulties having to do with the assertion of an identity. These difficulties may be interconnected in that an understanding of what it means to say that things are identical may require a prior understanding of what it means to say that something is true of an object. If such is indeed the case, then it will mean that the relation between our talk about objects and our talk about talk about objects may need to be reconsidered at a fundamental level.

1.1   Logical and nonlogical difficulties with identity

The logical problem of identity, or what I will subsequently refer to as simply "the problem of identity," has to do with the content of statements of identity. More specifically, the central concern of the problem is with that part of the content of identity statements that remains invariant regardless of the theoretical context in

which an identity statement occurs and is responsible for the validity of the most fundamental sorts of inferences involving identity statements. The contexts in which the substitutions of one term in an identity statement for the other are always judged to be plausible are typically considered to be just those contexts in which the basic inferences associated with identity statements are deemed plausible. At its core, the problem of identity does not involve that content of identity statements, of a specific sort, that is provided by a theory of individuation for a particular kind of thing. The familiar difficulties and consequences that attend the various proposed criteria for individuating such things as people, events, processes, psychological states, and physical objects undergoing change are not connected with the attempt to understand what justifies the substitution of co-referring terms. On the contrary, the alleged troubles with the criteria are generated from an application of the criteria to certain special cases by taking for granted the validity of inferences that involve only the substitution of co-referring terms and deriving problematic conclusions from identity statements viewed as the consequences of certain special hypotheses and the particular theory of individuation constituted by the criteria.

In an important sense, the logical problem of identity stems from uncertainties surrounding what is expressed by any identity statement, whereas the problem of individuating objects of various kinds (what may be called the "nonlogical problem of identity") arises from difficulties connected with the attempt to specify the precise conditions under which what is expressed by an identity statement should be accepted as true. The two problems (one concerning the logical content of identity statements

and the other concerning the nonlogical content of identity statements), though distinct, are, however, obviously interrelated in that one cannot solve the latter problem without presupposing a solution to the former problem. In other words, one cannot adequately advance a hypothesis concerning the identity and individuation of things of a certain kind unless one already knows those conditions that are independent of the particular subject matter of an identity claim and are always a part of the conditions sufficient and necessary for the truth of any identity statement. Unfortunately, the distinctions between the separate issues each problem raises are often muddled in discussions about identity, particularly, so it seems, in discussions about the identity of the indiscernibles and the indiscernibility of identicals.[3] In addition, more attention has apparently been devoted to the nonlogical problem of identity than to the logical problem of identity. Nevertheless, knowledge of when a statement of identity is true requires knowledge of both the requisite nonlogical conditions and the requisite logical conditions. Not knowing the logical conditions sufficient and necessary for the truth of an identity statement amounts to not knowing the first thing about identity and consequently not grasping the significance of the nonlogical conditions sufficient and necessary for the truth of identity statements of a particular sort.

The problem of identity (as opposed to the problem of individuation) received its modern formulation in Frege's "On Sense and Reference."[4] In that essay, Frege

---

3. See, for instance, Baruch A. Brody's, *Identity and Essence* and Peter Simons' entry "Identity of Indiscernibles" in *Routledge Encyclopedia of Philosophy*, 678-681.

4. This classical English rendition of Frege's essay, "*Uber Sinn und Bedeutung*," occurs in *Translations from the Philosophical Writings of Gottlob Frege* on pages 56-78.

asked whether identity is a relation and, if it is judged to be a relation, whether it is a relation between objects or a relation between names of objects.[5] Frege then noted that each of these two possible interpretations gives rise to certain perplexities. If identity is viewed as a relation between objects, then true statements of identity appear in an important sense necessary and trivial, given that anything is both necessarily and trivially identical to itself. If identity is viewed as a relation between the names of objects, then true statements of identity appear to be insignificant statements that are not about the world but are only about the use of language. Both interpretations are counterintuitive in that many identity statements are considered to be significant statements the truth of which, when recognized, contributes to our knowledge of the world. Since the questions Frege raised amount to questions about the very idea of identity, Frege's puzzle about identity did not concern uncertainties about what principles of individuation to adopt but was a puzzle about the very meaningfulness of identity statements.

In essence, then, the logical problem of identity amounts to the problem of interpreting what the assertion of any identity statement, regardless of its context, is normally an assertion about. Ultimately, it is the logical content of identity

---

5. The understanding of identity as a relation between objects must be kept distinct from the understanding of identity as a relation between names. Even though one could possibly hold that the appropriate relation holds between the names in a true identity statement only when a certain relation holds between the objects referred to (and vice versa), much confusion can result if it is not recognized that what identity statements are about is given different interpretations according to these two different understandings of the identity relation. Butchvarov, for instance, in his *Being Qua Being*, enters into a rather pointless discussion of the "apparent distinctness of identicals" by failing to appreciate the difference in kind between the distinctness or sameness of names and the distinctness or sameness of the objects to which the names refer. Contrary to Butchvarov, I do not see the apparent presumption of the distinctness of objects that he alleges is involved in the very use of distinct names. This may seem, and perhaps is, a minor point, but sometimes the elaboration and expansion on what is a small mistake ends up becoming a big mistake.

statements that is in question. It is clear that the identity predicate functions as a logical particle in many inferences, and the meaning attributed to the identity predicate when used as a logical word must be clarified in any proposed resolution to the problem of identity. The recognition of, among other things, whatever precisely we take statements of identity to be about, regardless of their specific theoretical underpinnings, should contribute to our understanding of the validity, and in some cases limitations, of the basic kinds of inferences customarily associated with identity statements. An understanding of the validity of such inferences amounts to an understanding of why and under what conditions the substitution of one co-referring term for another is legitimate. The legitimacy of such substitutions is not, in general, a consequence of the nonlogical content of particular identity statements. Rather, it is due to that component of the content of identity statements that is peculiar to all such statements. The problem of identity, as I shall more fully develop it in the next chapter, will therefore not be completely resolved until the interchangeability of co-referring terms is accounted for, since an explanation of the latter will necessarily figure in the solution to the former. Furthermore, because the intersubstitutivity of co-referring terms is due to the logical content of identity statements, an understanding of what identity statements are about sufficient to resolve the problem of identity will involve an understanding of the logical content of identity statements.

In very general terms, the logical content of a statement is that portion of the content of the statement that is considered to be responsible for the inferential connections between the statement and other statements. The logical content of

statements thus fixes the validity of arguments and determines what are the logical truths. The logical content of a sentence can be viewed as a consequence of the logical words it contains. Admittedly, the distinction between the logical and the nonlogical components of language is not always clear, and there is controversy regarding where precisely to draw the line.[6] Nevertheless, one should not conclude that, because the alleged distinction is not clear-cut and is in some sense arbitrary, there is in fact no distinction to be drawn. Such a conclusion would be no more warranted than the conclusion that, because, for any vague concept, the exact boundary between a case falling under the concept and a case falling outside of the concept is unclear and any particular line drawn always seems arbitrary, there is no distinction between the two cases. The distinction between the logical and the nonlogical is important. The development of formal theories of validity was prompted in effect by at least an implicit recognition of the distinction between the logical and the nonlogical components present in the premises and the conclusions of arguments. Moreover, the plausibility and the limitations of a formal theory of validity depend on the extent to which the logical behavior of all sentential particles that determine the truth conditions of, and therefore the inferential relations among, statements is taken into consideration. One of the limitations of traditional syllogistic logic, for example, stems from its difficulty, if not failure, in accounting for the

---

6. See, for instance, Steven T. Kuhn, "Logical Expressions, Constants, and Operator Logic"; McCarthy, "The Idea of a Logical Constant"; Christopher Peacocke, "What is a Logical Constant?"; and Roger Smook, "Logical and Extralogical Constants."

behavior of logical particles within the subject and predicate terms of categorical statements.

The general features of the logical particles can, I believe, be specified, and the specification will make it obvious that the identity predicate does function as a logical particle. I take it that it is reasonable to interpret a word or an expression as functioning as a logical word when it satisfies the following conditions.

(1) It does not belong to only the special lexicon in terms of which a theory is stated but is rather a part of the linguistic stock of the language in which the theory is expressed.

(2) It is an iterative device in the language that can be used to construct complex sentences out of all sorts of simpler sentential components.

(3) Its presence affects the truth conditions of complex sentences.

(4) Its occurrence in sentences determines, at least in part, the truth-value relations among different kinds of sentences.

(5) Its meaning is definable (and thus learnable) in terms of the truth-value relations that hold among sentences.

(6) Its meaning determines the validity of certain general types of inferences among sentences.

The nonlogical particles are, in contrast, those components of sentences the meanings of which do not determine the possible truth-values of sentences, and thus

the truth-value relations and inferential relations, among sentences. Nevertheless, the nonlogical particles generally fix the subject matter of statements and play an obvious role in fixing the actual truth-value of statements.

Thus, the problem of identity I wish to address will not be resolved until the role of the identity predicate as a logical particle is made clear. However, the clarity I seek does not involve settling the matter of so-called vague or indeterminate identity, a matter that seems to be of concern to some logicians nowadays. Issues connected with vagueness and with the problems vague concepts give rise to are important in themselves and can be, and I believe in most cases should be, handled as a separate concern. If there is vagueness to ascriptions of identity, that vagueness is not fundamentally different from the vagueness associated with the use of other terms and therefore can be dealt with as another extra matter concerning assertions of identity.[7]

1.2   The nature of the discussion to follow

Besides possessing merit in its own right, a resolution of the problem of identity is valuable in connection with other philosophical concerns, as I shall argue in the next chapter. For instance, Quine held that such referentially opaque contexts as belief contexts and other intensional contexts are problematic, and, as result, it is

---

7. Some have argued that identity statements cannot be indeterminate. Timothy Williamson (*Routledge Encyclopedia of Philosophy, s.v. "Identity"*) offers the following sort of argument. (I have interpreted and supplemented his argument somewhat.) Suppose that it is indeterminate whether a certain mass of rock and ice is Mt. Everest. It is, nevertheless, a determinate fact that Mt. Everest is Mt. Everest. If we assume for the sake of argument that it is true that the particular mass of rock and ice is indeed Mt. Everest, then (given the plausibility of substitution) it must be a determinant fact that the particular mass of rock and ice is Mt. Everest. However, this contradicts the original supposition. Thus, if the identity statement supposed to be indeterminate is in fact true, it is not indeterminately true (since it must in fact be false); hence, it is not indeterminately true after all.

difficult to talk sensibly about such things as beliefs, desires, and thoughts.[8] (Oddly enough, one should likewise conclude that talk about explanations is problematic since explanatory contexts are also referentially opaque.) Because substitutions into intensional contexts do not always preserve truth and particular intensions, if they exist at all, cannot really be specified since they lack clear identity conditions, Quine urged that talk of intensions be dispensed with and the underlying logic of science be regarded as purely extensional. Quine's rejection of intensions, and with that a rejection of intensional logic, is therefore founded on a notion of identity. So, an adequate understanding of identity is actually indispensable for judging whether or not Quine had a solid basis for his rejection of intensions.

I will now briefly outline what course my discussion of the relevant issues will take in the pages to follow. Subsequent to the present introductory chapter, I will give in the next chapter a more thorough presentation of both the problem of identity and what it will take to solve the problem. After that, in the next two chapters to follow, I will examine and evaluate more traditional approaches to solving the problem of identity. The primary goal of Chapter 3 will be an elaboration on and a critique of Frege's proposed solution, and the main intent of Chapter 4 will be to critically examine Kripke's solution to the problem. Both Frege's and Kripke's accounts of identity are objectual accounts. In Chapter 5, I will examine Lockwood's metalinguistic account of identity, which represents an attempt to solve the problem within a more general account of speech acts. Finally, in Chapter 6, I will summarize

---

8. See, for instance, Quine's essays "On What There Is," "Reference and Modality," and "Propositional Objects" as well as Chapters Four, Five, and Six of *Word and Object*.

the troubles with the sorts of solutions that have been proposed to resolve the problem of identity and give a general overview of a strategy that may possibly lead to a workable solution.

Chapter 2: The Nature of the Problem

2.0   Introduction

In the present chapter, I will discuss what I take to be the problem of identity. More specifically, I want to make clear in the next two sections what the problem of identity amounts to, what the significance of the problem is, and how two general approaches to understanding identity claims contribute to the problem. In the course of detailing the perplexities surrounding identity, I will specify the sorts of questions that will need to be addressed by any totally adequate solution to the problem of identity, and, by examining the two strategies often pursued in order to solve the problem, I will come to an assessment of what is required of an account of identity if it is to answer these questions.

2.1   The problem of identity

What I call "the problem of identity" concerns the manner in which we are to understand statements of identity.[1]   An adequate understanding of identity claims,

---

1.   Unless the context should indicate otherwise, whenever I employ such a phrase as "statement of identity," "identity statement," or "identity claim," what I will have in mind is any statement (occurring alone or as part of a longer statement) of the form $\ulcorner \alpha$ is identical to $\beta \urcorner$, $\ulcorner \alpha$ is the same as $\beta \urcorner$, or $\ulcorner \alpha = \beta \urcorner$, where α and β are any, not necessarily distinct, singular terms (i.e., proper names or possibly definite descriptions, depending upon the proposed analysis under consideration) or variables. The identity predicate can, of course, occur as a component of more complex sentences such as statements of uniqueness or other kinds of generalizations, but, as will emerge from my subsequent investigation, such statements are not what I will be focusing on primarily in discussing identity statements. However, the overall scope of the questions addressed and often the context of my discussions, particularly situations in which I do not want to beg any questions regarding an analysis, will force me to regard as an identity statement a statement that has an overall form such as $\ulcorner \alpha$ is identical to $\beta \urcorner$ and thus has at least the surface grammar of what I properly regard as an identity statement.

regardless of the context in which these claims occur, can only be had when three general sorts of questions about identity are answered. (1) What is the meaning, or import, of assertions of identity? In other words, what precisely are we indicating when we claim that **a** is identical to **b**? (2) If we regard identity, or sameness, as a relation, then must it be a relation that holds between objects, or a relation that holds between names or descriptions of objects? (3) Are identity statements actually, as opposed to just seemingly, informative? In asserting an identity statement are we, in the final analysis, merely indicating our decision to use a certain pair of terms interchangeably, or are we possibly informing our audience, in a practically useful manner, about the nature of objects in the world? Question 1 is what I shall call the "question of meaning"; question 2, the "question of reference"; and question 3, the "question of informational value." As I see it, each of these questions is a separate question, but the manner in which we answer any one of them will invariably make it more difficult to answer at least one of the others. For example, if, in our response to question 2, we regard identity as a relation that necessarily holds between every object and itself, and if we regard decisions about word usage as being merely a matter of convention, then, regardless of how specifically the question of meaning is answered, it is difficult to see how we could give an affirmative answer to the question of informational value, that is, how it is that identity statements could be used to convey knowledge about the world. These three questions are, surprisingly, not easy to answer, and for this reason reflection on such questions brings to light

certain conceptual problems associated with the notion of identity, a notion that, at least on the surface, seems so transparent.

Before discussing the attempts that have been made to answer the above questions, I need to clarify exactly what is at issue here. It is easy to get sidetracked, for the topic of the meaning of identity statements often raises issues which are actually peripheral and ought not to divert attention away from the main concern. As I indicated in the previous chapter, a specification of the content of any particular assertion of identity involves two separate, but related, concerns, and thus a careful distinction needs to be made between two different types of questions that are asked about the notion of identity. Our main concern is with the sort of questions that are asked in regard to an identity statement's logical content, that content that warrants the substitution of one term in an identity statement for another in transparent contexts and thus validates certain patterns of inference that are associated with statements of identity. As I noted in the previous chapter, the development of formal theories of deductive validity depends on the recognition that certain words (logical words) are put together with certain other words (nonlogical words) according to clearly definable patterns and the realization that the former words have a meaning that accounts for the inferential relations that obtain among sentences containing these words. The logical content of an identity statement is the content that remains invariant regardless of the theoretical context in which the identity statement occurs so long as the validity of the inference rules of identity elimination and identity introduction (or the natural language analogs of these rules) is preserved. Thus,

inquiries about the nature of, and the explanation for, the conditions in which it is permissible and the conditions in which it is not permissible to substitute for co-referring terms are questions that concern the logical content of identity statements.

The other sort of questions, those questions that are for the purposes of the present discussion peripheral, are asked in regard to the conditions under which objects of a certain kind are individuated. Depending upon the particular theory of the world in which identity statements occur (together with a particular understanding of identity statements), the specification of the conditions under which things are judged to be identical will vary. Thus, inquiries about what constitutes an individual or what constitutes more than one distinct individual are questions that concern primarily the nonlogical content of identity statements. The nonlogical content of identity statements is that component of the content of such statements that is provided by the contexts in which such statements occur and does not account for the inferential relations that are usually seen as holding among certain forms of identity statements and other statements regardless of context.

As I alluded to in the previous chapter, the questions that have been asked about what constitutes the identity of persons are questions that concern fundamentally the nonlogical meaning of the identity predicate as that predicate occurs in sentences that are used to make claims about persons. The traditional problem of personal identity did not arise in a context where the logical content of saying that person **a** is identical to person **b** was unknown, but instead the problem was connected with the metaphysical (and perhaps epistemological) difficulties

involved in adequately specifying when it is correct, given certain intuitions we may have about what constitutes being a person, to say that person **a** is identical to person **b**. The question of personal identity thus amounts to the question of the nonlogical content of the identity statement "Person **a** is identical to person **b**." Any proposed solution to the problem of personal identity will describe those conditions, unique to the theory of personhood under consideration, that must be satisfied in order for an assertion that person **a** is identical to person **b** to be correct, and such a solution must therefore presuppose an understanding of the logical content of any identity statement. Since an understanding of personal identity requires a prior understanding of the logical content of identity claims, the former is dependent on, even though distinct from, the latter.

Consistent with what was said in the previous chapter about the difference between the logical and the nonlogical words, the difference between the logical and the nonlogical content of identity statements is also reflected in a difference in the roles such statements play in making inferences. An identity statement can occur in a context where only the logical features of the identity predicate are exploited in the drawing of an inference. Any inference that depends on what gets expressed in a formal system as the rules of identity introduction (=I) and identity elimination (=E) constitutes an inference the plausibility of which reflects only the logical content of the identity predicate. In contrast, an identity statement can occur in a context where the identity predicate is being treated as a nonlogical particle in that the inferences

drawn are plausible only given the particular extralogical content of the particular identity statement.

The different sorts of inferences the typical set theorist makes from the statement that sets are identical reflect the difference between the logical and the nonlogical content of the identity predicate. The set theorist who, without appealing to an axiom governing the identity of sets, infers from some particular truth about **a** that there is a corresponding truth about **b** in virtue of the truth of the statement of identity "**a** = **b**" treats the identity predicate as a logical symbol, and thus the plausibility of the inference depends solely on the logical content of the identity statement. The set theorist can, however, appeal to the axiom of extensionality to infer the truth of the statement "**a** = **b**" from the fact that the elements of **a** are just the elements of **b**.[2] In this latter case, in contrast to the former, the plausibility of the inference depends crucially on the extralogical content of the identity statement as specified by the axiom since the conclusion is obtained as a deductive consequence of the axiom and not by an application of an inference rule governing identity. In this particular inference, the identity predicate "=" occurring in the axiom is being treated as any other predicate designating a binary relation, and the logical features of the identity predicate (those that collectively comprise the logical meaning of the identity predicate) do not play a role in reaching the conclusion. Even though the axiom of extensionality resembles a definition of the predicate "=", where the definiendum appears as the left side of the main biconditional and the definiens is the

---

2. The axiom of extensionality states that, for any sets x and y, x = y if and only if, for all z, z ∈ x if and only if z ∈ y.

generalization that appears as the right side of the main biconditional, the axiom is typically viewed as serving only the purpose of, in effect, defining the nonlogical content of the identity predicate. Clearly, within any development of set theory in which inferences regarding the identity of sets depend for their legitimacy on more than the logical content of identity statements and cannot be justified without an appeal to the axiom of extensionality, the axiom does not determine the logical content of the identity predicate.[3] Rather, in such circumstances, the axiom specifies the identity conditions for sets (by specifying the sufficient and necessary conditions under which, for sets, it will be said that set **a** is identical to set **b**); it does not unpack the logical content of identity statements, even if everything in the domain of the theory is regarded as a set. This is, however, no defect in the axiom since set theorists never intended it as a definition and the axiom usually appears in first order formalizations of set theory in a context where the understanding of the identity predicate as a logical symbol is already assumed. Without the treatment of the identity predicate as a logical particle in first order logic, the axiom of extensionality would not normally provide the criterion for the individuation of sets. This is the case since sets will be adequately individuated only when the inference rules governing identity are valid for the statements of set theory, which occurs only when the

---

3. For instance, any set theory in which "**a** = **c**" is not provable from "**a** = **b**" and "For all *x*, *x* is a member of **b** iff *x* is a member of **c**" by appealing only to logical laws and not also to the axiom of extensionality would be a set theory in which that axiom gave the nonlogical content of identity statements. In a version of set theory in which none of the inferences normally associated with identity could be made without an appeal to the axioms, the identity predicate could reasonably be viewed as a nonlogical word the content of which is fixed by the axioms. Such would be the case if the axioms of the set theory contained as the only primitive nonlogical particle a symbol for the membership relation and it followed that any set is a member of another set just in virtue of the membership of the former.

intersubstitutivity of co-referential terms is provided for, and this is not normally provided for whenever only the nonlogical features of identity statements are being exploited in drawing conclusions from identity statements, which occurs when the only meaning the identity predicate has is that provided by the axiom. Hence, sets are individuated only when the validity of the formalized language versions of =I and =E is provided for in the context of the set theory, and this typically occurs within a context where the logical meaning of the identity predicate is understood.

Questions regarding the logical content of identity statements are chiefly logical and semantical, since they deal with issues that concern the logical meaning of the identity predicate. Questions regarding the nonlogical content of identity statements tend right away to be metaphysical, for it is metaphysics, not logic, which provides the criteria for the identification and individuation of things that stand in the identity relation. The distinction is crucial, for the question of what an assertion of identity amounts to (from a logical perspective) is distinct from the question of when objects are the same (or when different appearances are appearances of the same object). It is only the former question that is of primary interest here and in the discussions to follow.

As I have indicated by the preceding remarks, in spite of the conceptual differences, an understanding of the logical and the nonlogical content of identity statements, and thus the separate bases for inferences from identity statements, are nonetheless interrelated. Any attempt to come to a complete understanding of the content of identity statements will involve both of two separate undertakings. One

must both come to an understanding of what is in general expressed by any identity statement (the logical content of an identity statement) and also come to an understanding of what commitments are signaled by identity statements in terms of the conditions, specific to the discipline in which the identity statements occur, under which the logical content, whatever is that content, of an identity statement is to be accepted. Assertions of identity and uniqueness are essential components of many different theories, and a complete understanding of the identity statements occurring in a particular theory will require one to be familiar with both the logical features of the identity predicate (those features that account for the inferences warranted by the natural language analogs of =I and =E) and the extralogical principles that, by specifying the circumstances in which an identity obtains, serve to individuate the objects discussed by the theory.[4] In order to understand the typical set theorist's use of the identity predicate, for instance, one must both understand the identity sign as a logical particle, as occurs when certain conclusions from the identity of sets do not depend on the axioms, and recognize how sets are individuated as specified by the axiom of extensionality. Any definition that specifies the extralogical conditions under which an assertion of identity should be accepted (and thus represents a specification of the nonlogical content of the identity predicate) will consist of a generalization stating that a logical statement of identity is true when and only when

---

4. Thus, the nonlogical content of an identity statement is that component of the statement's content that is provided by the theory, whereas the logical content of an identity statement is that component of the statement's content that is presupposed by the theory.

these conditions obtain.[5]   Thus, anyone who attempts to answer the metaphysical questions connected with the nonlogical content of the identity predicate must regard as unproblematic the logical content of the identity predicate since a specification of the nonlogical content of identity statements will invariably involve the at least implicit acknowledgement of the logical content of the identity predicate.   The problem of identity will, therefore, receive its more fundamental resolution when the problem is solved for the logical content of the identity predicate, and it is there that attention needs to be focused. What is desired here is to have answers to the purely logical and semantical questions that surround the notion of identity, for these questions are more basic, as I have just argued.

Now, how does the distinction, that has been duly noted, between these two notions of content affect the manner in which the original three questions (the question of meaning, the question of reference, and the question of informational value) are to be answered? The question of meaning needs to be addressed in terms of an explication of the meaning that attaches to the identity predicate when understood as a logical word. That is, an account needs to be given of the logical meaning of the identity predicate that explains the inferential relations that obtain between statements of identity and other statements, those relations being responsible for the validity of the natural language analogs of identity introduction and identity elimination. The question of reference will be answered when the reference of the

---

5. The axiom of extensionality in effect tells the set theorist that he should accept the logical consequences of claiming that set **a** = set **b** when and only when all elements of set **a** are elements of set **b** and vice versa.

terms (or the role of what appear to be terms) in any statement of identity is made clear at least with regard to the logical features of the identity predicate. The question of informational value will be settled when there is an elucidation of the general manner in which the truth-value of a statement containing the identity predicate understood as a logical symbol is invariably a reflection of certain extralogical considerations pertaining to matters of convention or to matters of empirical fact.

As I suggested in the previous chapter, there have emerged two principal strategies for resolving the questions about identity, Frege having mentioned each of these approaches as a distinct way of interpreting identity statements. Both of these ways of interpreting identity statements can be characterized in terms of the particular response given to the question of reference. The following discussion of these two approaches will help clarify the central issues involved in any attempt to answer the questions of meaning, reference, and informational value. As will be made clear in the discussion to follow, any successful reply to these questions must show how it is possible for there to be contingent statements of identity, and this latter can only be accomplished by an account of identity that satisfies three important requirements. Any account of identity that fails to satisfy any one of these requirements can then be rejected since it will fail to answer successfully the three basic questions.

2.2   The objectual and metalinguistic accounts of identity

Identity is often viewed either as a relation that holds between objects or as a relation that holds between names or descriptions of objects. An interpretation of

identity as a relation that holds among objects is what constitutes an objectual account of identity, and an interpretation of identity as a relation that holds among singular terms (proper names and definite descriptions) constitutes a metalinguistic account of identity.[6] Each one of these approaches seems to have its assets as well as its liabilities.

If identity is understood as a relation that holds between objects (or more precisely, between every object and itself), then, as Frege observed in "On Sense and Reference," every statement of identity appears to express a trivial truth since it is evidently a trivial matter that every object is identical to itself and to no other object.[7] The truth thus expressed by any true identity statement is what is often seen as a metaphysically necessary truth that is nonetheless trivial at least in the sense that it is trivially true that every thing is identical to itself. Furthermore, provided that different statements that are customarily used to make the same assertion are synonymous and the assertion made by the utterance of a statement is independent of the particular words occurring as singular terms in the statement but dependent on the referents of those terms, a true statement of identity of the form $\lceil \alpha = \beta \rceil$ appears to

6. Although Kripke may have originally coined the terms for these two interpretations of identity, I am borrowing the terms "the objectual account of identity" and "the metalinguistic account of identity" from Thomas V. Morris in his book *Understanding Identity Statements*. Morris used the former expression to refer to the first sort of interpretation of identity that I discuss and the latter to refer to the second sort of interpretation of identity that I discuss. Each of these two approaches is designed to give an analysis of the standard, or general, meaning of identity statements. One could, in principle, follow either approach and still admit that some identity statements in special contexts must be understood according to the alternative interpretation. By the way, Christopher Williams in his book *What is Identity?* rejected categorizing accounts of identity in such a fashion, but he did so, I believe, without adequate justification. Identity, if it is to be judged a relation, is clearly a two-place relation, and the terms that occupy those two places must both refer either to objects or to names of objects if ordinary identity statements are to be intelligible.

7. Gottlob Frege, *Translations from the Philosophical Writings of Gottlob Frege*, eds. Peter Geach and Max Black, 56-78.

have, regardless of what pair of co-referential singular terms flanks the identity sign, the same meaning as the corresponding identity statements of the forms $\ulcorner \alpha = \alpha \urcorner$ and $\ulcorner \beta = \beta \urcorner$.[8] However, this seems to be an intolerable consequence since such true statements as "The morning star is the evening star", "Benjamin Franklin was the first postmaster general of the United States", and "$2 = [8 - (4 + 4^{1/2})]^2 - 8^{1/3}$" are not self-evident but are such as to require verification through observation, investigation, or calculation. Also, each of these statements, with the possible exception of the last one, appears to express a claim about the world that seems not to be in any sense necessary, given an ordinary understanding of identity. In contrast, the statements "The morning star is the morning star," "Benjamin Franklin was Benjamin Franklin," and "$2 = 2$" all appear to be self-evident statements of necessity. How is it that each of the former seemingly contingent, nonobvious, statements has the same meaning as its corresponding self-evident necessary statement? According to what is suggested by the typical objectual account of identity, all true statements of identity are evidently necessary truths that ultimately tell us nothing significant about the world. However, this result stands in stark contrast to the usual understanding of the import of identity claims, for such assertions are often thought to convey at times important factual knowledge gained about the world.

---

8. This conclusion can be avoided by not granting these assumptions about meaning. It is possible to do this and still maintain an objectual analysis of identity, since one who opts for the objectual account of identity is not thereby committed to an objectual account of meaning. Frege, as I will argue later, abandoned the metalinguistic account of identity in favor of the objectual account, but he regarded the sense (meaning) of a sentence to be determined not by the reference of its terms but by the senses of its parts.

An objectual analysis does have certain points in its favor. It does allow for a rather straightforward specification of the truth conditions for an identity and in such a way that not all statements of identity are true. The identity statement "**a** = **b**" can be taken to be true when and only when "**a**" and "**b**" are co-referential. In addition, since identity claims are about objects in the world, such an analysis could in principle show how an assertion that **a** and **b** are identical conveys knowledge about the world if the problem of the apparent triviality of identity statements could be resolved. This is in contrast to any analysis of identity that would take identity statements to be only about singular terms and not about the referents of those terms, for in such a case identity statements would only be used to make assertions about language. Moreover, the typical sort of objectual account of identity does seem, at least initially, to lend plausibility to the inference rules governing identity. Since every object is identical to itself, any sentence of the form $\ulcorner \alpha = \alpha \urcorner$ should be true. Furthermore, it seems sensible to reason that, if "**a** = **b**" is true, then **a** and **b** are the same object so that a claim about **a** is also a claim about **b**, which means that the truth of any sentence that expresses a claim about **a** will entail the truth of the sentence that expresses the corresponding claim about **b** (and vice versa).

On the other hand, if identity is understood as a relation that holds between singular terms (or concerns only, as Frege put it, the "mode of designation" of its subject matter[9]), then statements of identity express, in essence, only a linguistic

---

9. "On Sense and Reference," *Translations from the Philosophical Writings of Gottlob Frege*, 57.

pronouncement to use certain pairs of terms interchangeably.[10] The agreement to use terms interchangeably may account for the validity of the inference rules governing identity (since one always uses a term interchangeably with itself and, presumably, the occurrence of one term in a sentence could be replaced by the occurrence of another if the two terms are being used interchangeably), but statements of identity only concern singular terms and not what those terms are about, what they designate, in spite of the fact that in any assertion of identity singular terms appear to be used instead of mentioned.[11] An identity statement only declares, in effect, that certain terms are being used to designate the same object, whatever that object happens to be.[12] According to a metalinguistic analysis of identity (that is, under an interpretation of identity as a relation that holds among singular terms), every true statement of identity thus represents what is in effect an acceptable stipulative definition, and every false statement of identity represents, presumably, a stipulation that cannot consistently be followed given the way the terms in the identity statement are used in other contexts. Such stipulations, though, appear to be arbitrarily producible and thus do not reflect knowledge about the world but only present to us

---

10. More precisely, they express in some way or another either a linguistic pronouncement to use terms in a certain way or a statement about how in fact speakers of a language are using terms.

11. In a formalized language, the form of a sentence reveals its logical content, but obviously this is not in general the case for natural language sentences. Conclusions about logical form should therefore not be made solely on the basis of the form of natural language sentences.

12. However, I want to avoid saying that under a metalinguistic account of identity the statement "**a** = **b**" means that the referent of "**a**" is the same as the referent of "**b**." Even though this way of giving the meaning of the identity statement does make mention of the names "**a**" and "**b**," the identity statement is not properly about those names but about the referents of those names if the content of "**a** = **b**" is specified in this manner. Also, this way of presenting a metalinguistic notion of identity makes such a notion more clearly dependent on a prior notion of identity holding between the referents of the terms in an identity statement.

conventions concerning language use.[13] Such is the case even if the stipulations of identity claims are, admittedly, established, or laid down, only after consulting the world, for as long as identity claims are not about the world but are only about language, they are not the vehicles by which knowledge about the world is, at least directly, communicated.[14] Therefore, in a sense, identity statements are just as trivial according to a metalinguistic account of identity as they are according to an objectual account of identity.

However, since the identity statement "**a** = **a**" declares, in effect, that the term "**a**" is to be used interchangeably with itself and the identity statement "**a** = **b**" declares, in effect, that the term "**a**" is to be used interchangeably with the term "**b**," the former statement of identity does differ in meaning from the latter statement of identity if identity is viewed as a relation holding among singular terms. After all, the former statement of identity expresses a claim only about the term "**a**," and the latter statement of identity expresses a claim about the two terms "**a**" and "**b**." It is for this reason that a metalinguistic interpretation does have an advantage over the typical objectual interpretation. In spite of this, the truth conditions of a statement of identity are not as obvious under a metalinguistic interpretation as they are under an objectual interpretation since, except for saying that the two terms are in fact being used

---

13. Statements of identity, on this view, are used either to introduce linguistic conventions or to talk about how language is being used. See footnote 10.

14. I say that statements of identity are established only after consulting the world in the sense that the terms occurring in identity statements also occur in a body of other statements that do make empirical claims and that must remain consistent upon the acceptance of an assertion of identity. Thus, we will accept the claim that the morning star is the evening star only if we are prepared to accept that all such empirically true statements about the morning star are also true statements about the evening star and vice versa. If the complex of empirical claims should prove to be inconsistent upon the acceptance of the statement of identity, then we would reject that statement.

interchangeably, it is not clear how one should specify the conditions under which two terms are always used interchangeably. Given a metalinguistic understanding of identity, the truth conditions for an identity statement will be about facts about linguistic usage, the most obvious of such facts being those the identity statement itself is in effect used to express. Thus, one who subscribes to the metalinguistic view would presumably claim that the statement of identity "**a** = **b**" is true when and only when the terms "**a**" and "**b**" are being used interchangeably in all purely referential (transparent) contexts. Clearly, though, the issue of circularity arises with such a specification since the statement of the truth conditions of a sentence is in this case virtually no different from what that sentence is actually used to assert. The circularity involved in stating the truth conditions may pose a distinct problem since one who opts for this approach will need to provide a precise and noncircular specification of a purely referential context into which these kinds of substitutions are always legitimate.[15]

The most perplexing aspect of the problem of identity, then, concerns the question how it is possible for there to be contingent, nontrivial statements of identity.[16] This is essentially the problem Frege saw with identity in "On Sense and

---

15. A metalinguistic identity theorist could not just say, for instance, that a purely referential context is one in which the substitution of "**a**" for "**b**" preserves the truth-value of a sentence as long as "**a** = **b**" is true. Incidentally, an objectual identity theorist is not necessarily immune to the difficulties attending the specification of a purely referential context, as will become evident in the discussion to follow, particularly if his account of identity depends crucially on notions, such as sameness of reference or sameness of sense, that occur in the metalanguage.

16. Although not the same problem, the problem of identity is similar in certain respects to another puzzle that has been called "the paradox of analysis." In any analytical definition, if the definiendum (the analysandum) and the definiens (the analysans) are synonymous, then the definition appears trivial; if they are not synonymous, then the definition seems inadequate. Given that the

Reference." Any attempt to successfully answer the questions of meaning, reference, and informational value must solve this problem. Any interpretation of identity statements that is successful at resolving this difficulty must at least show (1) how the identity statement "$a = a$" differs in content from the identity statement "$a = b$," (2) how it is that some identity statements are false (or how some identity statements are not necessarily true), and (3) how identity statements can be nontrivial in the sense that they can be used to convey information about the world. I shall call requirement 1 the "difference in content requirement," and I shall call requirements 2 and 3, respectively, the "nonnecessity requirement" and the "nontriviality requirement." (I should hasten to add here that any account of identity that completely meets these requirements will do so in a noncircular fashion in that it will not presuppose or rest upon a prior unanalyzed notion of identity.) According to an objectual account of identity, identity statements are about objects in the world but, when true, express what is often seen as a metaphysical necessity; according to a metalinguistic account of identity, identity statements are not necessary but express propositions that are about the use of terms instead of about the world. On the former interpretation, identity statements acquire their connection to the world (and thus could in principle be significant) but, when true, appear to lose their contingency; on the latter interpretation, they acquire their contingency but appear to lose their direct connection to the world. Thus, an objectual account partially meets the nonnecessity requirement (in that not all identity statements are true) and the nontriviality

---

definiendum and the definiens must be either synonymous or not, it evidently follows that every analytical definition is either trivial or inadequate.

requirement (only in that identity statements are about things in the world), while a metalinguistic account seems at best to meet only the difference in content requirement and the nonnecessity requirement. A metalinguistic account evidently fails to meet the nontriviality requirement in that the advocate of such an account interprets identity statements as statements that are only about the interchangeability of terms and not about what those terms designate.

What apparently is needed to solve the dilemma is some way to merge the best components of both views into a single interpretation that nonetheless avoids the worst components of both views. Such an interpretation would take identity statements to be about (nonlinguistic) objects, so that identity statements could be important statements about the world, but would also take identity statements to be in some manner about language, or the terms used in language, so that identity statements could be contingent statements even when true. In other words, it would be desirable to interpret true identity statements as statements about the world, as is suggested by the objectual view, without always acquiring a commitment to their necessity, and, at the same time, to interpret true identity statements as statements somehow connected with language use, as is suggested by the metalinguistic view, without severing the connection between identity statements and the world. The interpretation of the manner in which identity statements are about the use of language would, of course, have to dispel the pertinent sense of arbitrariness that is bestowed upon them by the latter view.

In the next chapter, I will begin my discussion of the proposed solutions to the problem of identity with a detailed examination of how Frege sought to resolve his puzzlement with identity. Both Frege's and Kripke's views on identity (the latter discussed in Chapter 4) are important if for no other reason than their explication helps to clarify further the problem of identity and the difficulties one faces in advancing an objectual account of identity.

Chapter 3: Frege's Solution

3.0   Introduction

In the present chapter, my focus will be on an examination and critique of Gottlob Frege's account of identity and a couple of more recent proposals that are within the spirit of Frege's analysis. I will examine in the next section Frege's solution, which in many respects represents the standard, or received, view of identity. I will then indicate in the subsequent section why I believe Frege ends up with an interpretation of identity statements that is not totally satisfactory. I will then close the chapter with a very brief look at two proposals that attempt to solve the problem of identity by in effect reintroducing Frege's sense/reference distinction.

Since Frege's views on identity have influenced the manner in which most philosophers of language today commonly think of identity, it seems only fitting that a critique of the traditional accounts of identity should begin with an examination of Frege's analysis.   In what follows, I will contend that, notwithstanding certain indications to the contrary, Frege ends up adopting an objectual account of identity, an objectual understanding that nonetheless attempts to meet the difference in content, nonnecessity, and nontriviality requirements by, in effect, providing for a distinction between metaphysical necessity and semantical and logical necessity. I believe that, in the final analysis, Frege settled on the view that the terms appearing in any statement of identity refer only to the objects that are their referents. Identity statements may appear trivial because, when true, they always contain reference to the same object so that from the standpoint of their referents they are, when true, necessarily true (in that it is metaphysically necessary that every object is identical to itself). Nonetheless, from the standpoint of their sense, identity statements are not

always necessarily true (and are thus not always trivial) in that a recognition of the sense of some identity statements does not always thereby constitute a recognition of the truth of such statements.

3.1   Frege's account of identity

In the *Begriffsschrift*, Frege clearly advances a metalinguistic interpretation of identity statements, but, later, after having made the distinction between sense and reference, he evidently changed his earlier views.[1] In section 8 of the *Begriffsschrift*, Frege characterizes identity as a relation that "applies to names and not to contents."[2] A is identical to B means that "*the sign A and the sign B have the same conceptual content, so that we can everywhere put B for A and conversely*."[3] For Frege, a sign (singular term) had a content, and to every sign there was associated a manner of determining a content.[4] Two different signs could both have the same content even though the two signs are associated with two different ways of determining a content. The fact that two different signs had the same content was not always obvious and might be a significant fact, so Frege felt compelled to introduce into his formal language the symbol for identity, which would be used to assert that the manner of determining a content associated with each of two signs determined the same content. Frege illustrated the need for an identity predicate with a geometrical example wherein two different ways of describing a point on the circumference of a circle

---

1. All citations to the *Begriffsschrift* (except for the citations to the original German text) are to the reprint in *Frege and Godel: Two Fundamental Texts in Mathematical Logic*, 1-82. (Italics are as they occur in that text.)

2. Ibid., 20.

3. Ibid., 21.

4. The association of a name with a determination of a content is what Frege evidently later saw as the expression by a name of its sense. What Frege meant by the manner of determining a content is perhaps best explained with the use of a mathematical analogy. Just as different mathematical descriptions of objects (numbers, points, angles, etc.) may reflect different mathematical procedures that can be used to pick out the referents of those descriptions, different names may reflect different ways the referents of those names can be determined or fixed.

were presented. The points determined by each of these descriptions were given two distinct names, and, when both descriptions corresponding to each of the names determined the same point, the content had by one of the names was the same content as that had by the other name. In such a case, we would correctly judge that the two ways of determining a content associated with the two names do in fact determine the same content. Frege declared, "Hence the need for a sign for identity of content rests upon the following consideration: the same content can be completely determined in different ways; but that in a particular case *two ways of determining it* really yield the *same result* is the content of a *judgment*."[5]

However, in the above-cited passage, the word "content" is apparently used in two different senses. Each occurrence of the word "content" in the passage is a translation of the same German word, "*Inhalt*," Frege used in the original, but the expressions "identity of content" and "same content" had a different meaning for Frege than did the expression "content of judgment."[6] In the geometrical example, the identity is (ultimately) between the point specified by the manner of determining a content associated with each name, and the two names have the same content when they both refer to the same point. Evidently, Frege used the terms "content" and "conceptual content" here to talk about what he later would call "reference," and, when he mentioned here a way of determining a content corresponding to a name, Frege had in mind what he would later call the "sense" of a name. However, in sections 2-4 of the *Begriffsschrift*, Frege presented a different notion of content when he described his notion of the content of a judgment.[7] In every judgment, there is both the expression of a content, or an idea or a thought, and the assertion that the

---

5. Ibid., 21.
6. See Gottlob Frege, *Begriffsschrift und Andere Aufsatze*, 1-88.
7. *Begriffsschrift*, in *Frege and Godel: Two Fundamental Texts*, 11-13.

content, or the idea or thought, is true. (For Frege, when one wants to deny that something is the case, one expresses its negation and then affirms that the latter is true.) Frege used, apparently interchangeably, in these sections some form of the German words "*Inhalt*," "*Vorstellung*," or "*Gedanke*" whenever he spoke of what could be expressed in a judgment.[8] These German words are frequently translated as, respectively, "content," "idea," and "thought." Not every content could be expressed in a judgment; Frege distinguished between contents that could become judgments from contents that could not become judgments. Given the examples Frege used to illustrate the difference, evidently the former are those contents that are complete thoughts and the latter are those contents that are not complete thoughts (or are ideas or concepts that are only associated with words and phrases and not with whole sentences). In section 3, Frege employed the expression "conceptual content" as a way of talking about the meaning synonymous sentences have in common. More precisely, two judgments (or propositions) were said to have the same conceptual content if and only if it is the case that, whenever both judgments each occur together with the same combination of certain other judgments, whatever follows from the first also follows from the second and vice versa. Frege used the same German words, "*begrifflichen Inhalt*," in section 3 as he used in section 8 to speak of what can be spoken of in English by the use of the words "conceptual content."[9] Frege's notion of content as it is associated with the phrases "content of a judgment" and "conceptual content" (as used in section 3) seems to be the notion of a thought or an idea, while Frege's notion of content as it is associated with the phrases "same content," "identity of content," and "conceptual content" (as used in section 8) seems to be the notion of

---

8. *Begriffsschrift und Andere Aufsatze*, 1-5.
9. Ibid., 2-4 and 13-15.

what Frege would later identify as an expression's reference.[10] The word "content" thus seems to undergo a shift in meaning from one place in the *Begriffsschrift* to another.[11] Perhaps the desire to avoid different uses of the same term played a part in forcing Frege later to make the distinction between sense and reference and to change his earlier analysis of identity.[12]

Frege began his famous essay "On Sense and Reference" by noting that the idea of equality gives rise to challenging questions that are not easy to answer.[13] Frege's initial discussion in that essay was concerned with one of the same questions with which I began my discussion: If identity is a relation, then must it be a relation that holds between objects, or a relation that holds between names or descriptions of objects? Frege attempted to solve the problem of identity by showing how the judgment represented by an identity statement of the form $\lceil \alpha = \beta \rceil$ could differ from the judgment represented by the corresponding identity statement of the form $\lceil \alpha = \alpha \rceil$. An essential component of Frege's analysis that allowed him to claim a difference in judgments was the distinction he made between sense and reference. The referent of

---

10. Admittedly, Frege used the expression "conceptual content" only once in section 8, in his definition of identity, but in section 8 and in other parts of the *Begriffsschrift* the way in which he spoke about two terms having the same content indicates that this expression was used in section 8 as just another way of talking simply about content.

11. After having distinguished between sense and reference, Frege later admitted on page x of his introduction to *The Basic Laws of Arithmetic* that he had previously used the expression "possible content of judgment" to mean either "thought" or "truth-value," the former being the sense of a declarative sentence and the latter being the reference of a declarative sentence.

12. Joan Weiner offers an interpretation of the situation that is somewhat different than mine. She maintains that Frege later abandoned his talk of "conceptual content" and recognized that included in the content of an expression is both its sense and its reference. Accompanying this change was, she contends, a change in Frege's understanding of identity: instead of viewing identity as a relation holding between expressions having the same sense and reference, Frege came to view identity as a relation holding between objects. According to Weiner, Frege continued the practice, introduced in his *Begriffsschrift*, of allowing sentential expressions to flank the identity predicate and was thus led to the view of a sentence as a name of its truth-value. On Weiner's interpretation, Frege held a view of identity in the *Begriffsschrift* that was, as she notes, implausible since sameness of content (or synonymy) is too strong a requirement for identity. See Chapters 5 and 6 in *Frege*, 72-116.

13. All references to "On Sense and Reference" in the following discussion are to the translation provided by Geach and Black appearing in *Translations from the Philosophical Writings of Gottlob Frege*, 56-78.

a singular term is the object the term refers to, the extension of the term, and the referent of a predicate is a concept or relation, which Frege understood technically to be a function that mapped one or more objects to a truth-value. The sense of an expression, wherein is contained what Frege called its "mode of presentation," corresponds closely to the contemporary notion of the intension of an expression (except that Frege conceived of the sense of an expression in terms of a function having the referent of the expression as its value). The sense of a singular term is grasped by anyone who is familiar with all of the true sentences in which the term appears, and the sense of a predicate is grasped by anyone who understands how the particular concept or relation referred to by the predicate is determined, or specified, by the predicate. It is in virtue of an expression having a sense that it also has a referent, although not every sensible expression has a referent. (The singular term "the largest prime number," for instance, has a sense but does not have a referent.) The relation between sense and reference is such that expressions with the same sense have the same reference but not all expressions with the same reference have the same sense. Frege subscribed to what can be called the principle of intersubstitutivity: expressions with the same sense can be substituted one for the other in the context of a sentence without affecting the sense of the sentence, and expressions with the same reference can be substituted one for the other in the context of a sentence without affecting the reference of the sentence. For example, if the word "bachelor" and the phrase "eligible unmarried man" are synonymous and the sense and referent of "Richard" is the same as, respectively, the sense and referent of "Dick," then the sentence "Richard is a bachelor" and the sentence "Dick is an eligible unmarried man" are synonymous and have the same referent (truth-value).

According to Frege's theory of sense, the sense of a sentence is a function of the senses of its component parts and how those parts are arranged. Since the sense of "**a**" can differ from the sense of "**b**" (after all, the senses or meanings of two singular terms don't have to coincide for them to be co-referential, but if they do coincide they are co-referential), the sense of "**a** = **b**" can differ from the sense of "**a** = **a**." The sense of a sentence was nontrivial for Frege if our mere recognition of that sense, our understanding alone of the sentence, is generally insufficient for us to determine the referent of the sentence (i.e., its truth-value).[14] Thus, the sense of "**a** = **b**" can be nontrivial, as occurs when the sense of "**a**" is entirely distinct from the sense of "**b**." So, since the sense of "the morning star" (the sense of "the brightest star seen in the early morning sky"[15]) is distinct from the sense of "the evening star" (the sense of "the brightest star seen in the early evening sky"), the sense of the sentence "The morning star is the evening star" is distinct from the sense of the sentence "The morning star is the morning star." Since a person's grasp of the sense of the latter sentence always suffices for him to recognize it to be true, while not everyone who grasps the sense of the former sentence thus recognizes it to be true, the sense of the latter is trivial, while the sense of the former is nontrivial.[16] It was in this fashion that Frege sought to satisfy the difference in content requirement.

---

14. In "On Sense and Reference," this is suggested by Frege's remarks in paragraphs 27 and 28 (pp. 57-58), paragraphs 32 and 33 (pp. 61-63), and paragraph 50 (p. 78).

15. A sense, like a set, is fairly easy to name but is very difficult to display. Frege characterized a sense as a thought or a part of a thought and conceived of it as a special type of function having a referent as its value. Since having the same sense is the basis for synonymy, in order to talk about the sense of an expression x (other than by using as a name the term ⌜the sense of x⌝) often the best that can be done is to find a synonymous expression y and to use, as a name of the sense of x, such terms as ⌜the thought that y⌝, ⌜the sense of y⌝, or ⌜the proposition that y⌝.

16. For Frege, any statement of the form ⌜$\alpha = \alpha$⌝ is an a priori truth (provided that it is a truth at all), but not all statements of the form ⌜$\alpha = \beta$⌝ are true (and not all of those that are true are a priori truths).

For Frege, the reference of complex expressions is fixed in a manner analogous to the manner in which the sense is fixed. According to Frege's theory of reference, the reference of a sentence, its truth-value, is a function of the references of its component parts. A statement of identity is true whenever both of its component singular terms have the same referent and is false otherwise. Thus, if the terms "**a**" and "**b**" refer at all, then "**a** = **a**" will always be true regardless of what "**a**" refers to and "**a** = **b**" will be true as long as the referent of "**a**" is the same as the referent of "**b**." However, since the referent of "**a**" is the referent of "**a**," but the referent of "**a**" is not necessarily the referent of "**b**," the former statement is logically necessary, in that all such statements having a truth-value are true, while the latter statement is logically contingent, in that some such statements having a truth-value are false.[17] Furthermore, "**a** = **a**" is (when true) always semantically necessary, in that our recognition of the sense of any such sentence is always sufficient for our recognizing that the sentence is true (when "**a**" refers at all), while "**a** = **b**" is not always semantically necessary, let alone true. Thus, Frege's analysis appears to meet the nonnecessity requirement.

Frege understood a judgment to be the movement from the sense of a sentence to its reference, its truth-value.[18] Since a sentence of the form $\ulcorner \alpha = \beta \urcorner$ can have a sense different from the sense of the corresponding sentence of the form $\ulcorner \alpha = \alpha \urcorner$ and sentences of the former kind can, in general, be either true or false but sentences of the latter kind cannot be false, different judgments are possible. The nontriviality (that is, the presence of informative value) of some identity statements is due to the fact that they possess a nontrivial sense and do not have to be true, given Frege's

---

17. The sentence "**a** = **b**" is logically contingent provided, of course, that it is not logically necessary that the terms "**a**" and "**b**" be co-referring.

18. Frege makes this claim at the end of "On Sense and Reference" (paragraph 50, p. 78).

theory of reference. A difference in judgments thus explains why the assertion that **a** = **b** may be cognitively significant, whereas the assertion that **a** = **a** is without informative value, even if true. The assertion of an identity statement with a nontrivial sense could inform us about the world because it could tell us something about its referents that we would not obtain by merely recognizing the sense of the sentence. Notwithstanding the fact that true statements of identity express metaphysically necessary truths, true statements of identity are not always logically necessary and are not semantically necessary unless they possess a sense that always enables one to pass directly from that sense to a truth-value (the True). If Frege's analysis does indeed satisfy the difference in content and nonnecessity requirements, then different judgments are possible and thus Frege's analysis also satisfies the nontriviality requirement.

In sum, then, Frege's theory of sense together with his theory of reference allowed Frege to give an account of identity statements that apparently satisfies the difference in content, nonnecessity, and nontriviality requirements by, in effect, providing for a distinction between different notions of necessity. Although Frege did not explicitly identify these different forms of necessity, they can be recognized as metaphysical, logical, and semantical necessity. Any instance of the form $\ulcorner \alpha = \alpha \urcorner$ is, if true, both logically necessary, since no instance of that form is false (i.e., the reference of no instance of that form is the False), and semantically necessary, since the terms flanking the identity sign have the same sense and thus our grasp of the sense of the sentence is all that we need in order to know that it is true. Any instance of the form $\ulcorner \alpha = \beta \urcorner$ is, if true, metaphysically necessary, since every object stands in the identity relationship with itself (i.e., the relation referred to by the identity predicate maps any object and itself to the truth-value the True), but some instances

of that form are not semantically necessary, since the sense of α may be distinct from the sense of β and thus our recognition of the sense of the sentence may not be sufficient for us to recognize its reference. Also, an instance of ⌜α = β⌝ may not be logically necessary because some statements having the same form as that instance are not true (i.e., the reference of some such statements is the False). Hence, even though all true identity statements are metaphysically necessary, the statement "**a** = **a**" may have a different content than "**a** = **b**" since the former, if true, is logically and semantically necessary, and thus is trivial, and the latter, even if true, may not be logically or semantically necessary, and thus may not be trivial.

3.2   The trouble with Frege's solution

Frege evidently thought that the clearest indication for the need to make the distinction between sense and reference was provided by the problem of identity. It was in virtue of this distinction that the problem could supposedly be solved, and without such a distinction the problem would remain. The only trouble with this proposed solution, as it is presented in "On Sense and Reference," is that it really does not completely answer the questions concerning identity Frege raised at the beginning of that essay. Is identity a relation that holds between any object and itself or is it a relation that holds between names of objects? What precisely are we indicating when we claim that **a** is identical to **b**? It is not quite clear in "On Sense and Reference" how Frege answers the first question, particularly since an advocacy of both interpretations of identity is suggested by different parts of his initial discussion. However, I do believe that, rather than adopting the view that statements of identity express a relationship between terms that holds whenever the senses of the terms determine the same reference, Frege took the simpler approach and supposed

that such terms refer to their customary referents so that the identity statement is true whenever the referents of the terms are the same. If Frege had advocated the former view after having made the distinction between sense and reference, then, given his definition of identity in section 8 of the *Begriffsschrift*, he would have been claiming that "**a** = **b**" means that the terms "**a**" and "**b**" are so related that the referent of "**a**" determined by the sense of "**a**" is the same as the referent of "**b**" determined by the sense of "**b**." Frege would then have understood identity as a relation, either simple or complex, between the terms in an identity statement and the referents of those terms.

The claim that Frege finally opted for the objectual, not the metalinguistic, interpretation of identity can be defended along two lines of thought. First of all, since terms with the same sense have the same reference and the terms in an identity statement refer not to themselves but to their customary referents, an objectual account would allow the replacement of one term in an identity statement with another term having the same sense to preserve the sense of the original. If an identity statement asserted a relation between singular terms, then the singular terms occurring in a statement of identity would presumably be referring to themselves.[19] However, if one such term in an identity statement were to be replaced by another term with the same sense, there would be no good reason to believe that the sense of the original statement would be preserved (unless, of course, one opts for the curious view that the names of two singular terms will always have the same sense provided that the two singular terms have the same sense).[20] In addition, when Frege

---

19. Frege acknowledges this in section 8 of the *Begriffsschrift.* Notice that, in his statement of the meaning of the identity statement "A = B," the signs "A" and "B" are being referred to even though the original identity statement uses the terms "A" and "B" themselves (presumably to refer to objects).

20. It would seem odd to say, for example, that the sense of the quoted expression "'the morning star'" is also the sense of the quoted expression "'the brightest star seen in the early morning

considered in "On Sense and Reference" the cases in which the principle of substitutivity appears to fail but doesn't because words do not have their customary reference, he never mentioned a case where a singular term refers both to itself and to its customary referent.[21] Frege noted that the words in direct quotation refer to the quoted sentence and that the words in some subordinate clauses refer both to their customary sense and to their customary reference, and this is why in these cases substitutions of co-referring terms that fail to preserve truth-value do not violate the principle of substitutivity. As one of Frege's examples of the latter phenomenon, the replacement of a term by a co-referring term in the sentence "Bebel mistakenly supposes that the return of Alsace-Lorraine would appease France's desire for revenge" may not result in a sentence with the same truth-value since the subordinate clause refers to both the sense and the truth-value of the embedded sentence.[22] Among the various other possible sentential constructions present in a language, a situation where features of these two cases are combined is never described in "On Sense and Reference." This would seem to be a significant omission, especially since identity was the subject with which he began the discussion in that article, if he still held to his earlier views on identity.

The second sort of reason why it can be argued that Frege later abandoned the metalinguistic account of identity concerns how he treats identity in the *Grundgesetze*, which was published after both the *Begriffsschrift* and "On Sense and Reference." In the *Grundgesetze*, Frege regards identity as a relation that is expressed by a functional expression having two argument-places.[23] Thus, Frege

---

sky'" in spite of the fact that the sense of the expression "the morning star" is also the sense of the expression "the brightest star seen in the early morning sky."

21. "On Sense and Reference," paragraphs 36-49, pp. 65-78.

22. Ibid., 76.

23. Sections of Volume I and Volume II of the *Grundgesetze* are reprinted in *Translations from the Philosophical Writings of Gottlob Frege*, 137-244. See especially sections 1-5 of Volume I of the *Grundgesetze*, 152-157.

used "$\xi = \zeta$" to refer to the function referred to in any identity statement, where what stands on the left side of the identity sign denotes the ξ–argument of the function and what stands on the right side of the identity sign denotes the ζ-argument of the function.[24]   A relation is a function that maps objects onto truth-values.[25]   Since concepts and relations are referred to by predicates, the "is" of identity must be a two-place predicate, in spite of the fact that it is a logical particle. Since an argument of a function is an object, identity is a relation that holds between objects, and whatever constructions occupy the two positions in the predicate must be proper names.[26] This latter fact accords with the manner in which Frege actually uses the identity sign to form sentences in the *Grundgesetze*. Frege thus used both "Δ" and "Γ" as names and specified the truth conditions for identity statements by stipulating that "$\Delta = \Gamma$" is to stand for the True whenever Δ and Γ are the same.[27]   Hence, where "**a**" and "**b**" are proper names and "**F**" is a two-place predicate, "**a** = **b**" is analogous in logical form to "**Fab**." Frege is therefore committed to the view that identity is a relation between objects, not a relation between names of objects.

Notwithstanding the difficulties involved in interpreting just how Frege would, in the end, answer the first question, Frege's solution clearly does not provide an answer to the second question: What precisely are we indicating when we claim that **a** = **b**? If we were puzzled about the meaning of such a statement as "**a** = **b**," it would not really help us to be told that such a statement means that the referent of "**a**" is the same as the referent of "**b**" or that there is a unique x (i.e., any y satisfying the same condition being the same as x) such that "**a**" refers to x and "**b**" refers to x. Similarly, it would not help to say (more in line with Frege's analysis) that the

---

24. Ibid., section 7, p. 158.
25. Ibid., section 3, p. 155.
26. Ibid., section 2, p. 154.
27. Ibid., section 7, p. 158.

statement "**a** = **b**" is true if and only if the referent of "**a**" is the same as the referent of "**b**." After all, the predicate, "is the same as," occurring in these parsed versions of the identity statement "**a** = **b**" is just another expression for the identity predicate. Thus, such metalanguage statements expressing the content of assertions of identity and such statements of the truth conditions for identities are no less obscure than are the identity statements themselves. Furthermore, we would, in most cases, never prefer the former over the latter since the former are decidedly more complex and still contain the problematic identity predicate.[28]

The trouble with Frege's solution to the problem of identity is that almost all the difficulties that attend the problem, even if solved for statements in the object language, are reintroduced in the metalanguage. In order to answer the question what it means for **a** to be identical to **b**, one must understand the truth conditions for the sentence "**a** is identical to **b**" and must therefore answer the question what it means for the referent of "**a**" to be the same as (identical to) the referent of "**b**." In order to answer the latter question, one must, in turn, rely upon a notion of identity occurring in the metalanguage wherein is contained the descriptions "the referent of '**a**'" and "the referent of '**b**'." The problem is thus never solved but is rather merely forestalled. This situation is objectionable since it means that we cannot rely on the metalanguage to employ a notion in terms of which identity can be usefully defined. Nonetheless, if we already have an adequate conception of the logical content of identity statements in the metalanguage, then that understanding can be usefully exploited in the development of an analysis of identity that will indeed satisfy the

---

28. I suspect Morris (*Understanding Identity Statements*) was led to understand Frege as espousing a metalinguistic analysis of identity by mistakenly interpreting Frege's statement of the truth conditions for an identity statement as being synonymous to the identity statement and wrongly conceiving of the former as a statement about the relationship between terms instead of the referents of terms. However, one should not take a statement to be about its terms just because the statement of its truth conditions makes mention of its terms (which was the overall mistake Morris made), otherwise every statement would be about its terms instead of the referents of its terms.

difference in content, the nonnecessity, and the nontriviality requirements and will answer the questions asked at the beginning of this chapter. Since the specification of the truth conditions for identity statements accounted for the fact that "**a** = **a**" is (logically) necessary when true and "**a** = **b**" is not, Frege's analysis relied fundamentally on a conception of identity in the metalanguage, as well as the notion of reference, in order to satisfy the nonnecessity requirement. Since differences in sense contributed to the fact that "**a** = **a**" is different in content than "**a** = **b**," Frege relied fundamentally on the theory of sense in order to satisfy the difference in content requirement. (Furthermore, since the difference in content was attributed to a difference in the senses of "**a**" and "**b**," Frege relied on a notion of two terms having different senses, which would presumably be dependent on a notion of two terms having the same sense.) Since differences in both sense and truth conditions accounted for the fact that "**a** = **a**" is trivial and "**a** = **b**" is not, Frege relied fundamentally on both the theory of sense and the theory of reference in order to satisfy the nontriviality requirement.

The fact that the identity predicate still occurs in the metalanguage statements and was never eliminated from the corresponding talk about senses and referents shows that the meaning of an identity claim has nothing essentially to do with the distinction between sense and reference. We should expect that if identity were to be expressible as some sort of relation between sense and reference, then the identity predicate as such would drop out of our considerations in the metalanguage, but such is not the case. Frege only succeeded in specifying the conditions under which an identity claim, which, when true, is to be understood as a claim about, in effect, a single object, could represent an important, nontrivial extension of our knowledge given a prior notion of identity in the metalanguage. Frege did not fully explicate the

meaning of the identity predicate, but, in all fairness to Frege, it should be remarked that this probably was never really his goal. Frege even argued that identity could not be defined. "Since any definition is an identity," Frege declared, "identity itself cannot be defined."[29]

Some may object that, since Frege did not make the metalanguage/object language distinction, it is unfair to criticize him in terms of machinery he was in no position to handle. However, such an objection is quite irrelevant.[30] The making of the metalanguage/object language distinction does not create the trouble with Frege's analysis. The problem would remain whether or not we explicitly recognized the difference between the language used and the language mentioned. The trouble may actually be more intuitively clear if we ignore the difference. Frege employed the very notion of identity in his presentation of the truth conditions for identity statements, and, if grasping the sense of the identity predicate is a necessary condition for understanding the truth conditions for an identity statement, then Frege presupposed a prior grasp of the sense of the identity predicate in his account of the truth conditions for an identity statement and thus relied on the very notion his analysis was ostensibly designed to explicate. Frege used discourse to talk about discourse, and this is true regardless of how his analysis is described. By making the metalanguage/object language distinction, we only clarify and elaborate upon what is already present in his analysis.

A more telling objection to my dismissal of Frege's analysis can nonetheless be made. It may be claimed that identity must be viewed as a primitive notion and as such is ineliminable. We should not expect to be able ultimately to define the logical

---

29. See Frege's review of Husserl's *Philosophie der Arithmetik*, reprinted, in part, in *Translations from the Philosophical Writings of Gottlob Frege*, 79-85.
30. Such an objection may also be quite mistaken as well. Given his criticisms of the works of some of the mathematicians of his day, Frege seemed to recognize the distinction between discourse about objects and discourse about the signs used to designate objects.

notion of identity, and, consequently, the best that can be achieved is to employ whatever resources are available in the metalanguage in order to specify the truth conditions of identity statements and to establish certain rules of inference that govern the way statements of identity operate in deductions. The identity predicate is, after all, a logical particle, and logical particles tend not to be completely eliminable. No one would object to the use of the word "and" to form conjunctions in spite of the fact that the metalanguage statement of the truth conditions for conjunctions standardly employs the very notion of conjunction: A sentence of the form ⌜A & B⌝ is true iff both the left and the right conjuncts are true. We should not be surprised then that the truth conditions for an identity statement cannot be stated without employing the very notion of identity.

However, the situation with the identity predicate is different than it is with other logical particles. First of all, the other logical particles are to some extent interdefinable, and this is not the case with identity. A universal generalization can be defined as the negation of existential generalization, and any one of the truth functional connectives can be defined in terms of one or more other truth functional connectives. The identity predicate, according to the standard account, is not definable in terms of any other logical particle. Secondly, and more importantly, the truth conditions of other logical particles are understood in the context of a more general account of how such particles determine the truth-value of sentences in which they occur. The truth conditions for universal and existential generalizations are stated in the context of the theory of quantification, and the truth conditions for negations, conjunctions, disjunctions, and conditionals are understood in terms of the theory of truth functions. It is in virtue of the latter theory that deviant conjunctions, disjunctions, and conditionals can be recognized in spite of the fact that the

conjunctive, disjunctive, and conditional particles are all taken as primitive. However, as should be evident from my discussion, Frege's account of identity provides no theory in the context of which the specification of the truth conditions for identity statements can help us understand the logical meaning of the identity predicate. The specification of the truth conditions for identity statements does occur in the context of Frege's theory of reference, but this theory is too general since it applies to both the logical and the nonlogical parts of the language and does not single out the strictly logical features of the identity predicate. The manner in which a logical predicate and a nonlogical predicate determine truth-value is accounted for in the same fashion by the theory of reference.

There is, though, a much more profound reason for not merely taking identity as primitive and letting it go at that. The distinction between an intensional and an extensional context, and thus the distinction between intensional and extensional logic, is made based upon a notion of identity.[31] If we lack a clear understanding of the relation of identity, then our idea of the distinction between intension and extension must ultimately be confused. The intensional idioms that introduce propositional attitude contexts such as "believes that," "desires that," "wishes that," "hopes that," "fears that," "says that," and "wonders whether" vary greatly in meaning and for this reason are difficult to categorize in terms of a common content, but they all share the feature of often being used to set up contexts within which substitutions of co-referring terms do not in general preserve truth-value. In a typical sentence in which one of these idioms is present, a sentential expression occurs in the subordinate clause following the idiom, and this sentential expression may contain one or more singular terms. The truth-value of the complex sentence will not always remain

---

31. Also, as I noted in the previous chapter, Quine presupposed an adequate understanding of identity in his rejection of intensions.

unchanged when a singular term occurring in the subordinate clause is replaced by another co-referring term. For example, even though Mark Twain is, or was, Samuel Clemens, the following sentences do not necessarily have the same truth-value.

Joe believes that Mark Twain wrote *The Adventures of Tom Sawyer and Huckleberry Finn*.

Joe believes that Samuel Clemens wrote *The Adventures of Tom Sawyer and Huckleberry Finn*.

Joe, for instance, may have an extensive knowledge of American literature but not know who Samuel Clemens was.

As Frege discussed at length in "On Sense and Reference," the principle of substitutivity appears to be violated by substitutions into these sorts of contexts.[32] Because the replacement in these contexts of one co-referring term for another may not preserve the truth-value of the sentence (the Fregean referent of the sentence) propositional attitude contexts are often said to be "referentially opaque."[33] In contrast, contexts in which co-referring terms can be freely substituted one for the other without altering truth-value are said to be "referentially transparent" or "purely referential." There are also those more problematic cases involving causal and evidential relations and explanatory contexts where transparent reference to objects

---

32. In addition, there are also modal contexts and contexts involving direct quotation where the principle of substitutivity likewise apparently fails. As I shall discuss in the next chapter, substitution into modal contexts can generate identity paradoxes.

33. This expression was coined by Quine in *Word and Object*. As I have previously mentioned, for Frege there are no violations of the principle of substitutivity; all alleged violations of the principle are only apparent violations. Instead of singular terms in subordinate clauses and in quoted sentences occupying referentially opaque positions, such terms, for Frege, involve "oblique reference" (i.e., they refer not to their customary referents, those referents the terms would normally have outside of these special contexts, but rather to their indirect referents, which are their customary senses).

seems to be made, and yet substitutions of co-referring terms are typically not permitted and, in the main, should not be permitted. We may, for instance, explain why bees are attracted to a certain flower by claiming that the bees are attracted to the flower because the color of the flower is blue. Notwithstanding the truth of the sentence "The color of the lamp signaling a blue light special is blue," we would not preserve the explanation in claiming that the bees are attracted to the flower because the color of the flower is the color of the lamp signaling a blue light special. The situation is complicated further by the fact that there are, however, special cases where some substitutions into these same sorts of contexts do seem to be legitimate. Thus, from the sentences "Jim provided evidence which indicated that Tom is guilty" and "Tom is Laura's husband" it seems to follow that "Jim provided evidence which indicated that Laura's husband is guilty." An understanding of the logical nature of propositional attitude and explanatory contexts will therefore require comprehending the extent of their opacity (or what I like to call the "degree of their opacity") and not merely the fact of their opacity.[34]

---

34. Frege's examination in "On Sense and Reference" of the cases in which the principle of substitutivity seemingly fails does not provide us with an adequate understanding of a purely referential context. Frege was primarily concerned there with defending his account of sense and reference against possible criticism, and so what he did was to examine apparent counterexamples and show that, in each case, the principle can be maintained as long as we take terms and component sentences as sometimes having a sense or reference that is not their customary sense or reference. Frege's remarks, though, do not constitute an adequate theory of transparent and opaque contexts. (Actually, contexts were never opaque for Frege, since intersubstitutivity never really failed according to Frege, but terms could have an indirect or oblique reference, instead of their customary reference, in certain contexts.) Frege never provided sufficient reason for claiming, for instance, that terms could have their customary sense as their indirect reference other than the fact that the principle of substitutivity is preserved if we sometimes take this to be the case. Thus, Frege characterized contexts in terms of what expressions with the same customary sense or reference could be substituted for corresponding expressions without altering the truth-value of sentences. Frege provided, therefore, no descriptions of contexts that were free of the notion of sameness (either sameness of customary sense or reference or sameness of indirect sense or reference). In addition, although I will not argue for this here, Frege's contention that the reference of sentences in propositional attitude contexts is to their customary sense is probably not in general true. A substitution of one term for another in such contexts should actually be legitimate provided that the substitution preserves the same saying, thought, belief, desire, etc. expressed by the original subordinate clause, rather than merely the same sense or proposition.

The apparent failure of intersubstitutivity has been taken as one of the tests for an intensional context. If a sentence contains a term t_1 and is such that the substitution for t_1 by a co-referring term t_2 can alter the truth-value of the sentence, then the context in which t_1 occurs is intensional, and, if this is not the case, then the context is extensional. Furthermore, the terms t_1 and t_2 are co-referring when and only when the identity statement $\lceil t_1 = t_2 \rceil$ is true. The notion of identity thus plays a crucial role in demarcating the boundary between extensional and intensional logic; the principle of substitutivity will always be understood to hold if the underlying logic of a theory is taken to be extensional, but this will not be the case if that logic is taken to be intensional. Within formal systems of extensional logic, the validity of the inference rules of identity introduction and identity elimination is a direct consequence of the commitment to the principle of substitutivity. If our understanding of identity is incomplete and does not allow us to see precisely why and in what circumstances intersubstitutivity holds and why and when it does not hold, then we have only a partial understanding of the distinction between intensional and extensional logic. It should perhaps also be noted here in this regard that, if the manner in which we view identity should change, then we might well have to redraw (or indeed even eliminate) the boundary between intensional and extensional logic.

Some may contend that all logic is extensional and that all contexts should be interpreted as being referentially transparent instead of opaque. For those who opt for this view, there is no distinction to be drawn between the two general sorts of logical inferences. The advocates of such a position may then not be troubled by the lack of a definitive solution to the problem of identity and may not be averse to regarding the identity predicate as an irresolvable primitive. However attractive this position may be from the standpoint of being able both to circumvent the problems posed by

propositional attitude contexts and to free oneself of the need to have a clear idea of identity, the denial that intensional idioms introduce contexts where there is some kind of problem with intersubstitutivity seems untenable. It clearly will not do to interpret contexts of direct and indirect quotation as referentially transparent. In the case of a statement involving direct quotation, the spoken or written words of another must be duplicated, in the same language or in the translator's language, nearly verbatim if the report on what was said or written is to be deemed true. In the case of a statement involving indirect quotation, the characterization of the spoken or written words of another is allowed greater variance, but an interpretation often can mischaracterize what was said or written if certain key words in the original are omitted, even if they are replaced by co-referring terms. If contexts of direct and indirect quotation should not in general be regarded as referentially transparent and such contexts are of the same nature as propositional attitude contexts, then the latter contexts should also not in general be regarded as referentially transparent. Thus, it often happens, and seems completely reasonable, that we accept as true such a sentence as "The Germans knew that the beginning of the Allied invasion of France would occur in June" at the same time we accept the falsehood of "The Germans knew that the Normandy invasion would occur in June," in spite of the fact that the beginning of the Allied invasion of France was the Normandy invasion. (The former, but not the latter, sentence accurately reports the Germans' state of knowledge.) It may seem convenient, for theoretical purposes, to dismiss, because of our acceptance of the truth of the former sentence, our acceptance of the falsehood of the latter sentence; however, whoever indulges in such a theoretical face-saving exercise is

inexorably left with the conclusion that a great deal of our talk about the propositional attitudes (as well as our talk about causes and explanations[35]) is nonsensical.

Although I have no desire to present here an *argumentum ad populum*, I find such a conclusion unwarranted. Admittedly, it is not the purpose of logical analysis to show the good sense of all that we say in ordinary language, but it is one of the primary goals of the formal analysis of arguments to account for the overall pattern of the valid inferences made by speakers of the language. The formal logician's analysis of the inference relations present in a language (or in a properly regimented version of the language) involves examining how the truth-value of one sentence is related to, or, in some cases, independent of, the truth-values of other sentences. The formal logician's task is to devise a grammar and a system of analysis to represent the logical form of statements in such a way as to allow him to make perspicuous both the truth conditions of and the inferential connections among statements, as those statements are understood by speakers of the language. Thus, holding to an analysis that is flagrantly at odds with common linguistic usage is both inconsistent with the spirit of logical analysis and counterproductive.

Furthermore, due to the aim of logical analysis, it can be said that a successful treatment of these special contexts that accommodates both the understanding of them as intensional and the understanding of them as extensional is, in general, to be preferred over an account of these contexts that always forces them to be read extensionally. The question whether the propositional attitudes should be understood

---

35. Biologist David Suzuki once remarked on the Discovery Channel's *The Nature of Things* that he was forced to spend part of his childhood in the 1940's in an internment camp "because of his genes." When understood literally, such a claim is surely incorrect. Even if it were possible to equate being of Japanese ancestry with the possession of certain genes (which is extremely doubtful at best), those who carried out the resettlement order did not send Americans of Japanese ancestry to such camps because of their genetic makeup, a variable which was certainly unknown to them. Thus, the interpreting of Suzuki's remark as being referentially transparent (which is involved in interpreting it literally) is clearly mistaken.

either as in some sense a relation between a person (a believer, a knower, etc.) and a particular proposition representing the content of the attitude or as a relation between a person and other persons or things is not one that can be legitimately settled just by deciding beforehand whether or not to accept both intensional and extensional logic. This question is central to the current debate over internalism and externalism in the philosophy of mind and needs to be addressed from a neutral perspective that recognizes at least the initial plausibility of both opaque and transparent readings of propositional attitude contexts.

3.3   Two other Fregean resolutions of the problem of identity

It is interesting to note that some contemporary philosophers, while evidently rejecting strictly Fregean senses, have nonetheless proposed in effect to resuscitate Frege's sense/reference distinction in order to solve the problem of identity. Unfortunately, these proposals have generally tended to reintroduce the sense/reference distinction (or something akin to it) without providing much, if any, elucidation of the concept of identity.

David Kaplan has suggested considering the meaning of any "fugitive" (as opposed to eternal or "timeless") sentence type that is an identity statement containing a demonstrative to be a function that maps an utterance of the sentence type at a particular time in a possible world to a content.[36] Kaplan construes the content of an utterance of a sentence type as a function that maps the utterance in a possible world to the truth-value the utterance would take on if it were made in that possible world. Thus, the meaning of a sentence type fixes the content of any particular utterance of that sentence type in a particular situation, and the content of

---

36. David Kaplan, "Dthat," in *Syntax and Semantics*, reprinted in A.P. Martinich, ed., *The Philosophy of Language*, 315-328.

an utterance of a particular sentence type in a particular situation fixes the truth-value of that utterance. Furthermore, for Kaplan, a content is necessary if the utterance having that content is a true utterance in every possible world, and a content is impossible if the utterance having that content is a false utterance in every possible world. The utterance of an identity statement containing a demonstrative has either a necessary content or an impossible content, and such identity statements are often contingent in the sense that their utterances have a necessary content in some contexts and an impossible content in other contexts. When the utterance of an identity statement expresses a true statement, the statement expressed is a necessary truth; when a false statement, the statement expressed is a necessary falsehood. However, knowing just the meaning of an identity statement is not sufficient in itself for knowing the content, and thus the truth-value, of any particular utterance of the identity statement.[37] Kaplan thereby accounts for how an utterance of an identity statement can express a necessary truth even though its truth is not known and is consequently an informative utterance. Thus, for Kaplan, an utterance of the sentence "That planet (the morning star) is identical to this planet (the evening star)" will express, when made under appropriate conditions and accompanied by appropriate behavior (a demonstration, such as pointing to or displaying something) to fix the reference of the singular terms containing the demonstratives, what is necessarily true, even though the sentence itself is not a necessary truth.

---

37. Kaplan's distinction between the meaning of a sentence and the content of an utterance also appears to be Simon Blackburn's later distinction between the character of an utterance (of a sentence) and the particular information expressed by an utterance of a sentence (i.e., the particular truth or falsity a sentence is used to communicate). As Blackburn illustrates with a case of a demonstrative utterance, the particular information conveyed by an utterance may depend upon the identification of the individual or individuals to which the speaker wishes to refer and under such circumstances is what Blackburn calls "identity-dependent" information. See page 303 of Blackburn's *Spreading the Word: Groundings in the Philosophy of Language*.

What is the case for demonstrative expressions may, Kaplan suggests, also often be the case for other singular referring expressions. The demonstrative expression serves only the purpose of enabling the speaker's utterance together with the demonstration to fix the speaker's intended reference. An understanding of how the demonstrative expression and the demonstration determine what is being referred to (the manner in which these are connected perhaps being amenable to a Fregean analysis) is not in general required for an understanding of the meaning of the sentence uttered. To understand what is being said specifically about particular things (i.e., to understand the content of a particular utterance), one needs to grasp the meaning of the sentence uttered and recognize the speaker's intended reference. Kaplan claims, although rather tentatively, that the distinction (and relationship) between the meaning of a sentence and the content of an utterance may well hold not only for sentences containing demonstratives but also for many sentences (or utterances of sentences) containing proper names or even definite descriptions. If Kaplan is correct, then an adequate understanding of the use of demonstrative phrases may thus be basic to an understanding of the use of all singular terms.

Kaplan's analysis of the meaning/content of any particular token of a sentence is overall Fregean in its approach in spite of the fact that Kaplan downplays the role played by what Frege would regard as senses. For Kaplan, as was the case for Frege, there is associated with any particular declarative sentence a composite function from that sentence, as uttered by a speaker in a particular context, to a truth-value, that function representing the meaning/content of the sentence. Kaplan's proposal is thus vulnerable to the sorts of criticisms that can be leveled against Frege's analysis when it comes to the matter of identity statements. As a result, Kaplan has not really elucidated the concept of the identity relation, particularly since he never specifies the

functions that represent the meaning of an identity statement and the content of an utterance of such a statement. Kaplan has just indicated in a general way how it is possible for an identity statement to be considered contingent and informative even though any particular utterance of such a statement, if true, will necessarily be true.

Searle takes a somewhat different approach to the problem of how it is possible for identity statements to differ in content and significance.[38] Searle's solution relies fundamentally on introducing a notion that corresponds to Frege's notion of sense. According to Searle, a proper name is associated with a complex of descriptions a sufficient number of which must be satisfied by exactly one individual in order for the name to have a referent. The descriptions with which proper names are associated will vary, and some proper names that have the same referent will nevertheless be associated with significantly different descriptions. When the proper names in an identity statement are associated with the same descriptions and there is exactly one thing that is the referent of both names, the identity statement is a trivial analytic truth. However, when the proper names in an identity statement are associated with different descriptions and yet there is exactly one thing that is the referent of both names, the identity statement is a synthetic truth.

For Searle, the sentence "Tully is Tully" and the sentence "Tully is Cicero" can both be used to express an analytic truth, and the ability of a speaker to use either one to express such a truth is a reflection or illustration of contingent linguistic rules governing the use of proper names generally and the use of "Tully" and "Cicero" to refer specifically. Without linguistic rules ensuring that successive occurrences (tokens) of the same proper name (type) in a sentence will refer to the same individual and without linguistic rules regarding how "Tully" and "Cicero" are used

---

38. John R. Searle, "Proper Names," *Mind* 67 (1958), also reprinted in A.P. Martinich, ed., *The Philosophy of Language*, 270-274.

to refer, neither sentence could be used to make a claim that is analytically true. Searle, though, admits that the rules that allow the use of "Tully is Tully" to express an analytic truth are more universally conventional in the use of language than are the rules that allow the use of "Tully is Cicero" to accomplish the same feat. The two identity statements therefore differ in significance at least in the sense that each is a reflection, or illustration, of different sets of linguistic rules. This observation, while not ruling out the possibility that identity statements containing distinct proper names can express significant synthetic truths, nevertheless does not enable us to see how such statements are possible. The issue of just that sort of possibility is, of course, at the heart of the problem of identity. Searle acknowledges that those who argue, for instance, that Shakespeare was Bacon are not advancing a thesis about language but are intending by their use of the sentence "Shakespeare is Bacon" to express an important historical truth.

Two strategies have traditionally been pursued in order to explain the manner in which proper names are connected with the individuals to which speakers intend to refer, and thereby to account for the nature of the rules governing the use of proper names to refer. These two approaches represent for Searle two extremes; Searle rejects both and instead offers a proposal that is a compromise between these extremes. On the one hand, some have considered proper names to be devices used exclusively to refer and thus to be devoid of any descriptive content whatsoever. On the other hand, others have viewed proper names as terms always having a descriptive content in virtue of which they have, or do not have, reference.

If proper names necessarily have reference but lack anything corresponding to sense (which would seem to be the case if proper names are completely devoid of descriptive content), then it is difficult to see how the rules for the correct use of

proper names to refer to particular individuals could be learned and it is difficult to assign an adequate meaning to any denial that a proper name has a referent. In order to learn to how to use a particular name to refer, one must learn the connection between the use of the name and the object to which speakers intend to refer. The connection between the name and the intended referent is ultimately established by describing the latter, even in cases where a speaker indicates the intended reference by an act of ostension. Since the descriptions associated with the proper name fix the intended referent of the name and thereby determine in effect the rule for the correct use of the name, these descriptions can be construed as collectively constituting the sense of the proper name. Hence, if one rejects the notion of proper names having a sense, then one would also have to reject the notion of there being learnable descriptive rules for the use of proper names.

Perhaps even more importantly, if proper names have no descriptive content, then any denial that a name has a referent appears not to have the meaning and significance normally attributed to such a claim. If, for example, someone were to utter the sentence, as evidently some people actually have, "Socrates never existed," what the speaker would most likely be intending to say is that the features customarily ascribed to Socrates were never possessed by any one individual or were possessed collectively by different people living at different times or places. The speaker should not be interpreted as saying merely that no one named "Socrates" ever existed (which is false), but should rather be interpreted as saying something like, "There was no influential ancient Greek philosopher who lived in Athens, was Plato's mentor, was tried and convicted of impiety and corrupting the youth, and was executed in 399 B.C." Such an interpretation of the denial that a proper name has a referent is precluded, though, if the use of a proper name does not involve the at least

implicit attempt to describe or allude to some of the features thought to be possessed by the referent of the name.

However, if proper names necessarily have sense and only contingently, if at all, have reference, then there are a whole host of other difficulties of a different sort. If proper names have a meaning or sense in that there is associated with any name a description that must be satisfied by an object in order for the name to be used to refer to that object, then the particular meaning attached to a name may vary from one person to another and that meaning may have to change in order for the name to have the same referent, or any referent at all, if the object the name is used to refer to should change its attributes. Furthermore, if proper names have descriptive content, then they function in effect as shorthand for their descriptions, which means that some statements that involve the name in simple predications are trivial analytic truths and others are self-contradictory. Thus, for instance, if at least part of the meaning of the name "Aristotle" is given by the definite description "the famous philosopher from Stagira who tutored Alexander the Great," then the sentence "Aristotle was once a tutor" can be used to express a truth that is both trivial and analytic. However, this does not seem right; as Searle notes, it is only a contingent matter of fact that Aristotle ever went into pedagogy.

What I suspect is of even more importance for Searle is the fact that, if proper names were to function in effect as descriptions, then the linguistic rules governing the use of proper names to refer to particular individuals would be so precise that a name could be used to refer to its intended referent only if the latter satisfied a certain definite complex of descriptions, which would mean that proper names are logically equivalent to that complex of descriptions and are therefore superfluous. Under such circumstances, there would be no difference between referring and describing, since

one could only succeed in referring by describing. That difference, though, is what marks, for Searle, the distinction between proper names and descriptions. If the rules for the use of proper names are precise, then the descriptions that determine the referent of a proper name must specify what it is to be identical to the referent of the name (i.e., they must specify the identity conditions for the referent of the name). The descriptions that collectively give the sufficient and necessary conditions for applying a proper name will not specify merely that the referent of the name is a particular thing that happens to bear the name as its label, or, in other words, that may be or has been referred to in speakers' referring uses of the name. (Thus, for example, the fact that there lived in the twentieth century a certain wealthy Greek tycoon is not a sufficient condition for saying that philosophers' use of the name "Aristotle" succeeds in referring.) However, Searle contends that the great pragmatic significance of proper names lies precisely in the fact that speakers who use proper names do not need to come to any prior public agreement on what exactly constitutes the identity conditions of the things to which they use proper names to refer. In most cases, the issue of what constitutes the identity of the intended referent is never even raised by a speaker's referring use of a proper name. In contrast, a definite description is always used to specify the condition that must be satisfied by the unique referent of the description. So, if proper names function ultimately as definite descriptions, then the former are eliminable in terms of the latter and thus lack their important role in speech, a role that seems to distinguish them from descriptions.

What Searle proposes as a solution to these difficulties is in important respects an intermediate between the view of proper names as purely referential and the view of proper names as fundamentally descriptive. According to Searle, proper names, with but few exceptions, are not strictly equivalent to descriptions, but proper names

nevertheless do possess a sense in that their referring uses always involve certain "descriptive presuppositions."[39] For Searle, a speaker who uses a proper name to refer presupposes, but does not assert, that some of the descriptions of an unspecified complex of descriptions hold true of the individual to which the speaker intends to refer, this complex of descriptions detailing the attributes the speaker believes the intended referent to possess. Since presumably what is known, or at least is believed to be true, about an individual will vary somewhat over time and among different people, the descriptions associated with a proper name will vary accordingly. Certain of these descriptions a speaker will regard as more important in that their fulfillment is considered more crucial to a specification of what amounts to the identity conditions of the individual to which the speaker intends to refer, these descriptions collectively constituting the descriptive presuppositions involved in a speaker's referring use of the proper name of that individual. Still, since what a speaker takes to be an essential fact about an individual and thus a component of the identity conditions of the individual is rather vague and will vary from speaker to speaker, the extent and nature of these descriptive presuppositions is nevertheless left rather indefinite in a typical referring use of a proper name.[40] The absence of an explicit specification of these descriptive presuppositions marks, according to Searle, the distinction between referring and describing, and the consequent lack of precise criteria for applying proper names is what for Searle will in general serve to

---

39. The exceptions involve proper names that assume the form of a description or that have acquired a strict descriptive use. A proper name such as "The Bank of England" (Searle's example) or "The artist formerly known as 'Prince'" (my example) carries, in this case, what would customarily be its presuppositions as a very part of its meaning. For an example of a proper name that is commonly used as a substitute for a description, consider that traditional Western theists have come to use the word "God" to mean "the omnipotent, omniscient, omnipresent, all-loving being who created the universe but is separate from it."

40. This conclusion is reminiscent of the same sort of conclusion reached by Ludwig Wittgenstein in section 79 of his *Philosophical Investigations*. See, for instance, *Philosophical Investigations*, trans. by G.E.M. Anscombe, 3rd ed., pp. 36e-38e.

distinguish proper names from descriptions. Therefore, proper names, Searle claims, function not as descriptions, but as "pegs upon which to hang descriptions." Furthermore, the complex of descriptive presuppositions associated with a proper name functions like a Fregean sense in that a speaker who uses the name to refer can be considered as successfully referring in virtue of the fact that a sufficient number of the descriptions he presupposes as satisfied are indeed true of the intended referent.[41] Similar to what is the case with Fregean senses, the issue of what exactly are the presuppositions associated with a proper name and precisely how many of a speaker's presuppositions need be satisfied in order for referring uses of the name to be successful is, however, in general never raised (nor is a definitive answer presupposed) by a speaker's referring use of the name. This approach, according to Searle, can yield an adequate account of what a speaker asserts in denying that a proper name has a referent. In denying that a proper name has reference, a speaker in effect asserts that a sufficient but unspecified number of the descriptions conventionally associated with referring uses of the name (more specifically, all those descriptions that are conventionally presupposed and regarded as essential to fixing the alleged reference of the name) are true of no one individual.

Thus, Searle, if I am interpreting him correctly, considers the sense a speaker attaches to a proper name to be the indefinite complex of descriptive presuppositions with which the speaker associates the name, while the sense of a proper name in general is the indefinite complex of conventional presuppositions with which speakers have come to associate the name. In addition, the linguistic rules for the use of a proper name must be based upon a specification of the identity conditions of the name's referent since for Searle it is these identity conditions that connect the name

---

41. The idea of the referent of a proper name being fixed by descriptions has also been endorsed by Paul Ziff in his *Semantic Analysis*, 102-105.

with its referent. Hence, the sense of a proper name, for Searle just as it was for Frege, determines the reference of the name, since, for Searle anyway, the sense of a proper name amounts to the complex of descriptive presuppositions that constitutes the identity conditions of the name's referent and figures in the linguistic rules for the use of the name.[42]

Given this understanding of proper names, Searle then offers a solution to the problem of identity that parallels Frege's solution. The terms used in an identity statement can have the same referent, in which case the identity statement is true. Searle's commitment to an objectual account of identity is clearly revealed by his observation that, in order for the proper name "Aristotle" to apply correctly to an object, it is both sufficient and necessary that the object in question be identical with Aristotle (instead of merely being identical to some individual named "Aristotle"), which in turn will be the case when and only when that object satisfies the descriptive presuppositions associated with the name "Aristotle" (or, in other words, satisfies the conditions both sufficient and necessary for an object to be Aristotle). However, the terms used in an identity statement can be, or can be associated with, different descriptions that single out unique individuals, in which case the identity statement,

---

42. Although Searle is not altogether careful to spell out all the steps of his reasoning, his main argument can, I believe, be summarized as follows.

A speaker follows linguistic rules and uses a proper name to refer successfully to an object when and only when the object he intends to refer to satisfies the identity conditions of the object denoted by the name he uses.

The sense of a proper name is what connects the name with the unique object to which speakers use the name to refer.

The condition both sufficient and necessary for the object to which the speaker intends to refer to be identical to the object denoted by the proper name is that the former object satisfy the complex of descriptions that all fit only one individual and that establish the connection between the name and the object denoted by the name.

When a speaker uses a proper name to refer, he typically presupposes that the requisite connection exists between the proper name and the object to which he uses the proper name to refer.

Hence, when a speaker follows linguistic rules and uses a proper name to refer, he typically presupposes that the object to which he intends to refer satisfies the complex of descriptions that constitutes the sense of the proper name.

even if true, may nevertheless be neither analytic nor trivial. Indeed, the recognition of the truth of an identity statement made with proper names that are associated with very different descriptive presuppositions can amount to the acquisition of a significant bit of knowledge. Thus, the sentence "Tully is Cicero" is true, and for most people would typically be used to make an analytic statement, assuming, of course, that for most people the two names are associated with the same complex of descriptive presuppositions. In addition, the sentence can be uttered as a standard way of making what is an analytic statement, provided that the general sense of the two names is the same. However, the sentence "Shakespeare is Bacon," even if it should prove to be true, would not be used by most people to make an analytic statement, given the fact that most people do not associate the two names with the same complex of descriptive presuppositions. Furthermore, the sentence cannot be uttered as a standard way of making what is an analytic statement since the general senses of the two names are very different. Instead, a typical utterance of the sentence would only be used (both standardly and for most people) to make a synthetic statement, a statement that, if discovered to be true, would represent a valuable extension of our knowledge.

Unfortunately, Searle's proposed solution to the problem of identity does not provide us with an explication of the concept of identity and can be faulted, as a theory of identity, for the same basic reasons as can be Frege's proposed solution. As a theory of identity, Searle's solution, as does Frege's, suffers from a major defect. According to both accounts, an identity statement is true if and only if its singular terms have the same referent, and a true identity statement is to be deemed analytic when its singular terms have the same sense. Furthermore, on Searle's account, the sense that attaches to a proper name is its set of descriptive presuppositions that

represent the identity conditions of the name's referent. Thus, we can judge Searle's solution, no less than Frege's, to be successful only if we already have at our disposal a separate understanding of identity.

Searle's solution is questionable for other reasons as well. According to Searle, proper names are typically used to say something about an individual, and their characteristic use involves presupposing, but not asserting, that the individual or individuals to which a speaker intends to refer satisfy certain descriptions. However, the relevant presuppositions become part of what is asserted in cases where the proper name has no referent and the speaker is using the proper name to assert that the name has no referent. Furthermore, Searle has suggested that, at least in some cases where a speaker uses a proper name to attribute, quite mistakenly, a property to an individual that does not actually possess the property, the proper name can be legitimately interpreted to be functioning as a description singling out the individual that in reality does possess the property, even if that individual is unknown to the speaker.[43] So, in some cases the use of a proper name involves descriptive presuppositions that are neither severally nor jointly equivalent to the name, while in other cases the use of a proper name involves descriptions that do figure in what a speaker is actually in effect saying. In order for Searle's analysis of these two cases not to seem ad hoc, Searle needs to clarify what the difference is between the two uses of a proper name that would account for the two distinct interpretations. For Searle, such an account would be given within his general theory of speech acts. Presumably, one could plausibly say that the former uses amount to a referring use of a proper name and the latter uses amount to a nonreferring or descriptive use of a proper name. Unfortunately, Searle never elucidates such a distinction. I suspect that

---

43. See Chapter 6 of Searle's *Expression and Meaning: Studies in the Theory of Speech Acts.*

such a distinction would need to be explained ultimately on the basis of a more general account of how speakers use proper names to indicate what it is that their pronouncements are about and what may be the result when the conditions necessary for reference to be successful are not met. Without an explanation of the distinction on such a basis, Searle's account appears to represent more of a series of observations on the uses of proper names than a general theory on the semantics of statements containing proper names.

Searle also leaves unanswered important questions about the nature of the descriptive presuppositions that he alleges are associated with the use of proper names. Even though the extent of these presuppositions is for Searle always, by the very nature of proper names, rather vague, Searle still does not make sufficiently clear what he considers to be the conventional descriptive presuppositions associated with a proper name. Do the conventional presuppositions that are involved in the use of a particular proper name relate only to what is commonly known among people who have expert knowledge about the referent of the name, or to what is commonly known among all people who just have a certain familiarity with the referent of the name? Since what is known and considered important about the individuals speakers use proper names to refer to often changes over time, should we say that the Searlean sense of a proper name changes over time? The precise relationship between the descriptive presuppositions and an understanding of the content of an utterance involving the use of a proper name also needs to be clarified. Do those who have very little knowledge of the individual to which they use a proper name to refer presuppose no more than what they say literally when they use the name in an utterance intended to report all that they know about that individual? For instance, would a person who knew that Aristotle was an ancient Greek philosopher but knew

nothing else about Aristotle presuppose exactly what he asserted in uttering the sentence "Aristotle was an ancient Greek philosopher"? If such a person were to presuppose what he asserted, then, oddly enough, he would only know about Aristotle what is according to Searle an analytic truth. It is hard to imagine that speakers who lack a certain familiarity with the objects to which they use proper names to refer presuppose anything about the referents of those names. It is even more difficult to imagine that an understanding of the content of a statement made using a proper name, an understanding of such a statement being necessary if any knowledge at all concerning the referent of the name is to be obtained or communicated to others, should require one to recognize beforehand what is being presupposed about that referent. A theory that has it that the reference, if not also the meaning, of a proper name is determined or given by a single description or a cluster or family of descriptions is a description theory of names. Saul Kripke has argued against such a theory of reference or meaning, and it is Kripke's views on identity that constitute the main topic of the next chapter.[44]

---

44. In *Naming and Necessity*, Kripke presents a sustained attack on the description theory of names, including a criticism of Searle's position. Kripke, though, charges Searle with advancing a view that, in all fairness to Searle, I believe Searle does not really hold. In "Proper Names," Searle claims that "... it is a necessary fact that Aristotle has the logical sum, inclusive disjunction, of properties commonly attributed to him: any individual not having at least some of these properties could not be Aristotle." Kripke takes this remark as indicating that Searle holds that it is epistemologically or metaphysically necessary that Aristotle had the properties commonly attributed to him. However, given the specific example and the wider context of Searle's remark, I highly suspect that Searle was using the phrase "necessary fact" only to express his very high degree of confidence in the correctness of his account of how proper names become attached to their referents.

Chapter 4: Kripke's Solution

4.0   Introduction

I wish to discuss in this chapter and the next chapter two additional strategies that philosophers have pursued in order to solve the various difficulties that must be surmounted in order to have an adequate analysis of identity. The first strategy is to follow Frege's lead and offer a solution that relies on a distinction between different types of necessity. More specifically, the next account of identity to be examined involves drawing a distinction between statements that are necessarily true and those that are known to be true a priori and regarding all identity statements as statements, about objects, that are necessary but not always known to be true a priori. The second strategy is to treat identity statements as statements that are, at least in part, about their terms and involves modifying the traditional notion of a purely referential context and its associated principle of substitution in order to account for the differences in the informational value of identity statements. Kripke is one who has opted for the first of these just-mentioned strategies, and, as I indicated at the end of the previous chapter, his account of identity statements is the primary concern of the present chapter.[1] Although Kripke most likely never set out specifically to solve the problem of identity, as perhaps a means of going on to solve further philosophical quandaries, there nevertheless emerge from what Kripke has said about the notion of identity a distinct view of identity and manner of solving some of the difficulties to which the notion gives rise. It is this stance that Kripke takes with regard to the issues that are relevant to my discussion of identity that I will have in mind when I

---

1. Saul Kripke, "Identity and Necessity," in *Identity and Individuation*, 135-164. See also Kripke's *Naming and Necessity*.

use such descriptions as "Kripke's account of the nature of identity statements" or "Kripke's solution to the problem of identity."

4.1   Kripke's proposals

Kripke regards identity as a relation that always holds between every object and itself (i.e., a relation that holds between an object and itself in every possible world in which the object occurs). According to Kripke, when the identity relation is discovered to hold between the referents of two terms, what is discovered is a relationship between objects that could not be otherwise, even in cases where, in the absence of empirical investigation, we could not have ascertained beforehand whether this relation obtained. Kripke maintains that it is entirely plausible to regard any true identity statement (or at least what such a statement says about the referents of its terms) as being in at least one sense necessary, in spite of the fact that many true identity statements are not true just in virtue of the meanings of their words or are not known to be true prior to experience. More specifically, if the notion of *de re* necessity is intelligible (and evidently Kripke thinks it is), we can take true statements of identity to express metaphysically necessary truths about the referents of the terms in identity statements.[2] Kripke regards true identity statements as being what he calls "weakly necessary": they are true in all possible worlds in which the referents of their terms exist. What true identity statements always say about the referents of their terms is weakly necessary, i.e., it is true of those referents (particular objects) in every possible world in which they exist. Thus, in asserting the sentence "Mark Twain is

2. *De re* ("due to the world") necessity is supposed to be necessity due to the nature of the things talked about; this is in contrast to *de dicto* ("due to the word") necessity, which is necessity due to the manner of talking about things. If every object has the property of necessarily being identical to itself (which is presumably the case if we correctly use the identity predicate only to express a metaphysical relation between a thing and itself), then "**a** = **a**" is both necessary *de dicto* (since it is tautologous) and necessary *de re*.

Samuel Clemens," one claims that a necessary metaphysical relation holds between the referent of "Mark Twain" and the referent of "Samuel Clemens." Likewise, in asserting the sentence "The richest man in Seattle is the richest man in the world," one claims that a necessary metaphysical relation holds between the referent of "the richest man in Seattle" and the referent of "the richest man in the world."

Since some identity statements occur in generalizations that do not contain proper names, Kripke sees the view of the identity relation as a relation between names or terms as unintelligible. Furthermore, even if we were to understand the identity relation as a relation between terms, we could, Kripke contends, still define what would be an artificial relation (a relation Kripke calls "schmidentity") holding between every object and itself.[3] Kripke notes that the sorts of problems with statements of identity that have led some people to take identity to be a relation between terms would also arise for statements of schmidentity. As a consequence, Kripke concludes that the existence of these problems does not constitute evidence for a metalinguistic account of identity. Nevertheless, since a typical objectual analysis of identity fails the difference in content requirement (i.e., fails to draw a distinction between the content of "$a = a$" and the content of "$a = b$") and satisfies only partially the nonnecessity and nontriviality requirements, Kripke's position encounters the difficulty that I mentioned previously in regard to any objectual account of identity: it appears to be counterintuitive in that it seems to deny the possibility of there being true statements of identity that are only contingent statements. After all, if true identity statements only assert a relation between objects,

---

3. See *Naming and Necessity*, Lecture III. Presumably, Kripke has in mind defining schmidentity in terms of how we currently understand the identity relation. In the hypothetical case, though, Kripke suggests calling that relation something else and reserving the term "identity" for the relation that holds between terms. In the hypothetical case, I take it that the referent of "a" is schmidentical to the referent of "b" if and only if "a" is identical to "b."

a relation that always and trivially holds and that must hold between every object and itself, then it is hard to see how such statements could be contingently true.

Before presenting the details of Kripke's position, I want first to discuss Kripke's version of the problem of the contingently true status of identity statements and then examine in some depth Kripke's suggestion that an appeal to Bertrand Russell's theory of definite descriptions can be employed to solve the puzzle. When definite descriptions occur in identity statements, the situation becomes more complex than when identity statements contain only proper names. Thus, some elaboration on the use of definite descriptions will be in order. Finally, after clarifying some distinctions Kripke makes that are important to understanding his views on identity, I will then present, discuss, and critique those views.

As Kripke points out in *Naming and Necessity*, the view of identity statements as necessary introduces an additional perplexity. If it is true that, for any *x* and *y*, to say that *x* and *y* are identical is to say that all properties of *x* are properties of *y* and vice versa (which means that either one of the singular terms occurring in a true identity statement can be substituted for one another in all contexts without altering truth-value), then it would seem that we would have to accept as valid the following inference.

Benjamin Franklin was the first postmaster general.

Benjamin Franklin was the inventor of bifocals.

It is necessarily true that Benjamin Franklin was Benjamin Franklin.

Therefore, it is necessarily true that the inventor of bifocals was the first postmaster general.[4]

This inference seems specious since the conclusion of this argument looks false: it appears in no way necessary that the inventor of bifocals should be the first postmaster general. Presumably, the historical facts could have been different than they are, and it might not have been the case that the same person both invented bifocals and headed the first U.S. post office. As a consequence, the statement "The inventor of bifocals was the first postmaster general" certainly appears to be a contingent truth. However, if we regard identity statements as being necessarily true when true, then we seem to be forced to accept both the truth of the conclusion and the validity of the inference. Some may contend that what this shows is that modal contexts, just as propositional attitude and explanatory contexts, are evidently not transparent. Nevertheless, even if we have qualms about substituting co-referring terms for one another in a modal context, by regarding true identity statements as necessary the first two premises of this argument are necessary so that the conclusion is not derived from at least one contingent statement. The modality of necessity in the conclusion is thus warranted since the conclusion is derived only from necessary statements.[5] Hence, the apparent transparency of at least necessity contexts is assured if true identity statements are always taken to be necessary.

Kripke contends that Russell's treatment of definite descriptions provides a solution to the above problem, although the account of identity Kripke ultimately proposes does not rely fundamentally on Russell's theory of descriptions. According

---

4. There is another way to reach this conclusion. Since the first postmaster general is necessarily identical to the first postmaster general and the inventor of bifocals is the first postmaster general, the inventor of bifocals is necessarily identical to the first postmaster general.

5. The inference may not be reasonable, though, if different notions of necessity are involved.

to Kripke, by analyzing definite descriptions in the Russellian fashion and by paying close attention to scope, the apparent conflict between our ordinary understanding of necessity and the alleged truth of such a sentence as "Necessarily the inventor of bifocals was the first postmaster general" can be resolved. The trouble with the conclusion of the above problematic inference is that the scope of the necessity operator appears to be the entire sentence when actually it is not. The precise scope of the necessity operator is revealed when the sentence is rendered into one of the forms suggested by Russell.[6] The two possible readings of the sentence, one wherein the necessity operator has wide scope (or what Kripke calls "*de dicto* scope") and the other wherein the necessity operator has narrow scope (or what Kripke calls "*de re* scope"), are given by the following two sentences.

1. Necessarily, it is the case that there was both a unique person who invented bifocals and a unique person who held the original office of postmaster general and the former person is identical to the latter person.

2. There was both a unique person who invented bifocals and a unique person who held the original office of postmaster general, and necessarily the former person is identical to the latter person.[7]

Sentence 1 is false since it is not necessary that there was a unique person who invented bifocals and a unique person who held the original office of postmaster general. Sentence 2 is true since both definite descriptions serve to pick out

---

6. Bertrand Russell, "On Denoting," *Mind* 14 (1905), reprinted in *Contemporary Readings in Logical Theory*, ed. Irving Copi and James Gould.

7. The first sentence has the form of "$\sim\Diamond\sim(\exists x)(\exists y)(((Bx \ \& \ (z)(Bz \rightarrow x=z)) \ \& \ (Py \ \& \ (z)(Pz \rightarrow y=z))) \ \& \ x=y)$," and the second one has the form of "$(\exists x)(\exists y)(((Bx \ \& \ (z)(Bz \rightarrow x=z)) \ \& \ (Py \ \& \ (z)(Pz \rightarrow y=z))) \ \& \ \sim\Diamond\sim x=y)$."

Benjamin Franklin and that individual is necessarily identical to himself. The definite descriptions occurring in sentence 1 are both subordinate to the necessity operator and thus have a secondary occurrence in that sentence. In contrast, the necessity operator is subordinate to both definite descriptions occurring in sentence 2 so that these descriptions have a primary occurrence in that sentence. The apparent implausibility of the above-mentioned argument vanishes provided that we understand all definite descriptions in the Russellian fashion and interpret the definite descriptions occurring in the conclusion as having a primary occurrence in that sentence. The inference is indeed problematic, though, when the conclusion is interpreted as the first of the above two sentences.

The strategy Kripke follows to solve his puzzle about the inventor of bifocals and the first postmaster general may suggest to an objectual identity theorist a way to account for the contingency of at least some true statements of identity. Apparently, by appealing to Russell's theory of descriptions, an objectual identity theorist could plausibly maintain that true identity statements always express metaphysically necessary truths without denying that there are statements of identity that are true only as a matter of contingent fact. An advocate of an objectual account of identity may defend the view that what a true identity statement says about the object to which the terms flanking the identity sign both refer is necessarily true of that object, even though the identity statement itself may only be contingently true due to other component claims implicit in the statement.[8] Some true identity statements will be contingent because they contain a definite description, or a name functioning as a disguised description, and these descriptions will serve to specify uniquely

---

8. One who advances such a view will actually maintain that contingently true "identity statements" are not, properly speaking, identity statements but do contain a component that is a genuine statement of identity.

individuals that are necessarily identical to themselves. The element of necessity will enter into a contingently true identity statement only in the assertion that certain individuals are identical, those individuals being singled out by one or more definite descriptions in the statement. This means that the necessity operator implicitly associated with an identity statement containing a definite description must always be understood as having the smallest possible scope. Thus, all true statements of identity express necessary truths, but some true identity statements say more than that. Any identity statement containing a definite description that is resolvable in the Russellian fashion also involves a component claim that is not necessary (i.e., the claim that there is a unique individual satisfying the description), and such identity statements may differ in content due to differences in the content of the definite descriptions occurring in them. Furthermore, no identity statement that contains a definite description, or a disguised definite description, that can be understood in the Russellian manner will be trivial in the sense that it will at least assert that there is a unique individual that satisfies the description.[9]

Unfortunately, the troubles are not over for the objectual identity theorist who, like Kripke, wishes to maintain that the identity relation holds necessarily whenever it obtains in spite of the fact that many identity statements are contingent. I contend that the appeal to Russell's theory of descriptions, and the notion of scope, in order to

---

9. In virtue of his theory of descriptions, it would appear (although I think this is only an appearance) that Russell was thus able to solve the difference in content, nonnecessity, and nontriviality requirements provided that all identity statements contain only definite descriptions or names functioning as definite descriptions. The difference in content requirement is at best solved in an unusual manner. The statement "$a = a$" differs in content from the statement "$a = b$" (provided that "a" and "b" are not synonymous definite descriptions) because the uniqueness claim associated with "a" is different from the uniqueness claim associated with "b," but the statement "$a = a$" is not a tautology. The way in which the nonnecessity requirement is satisfied is also suspect. In general, an identity statement that contains a definite description appears to be contingent not just because a description only contingently denotes something but also because the identity of the referents of the terms flanking the identity predicate itself seems contingent.

explain the contingency of identity statements containing definite descriptions does not by itself solve, in its entirety anyway, the problems regarding the alleged necessity of true identity statements. The logical contingency of identity statements containing a definite description can be accounted for on Russell's analysis since such statements are never interpreted as being logical truths, even in the case of a statement in which the same definite description flanks the identity predicate. Accordingly, the sentence "Benjamin Franklin was the first postmaster general" will yield, when parsed in the standard Russellian fashion, the following existential generalization, sentence E.

E.   There was a person, and only one person, who was the first postmaster general, and Benjamin Franklin was that person.

Sentence E can, in turn, be rendered into symbolic notation as the following formal expression, sentence E'.

E'. $(\exists x)((Px \mathbin{\&} (y)(Py \rightarrow x = y)) \mathbin{\&} x = b)$

Since sentence E is not a tautology, as revealed by sentence E', and it is a contingent matter of fact that there was a unique person who was both the first postmaster general and identical to Benjamin Franklin, the logically contingent and overall metaphysically contingent nature of sentence E is secured. Each implicit necessity operator associated with an occurrence of the identity predicate in sentence E' is understood to have the smallest possible scope, and for this reason the entire sentence E', and hence also the entire sentence E, is not within the scope of an implicit

necessity operator. Consequently, the statement "Benjamin Franklin was the first postmaster general" is, according to a Russellian analysis, a contingent statement that nevertheless, according to the objectual view of identity under consideration, involves the ascription of a metaphysically necessary relation between Benjamin Franklin and the first postmaster general.

However, something still seems lacking here. Since it is an accepted fact that Benjamin Franklin was in actuality the first postmaster general, the description "the first postmaster general" can be used to refer to Benjamin Franklin. In addition, it seems necessary that the first postmaster general, Benjamin Franklin or whoever else that person happened to be, headed the first U.S. post office and was unique in so doing. Thus, any value of the variable "x" in sentence E' that is identical to Benjamin Franklin is, evidently, guaranteed to have headed the first U.S. post office and to have been unique in so doing. As a result, sentence E, and hence the original identity statement, presumably is true in every possible world in which Benjamin Franklin exists and is thus weakly necessary. Therefore, it would appear that the sentence "Benjamin Franklin was the first postmaster general" and the seemingly innocuous sentence "Necessarily, the first postmaster general was the first postmaster general" entail the weak necessity of the original identity statement. Furthermore, there is nothing special about that statement; an argument analogous to the one I just gave can be given to establish the weak necessity of any such identity statement. Kripke's problem concerning the necessity of what appear to be contingent statements of identity is thus resurrected. What needs to be introduced is some additional apparatus in virtue of which it can be shown that the above inference is faulty. The appeal to Russell's analysis of descriptions, and the notion of scope, is, by itself, inadequate to

give a successful account of the contingent nature of identity statements containing definite descriptions.[10]

One way of remedying the problem (a strategy Kripke would, however, reject) would be to introduce into an account of the use of definite descriptions a way of making plausible the treatment of the description as, in some cases, a proper name (or something like a proper name) and, in other cases, a tacit assertion of uniqueness or a disguised universal generalization. The manner in which the description is functioning in an identity statement may sometimes affect the scope of the implicit, or explicit, necessity operators in the statement. The inclusion of such a distinction as the one Keith Donnellan makes between the referential and attributive uses of a description would be an important element of an account that would warrant the requisite distinction in the treatment of definite descriptions.[11] For Donnellan, when it is said that we can use the description "the first postmaster general" to refer to Benjamin Franklin, what is meant is that those who specifically wish to refer to Benjamin Franklin can use that description, as well as any other description that singles out Benjamin Franklin as its unique referent, as a name of Benjamin Franklin. Such a use of the description to refer to Benjamin Franklin would be a referential use

---

10. It should be fairly obvious that the suggested approach to identity statements, when understood as an account of identity, is also incomplete. One who takes this approach will interpret any identity statement that contains a definite description as a generalization, and that generalization will contain the identity predicate within the scope of a quantifier. This latter occurrence of the identity predicate (in what properly speaking is an identity statement) will need to be understood if the occurrence of the identity predicate in the unanalyzed identity statement is to be understood. Presumably, the generalization will ultimately have (or can be viewed as, in effect, having) as its instances sentences containing only proper names, and the identity predicate occurring in those instances will be flanked by proper names. Only if an adequate account can be given of the meaning of these identity statements in the instances of the generalization will an adequate account be had of the meaning of that component of the generalization containing the identity predicate (i.e., what constitutes the identity statement proper). Hence, the suggested Russellian approach to analyzing apparent identity statements needs an account of identity statements proper in order to be complete.

11. Keith Donnellan, "Reference and Definite Descriptions," reprinted in A.P. Martinich, ed., *The Philosophy of Language*, 235-247.

of the description. A definite description is being used in a referential manner when its occurrence in a sentence serves only the purpose of enabling the speaker to make reference to a particular person or thing the speaker has in mind. When a statement contains a definite description used in this manner, the occurrence of the definite description is not essential to the meaning of the statement in the sense that any other description that uniquely specifies the same object as the original description could have been used instead without a change in the content of the statement. A speaker who uses a definite description referentially uses it merely as a device that allows his audience to pick out the particular object about which he intends to say something. The use of a definite description in this manner is, roughly speaking, the use of the description as a proper name. According to Donnellan, an attributive use of a definite description is something different. When it is said that the first postmaster general was necessarily the unique person who headed the first U.S. post office, what is meant is that whoever was the first postmaster general (the referent of the description "the first postmaster general") was the head of the first U.S. post office solely in virtue of being the first postmaster general. In other words, it is being claimed that anyone who was the head of the first U.S. post office was the one who was the head of the first U.S. post office. (If the attributive use of the description in the sentence "The first postmaster general was the head of the first U.S. post office" is understood to render the sentence a universal generalization, then the sentence is understood to be a logical truth.) What is not meant is that someone whom the speaker specifically has in mind, and is being referred to in the speaker's use of the description "the first postmaster general," necessarily was the head of the first U.S. post office. The use of the description as illustrated by sentence E is an attributive use of the description. A definite description is being used in an attributive manner when its occurrence in a

sentence serves the purpose of allowing the speaker to say something about whatever, or whoever, satisfies the description. When a definite description is used in this manner, its occurrence in a statement is essential to the meaning of the statement in the sense that the content of the description is part of the content of the statement. The content of the statement will therefore not in general be preserved if such a description is replaced by another description that characterizes the same particular thing that the original description characterizes. The attributive use of a definite description is successful when a unique object is being referred to (specified or picked out) in virtue of that object uniquely satisfying the description, but the speaker who uses a definite description attributively does not convey to his audience that he has a particular object in mind, an object, that is, that he could refer to in any number of different ways.

The Russellian strategy Kripke suggests as a way of solving the puzzles that arise from substituting co-referring terms into necessity contexts is a plausible solution to his particular puzzle provided that the required scope distinctions can be justified. However, in spite of the merits of a Russellian approach that could justify the needed scope distinctions, such a proposal would still not go far enough (as I have already suggested by some of the above remarks) to dispel the air of paradox surrounding the necessary yet contingent status of identity statements containing a definite description. I believe that an adequate solution can only be had in the context of a more thorough account of the different meanings that can attach to a definite description.[12] Russell's theory of descriptions together with the recognition of Donnellan's distinction between the referential and attributive uses of a definite

---

12. However, I hasten to add that, even though I will subsequently find fault with Kripke's rejection of Donnellan's referential-attributive distinction, the general strategy that I will ultimately suggest as a possible way of developing an adequate solution to the problem of identity does not depend on the enriched account of definite descriptions I outline below.

description is, I believe, necessary but is still not sufficient to account for all uses of a definite description. Both Russell's and Donnellan's analyses can, nevertheless, help to motivate a more inclusive account of the content of statements containing a definite description. The occurrence in a sentence of a definite description can actually serve a number of assertory purposes. In very general terms, when one uses a statement containing a definite description to make an assertion, there are, ignoring variations in the scope of operators occurring outside the description, five different ways of understanding what is being said. The content of a statement having the gross grammatical form, "The F is G," could be revealed, in terms of logical form, by rendering the statement as one of the following formal expressions (where "F" and "G" are predicates, "n" is a proper name, and "d" is the definite description, "the F").

(1) Gd

(2) $(Fn \ \& \ (y)(Fy \rightarrow n = y))$

(3) $(\exists x)((Fx \ \& \ (y)(Fy \rightarrow x = y)) \ \& \ Gx)$

(4) $(x)((Fx \ \& \ (y)(Fy \rightarrow x = y)) \rightarrow Gx)$

(5) $(x)(Fx \rightarrow Gx)$

An example of a sentence containing a definite description is "The President of the United States is a Republican." If the description "The President of the United States" is being used in that statement merely to refer to a particular person the speaker has in mind, then the description is functioning more or less as a proper name and the sentence has the form of (1). Someone might utter the sentence understood in this way if, for example, the speaker was intending to say of George W. Bush that he is a Republican. However, if the description is functioning in the manner envisioned

by Russell, then the description is being used to assert both that there is a unique person who is President of the United States and that that person, whoever he happens to be, is a Republican. In this case, the sentence has the form of (3). The sentence might have this meaning if, for example, it was to be uttered by a speaker who wished to inform an audience about current U.S. presidential politics. The definite description in the sentence could, however, be serving the purpose of allowing the speaker to assert that anyone who is President of the United States, and is unique in so being President of the United States, is a Republican. In such a situation, the speaker would not be asserting (but might be presupposing) that there is someone who satisfies the description "the President of the United States." If the sentence is being used to make such a claim, then the sentence can be interpreted as having the form of (4). A person might intend that his utterance of the sentence be understood in such a fashion if, for instance, he believed that, given the current political climate, only a Republican could be elected President. It is possible, though, that a speaker may use the description in the statement as a means merely of indicating that the class of Presidents of the United States is included in the class of Republicans. In such a situation, the speaker would not be asserting (but might be presupposing) that the class of Presidents of the United States is not empty. In this case, the sentence would need to be interpreted as having the form of (5). A speaker might utter the sentence with that meaning if he believed, for instance, that some constitutional law required all those who are elected to the office of the presidency to be Republicans. Alternatively, a speaker who utters the sentence "The President of the United States is George W. Bush" would probably be using the description only to attribute to George

W. Bush an outstanding unique feature. In such a case, the sentence uttered would need to be interpreted as having the form of (2).[13]

Russell's contextual analysis of the content of definite descriptions could only be applied to definite descriptions when they had an attributive role in the sentences that contained them. Russell's preferred rendering of such a sentence as "The President of the United States is the person who garners the most votes in the electoral college" would be "There is a unique person who is President of the United States and a unique person who garners the most votes in the electoral college, and the former is identical to the latter." We would still at least be within the spirit of Russell's analysis if we instead interpreted the sentence as "Anyone who is President of the United States and is unique in being President is also a person who garners the most votes in the electoral college and is unique in so doing." (The latter rendition differs from the former rendition only in being a universal instead of an existential generalization.) Russell's analysis can thus be brought to bear on sentences containing definite descriptions that can be interpreted as having the form of either the third or the fourth of the above cases. Russell's analysis is also applicable to

---

13. Besides being cognizant of the differences in the use of a definite description as represented by the above five forms, one also needs to pay particular attention sometimes not only to the scope of an operator occurring outside a definite description but also to the scope of a quantifier occurring within a definite description. Morris (*Understanding Identity Statements*, p. 96) mentions a fallacious inference that turns on an ambiguity in the scope of a quantifier.

"... Ronald Reagan is now one and the same individual as the President of the United States. The President is elected every four years. And, surprisingly successful as he may be, it is not true of Ronald Reagan that he is elected every four years."

The conclusion that Ronald Reagan is elected every four years is unwarranted because the sentence "The President is elected every four years" should not be read as "There is one and only one person who is President, and that person is elected every four years" but rather as "Every fourth year, there is one and only one person who is elected President that year." The quantifier "every" should thus be understood as having wide scope, while the description should be understood as having narrow scope. (In addition, the predicate "is elected President on a year" needs to be understood differently than the predicate "is President and is elected on a year.")

sentences containing a definite description that can be interpreted as having the form of the second of the above five cases.[14] The sentence "The President of the United States is George W. Bush" can, after all, be rendered as the sentence "There is a unique person who is President of the United States and George W. Bush is that person," which in turn is equivalent to the sentence "George W. Bush presides as President of the United States and is unique in so doing."[15] Identity statements that contain definite descriptions that are being used in the fashion in which the description is used in sentences having the form of the fifth of the above cases present no special problem to an account of identity since, when properly understood, they do not actually contain the identity predicate. The sentence just mentioned about the President and the electoral college would be understood as "Anyone who presides as President of the United States garners the most votes in the electoral college" if the descriptions are being used, as in the fifth case, to refer to two classes of individuals. However, Russell's theory of descriptions did not apply to those cases that involve the purely referential use of the definite description. An objectual identity theorist who believes in the "contingent and necessary" status of true identity statements will thus be forced to provide a nonRussellian account of the content of those identity statements that contain definite descriptions used in the manner of sentences having the form of the first of the above cases. The distinction between, on the one hand, the use of the definite description as it occurs in sentences having the form of the first of the above cases and, on the other hand, the use of the definite description as it occurs in sentences having the form of the third or fourth cases is derived from Donnellan's

---

14. I have actually already assumed this to be the case in obtaining sentence E and sentence E' from the sentence "Benjamin Franklin was the first postmaster general."

15. This is the case since "$(\exists x)((Fx \ \& \ (y)(Fy \to x = y)) \ \& \ x = n)$" is equivalent to "$(Fn \ \& \ (y)(Fy \to n = y))$."

distinction between the referential and attributive uses of a definite description. Since the referential use of a definite description is very much like (although not quite the same as) the use of the description as a proper name, most likely an appropriately enriched account of definite descriptions will apply not only to statements having the form of the first, third, or fourth of the above cases but also to identity statements containing proper names.

In light of the more complete analysis of definite descriptions presented above, the argument I presented previously for the weak necessity of the sentence "Benjamin Franklin was the first postmaster general," and, by extension, for the weak necessity of any identity statement, can be shown to be faulty. I argued above that, despite Kripke's resolution of his particular puzzle, the following troublesome inference, which resembles Kripke's problematic inference, appears reasonable.

Benjamin Franklin was the first postmaster general.

Necessarily, the first postmaster general was the first postmaster general.

Therefore, necessarily Benjamin Franklin was the first postmaster general.

A speaker who uttered the first premise would most likely only be using the description to ascribe to Benjamin Franklin a property that distinguished him from other people. Since the sentence "Benjamin Franklin was the first postmaster general" is obviously equivalent in meaning to the sentence "The first postmaster general was Benjamin Franklin," the first premise should be interpreted as having the form of (2) in the above forms. The second premise is clearly true only on the condition that the sentence "The first postmaster general was the first postmaster general" is interpreted as a tautology having the form of (4) in the above list of forms.

The necessity operator in the conclusion seems quite naturally to have wide scope: anyone who asserted the conclusion would be saying, in effect, that the property ascribed to Benjamin Franklin by the first premise necessarily pertains to Benjamin Franklin. Thus, the conclusion needs be interpreted as having the same form, except for the necessity operator, as the first premise. The above argument can therefore be rendered as the following formal argument.

$(Pb \ \& \ (y)(Py \rightarrow b = y))$

$\sim\lozenge\sim(x)(y)(((Px \ \& \ (z)(Pz \rightarrow x = z)) \ \& \ (Py \ \& \ (z)(Pz \rightarrow y = z))) \rightarrow x = y)$

$\therefore \ \sim\lozenge\sim(Pb \ \& \ (y)(Py \rightarrow b = y))$

Rendered in the above symbolic form, the argument is clearly invalid. The necessity in the conclusion is unwarranted since the premises are not both necessary propositions. (If the necessity operator in both sentences designates logical necessity, then it is clear why the inference fails, since the substitution of one co-referring term for another in a tautology will not always produce a tautology.) The conclusion is false since the first premise is a metaphysical (and logical) contingency inasmuch as the particular person (namely, Benjamin Franklin) who in reality satisfies the description "the first postmaster general" need not necessarily have been the one who was the head of the first U.S. post office, in spite of the fact that whoever was the first postmaster general was, of course, the person who was the head of the first U.S. post office.[16]

---

16. It is best to interpret the necessity in the second premise as logical necessity and the necessity in the conclusion as metaphysical necessity. (So the argument may involve equivocation.) Some objectual identity theorists who hold the view of identity as a necessary metaphysical relation may want to appraise the original natural language argument by interpreting the explicit necessity operator in the conclusion as having small scope. For these identity theorists who interpret things that way, the conclusion is actually true but is no different in content than the first premise, since that

Regardless of what theory of singular terms ultimately proves to make the best sense out of identity statements containing a definite description and successfully avoids identity paradoxes involving the modalities, any completely articulated objectual account of identity must still surmount all those previously mentioned difficulties that are associated with an objectual analysis of identity. Even if such a well-formulated account of identity does allow the objectual identity theorist to maintain that all true identity statements are in some sense necessary, even when they are logically contingent, that account must go on to solve the problem of identity not only for identity statements containing a description but also for identity statements containing only proper names. Hence, the objectual identity theorist will still need to explain why such statements as "Hesperus is Phosphorus," "Mark Twain is Samuel Clemens," and "Cicero is Tully" are not trivial, are in some sense not necessary, and are different in content than, respectively, the sentences "Hesperus is Hesperus," "Mark Twain is Mark Twain," and "Cicero is Cicero" in spite of the fact that the former type of statements seem to be different in nature than the latter type of statements and yet an utterance of any of these statements seems more clearly to amount to only an assertion that a relation holds between a particular object and itself. It is to the issue of how Kripke specifically regards both kinds of identity statements that I must now turn.

As I mentioned at the beginning of the present chapter, Kripke maintains that identity is a metaphysical relation that necessarily holds between every object and itself. Any true identity statement expresses a necessary truth, at least to the extent that what the statement says about the referents of its terms (i.e., the particular

---

premise in effect contains an implicit necessary operator in the same position. Under such an interpretation, the argument is valid but circular, and the second premise plays no role in the inference.

individual referred to or uniquely described by each of its terms) is true of that individual and itself as a matter of metaphysical necessity.   However, because an identity statement may contain a definite description or two distinct proper names, a true identity statement may not overall be necessary (since it may not express just a necessary metaphysical truth) or may not be necessary in another sense of necessity. A true identity statement may thus on the whole be contingent and, as a result, not known to be true prior to experience, even though what it says about the relationship between the referents of its terms is metaphysically necessary. As I also indicated earlier, Kripke's solution to the problem of identity does not depend in any fundamental way upon interpreting definite descriptions according to Russell's theory of descriptions. Nevertheless, Kripke does believe that it is in virtue of an individual satisfying the Russellian truth conditions (including the uniqueness condition) associated with a definite description that the description successfully attaches to the individual.[17]   Moreover, Kripke understands the *de re* necessity always associated with a true statement of identity to be akin to the necessity indicated by an implicit or explicit necessity operator, having narrow scope, associated with identity statements that are analyzed in the manner suggested by Russell. (That is why Kripke calls the narrow scope of the necessity operator the "*de re* scope" of the operator.) In addition, Kripke does not accept as a valid semantic distinction Donnellan's distinction between the referential and the attributive uses of a definite description.   (The rejection of that distinction is an issue that I will discuss later in the present chapter.) Kripke thus does not rely upon the apparatus suggested above for analyzing identity statements containing a definite description.

---

17. There are, though, singular terms that look like definite descriptions but do not attach to their referents by describing something uniquely. Such expressions as "the Statue of Liberty" and "the wicked witch of Wescoe Hall," although perhaps suggestive of their referents, will for Kripke acquire their reference in the manner that proper names acquire their reference.

Instead, Kripke's solution relies upon seeing proper names and definite descriptions as two different types of designators based on how such terms refer in counterfactual situations. Because Kripke understands proper names to be one kind of designator and definite descriptions to be another kind of designator, Kripke treats identity statements containing only proper names in a different manner than he treats identity statements containing at least one definite description. Kripke contends that, with regard to at least one sense of necessity, identity statements of the former kind are, when true, necessarily true and, when false, necessarily false and are never contingently true or contingently false. The necessity that always pertains to true identity statements containing only proper names is due to the fact that a proper name functions as what Kripke calls a "rigid designator."[18] A singular term functions as a rigid designator if the term does not refer to one thing in one possible world (description of a counterfactual state of affairs) and another thing in another possible world. A term rigidly denotes if it denotes the same thing in every possible world in which it has a referent. In contrast, a singular term functions as a nonrigid designator when it denotes one thing in one possible world and a different thing in another possible world. A nonrigid designator does not maintain the same reference from one possible world to another. Since definite descriptions, or at least most definite descriptions, function as nonrigid designators, identity statements containing a definite description are typically not true in every possible world. That is why Kripke contends that most true identity statements containing a definite description are only contingently true, even though the identity relation holds between any individual and itself in every possible world.

---

18. See Lecture I in *Naming and Necessity*.

Unfortunately, Kripke's notion of rigid designation (as well as its significance) and why Kripke regards proper names as denoting rigidly can easily be misunderstood and will need to be clarified. In claiming that proper names denote rigidly, Kripke does not mean to claim that speakers in other possible worlds must always use the proper names we use in the actual world to refer to the same things we refer to in using those names in the actual world. Kripke would not deny, for instance, that people in some other possible world might use the name "Hesperus" to refer to something other than the planet Venus, or Phosphorus. (In other words, Kripke acknowledges that our words might have had a different meaning and use other than those they actually do have.) As a consequence, for Kripke, an identity statement that expresses a truth when uttered by us in the actual world might not express a truth when uttered by a speaker in some other possible world. "Hesperus is Phosphorus" is true (and expresses a necessary metaphysical truth) according to our use of the names "Hesperus" and "Phosphorus" in the actual world, but in some other possible world that identity statement might not be true given the way in which speakers in that possible world use those names in that other world. Rather, in claiming that proper names denote rigidly, Kripke means to claim that speakers in the actual world use a proper name to denote a particular individual in the actual world and continue to use that proper name to denote that same individual when considering counterfactual states of affairs in which the term refers and that individual thus exists. So, for Kripke, it is our use of a rigid designator to refer in the actual world that determines how we use that designator to refer in other possible worlds in which the term refers at all. Identity statements containing only rigid designators that are true for us in the actual world are therefore true for us in every possible world. In contrast, our use of a nonrigid designator to refer to a particular individual in the

actual world does not, according to Kripke, determine to what we (or, for that matter, anyone in some other possible world) might use the description to refer in considering a counterfactual state of affairs. As a result, identity statements containing a nonrigid designator that are true for us (i.e., that we use to assert that a necessary metaphysical relation holds between an object and itself) in our talk about the actual world may not be true for us (i.e., may not be used by us to make the same assertion) in our talk about some other possible world.

For Kripke, the proper name "Benjamin Franklin" refers to Benjamin Franklin in a real-world description and does so in every counterfactual description in which the term refers, which will always still be a description in which Benjamin Franklin exists. Kripke regards as mistaken the view that in other possible worlds such a name does not refer to what it does in the actual world but instead refers to someone, or something, in that possible world who is, or that is, similar in all important respects to the referent of the name in the actual world. Kripke thus rejects David Lewis's counterpart theory that construes proper names as referring in other possible worlds to the counterparts, in those other worlds, of what they refer to in the actual world.[19] For Kripke, when we make reference to Benjamin Franklin, for instance, in some counterfactual statement (such as, "Benjamin Franklin would have made a fine President") we are still referring to Benjamin Franklin and not to some person who in that counterfactual state of affairs resembles Benjamin Franklin in enough of the details to enable the name "Benjamin Franklin" to be attached to him. If the name "Benjamin Franklin" did not denote rigidly, then it would make sense to talk about a possible world in which Benjamin Franklin was not Benjamin Franklin. However,

---

19. See David Lewis's "Counterpart Theory and Quantified Modal Logic" in the *Journal of Philosophy*.

Benjamin Franklin could have been no individual other than Benjamin Franklin, so talk about such a possible world is nonsense. Thus, the name "Benjamin Franklin" denotes rigidly, and the same can be said for any proper name.

Kripke sees proper names as denoting rigidly due to the manner in which proper names acquire their reference. With regard to proper names, Kripke subscribes to a theory of reference that has been called "the causal theory of names."[20] According to Kripke, an individual is given a proper name as a result of a special act involving reference to the individual and the use of the name, and it is in virtue of this "christening" that the custom of using the name to refer to the individual is inaugurated. The particular individual that gets named may be indicated to an audience through an act of ostension or by the use of a definite description. In cases where a description is used to single out an individual, the individual does not actually need to satisfy the description as long as the speaker succeeds in making clear to his audience exactly what individual is being given the name. (In most cases where a description is used to specify the individual that gets named, the audience nevertheless believes that the individual satisfies the description.) The name-giver's habit of using the name to refer to the individual named is picked up by the audience and is subsequently transferred to other people who come to understand what the referent of the name is. Because the custom of using the name to refer to a particular individual can then be transferred to subsequent generations, people who are far removed from the original users of the name with respect to both time and distance can come to use the name, or some corrupted version of it, perhaps in a different language, with the same intention to refer as that once had by the original name giver. Proper names are causally linked in this manner to their referents, and the causal link

---

20. See Gareth Evans' essay "The Causal Theory of Names."

between a name and its referent is maintained even when the name is used to refer in talk of counterfactual situations. A definite description, by contrast, is used successfully to refer only if its referent satisfies the description (or at least people believe that its referent satisfies the description). Since what satisfies or is considered to satisfy a description may well vary from one state of affairs or description of a state of affairs to another, definite descriptions do not in general keep the same reference when changes occur in the actual world or when other possible worlds are considered. Thus, definite descriptions denote nonrigidly, while proper names denote rigidly.[21]

Given that all the singular terms in a true identity statement are rigid designators, Kripke contends that the weak necessity of the statement follows. Kripke argues that, if it is true that **a** is identical to **b** and "**a**" and "**b**" are both rigid designators, then, since "**a**" and "**b**" refer to the same thing in the actual world and "**a**" refers to the same thing in every possible world in which it has a referent and "**b**" refers to the same thing in every possible world in which it has a referent, "**a**" and "**b**" must both refer to one and the same thing in every possible world in which one, and thus both, of the terms have a referent, which is a world in which **a** and **b** exist.[22] Thus, if **a** is identical to **b**, then necessarily **a** is identical to **b**. In other words, if "**a** = **b**" is true, "**a** = **b**" is necessarily true.[23] Hence, any identity statement containing only

---

21. Actually, it is more accurate to say that all proper names are rigid designators while the vast majority of definite descriptions are nonrigid designators. For Kripke, some definite descriptions, such as "the sum of two and two," denote rigidly due to the necessity that Kripke alleges attaches to mathematical truths. In addition, Kripke admits that there are expressions that look like proper names, but are not, and are nonrigid designators. The term "Jack the Ripper," for instance, appears to be a name of a particular individual whose precise identity was never discovered. However, given the circumstances under which the term entered the vernacular, it actually is used as a way of talking about the person, whoever he was, who committed most or all of a series of grisly murders in the Whitechapel area of London's East End from April to August of 1888. The term is therefore functioning as a disguised definite description and not a proper name.

22. "Identity and Necessity," in Milton Munitz's *Identity and Individuation*, 154.

23. Kripke would therefore accept some version of the Barcan formula (that is, something like the formula "**a** = **b** → ~◇~**a** = **b**") endorsed by Ruth Barcan Marcus in "Modalities and Intensional Languages."

rigid designators that is true in the actual world will be true in every possible world in which the referent of its terms exists. However, if either "**a**" or "**b**" is a nonrigid designator, then, even if the referents of the two terms are identical in the actual world, it is not the case that what one term refers to is the same as what the other term refers to in every possible world, since in some possible world the two designators are not used to specify the same object in that possible world. Therefore, if either "**a**" or "**b**" is a nonrigid designator, then it is not the case that, if "**a** = **b**" is true, then "**a** = **b**" is necessarily true.

In order to understand fully Kripke's position in regard to both kinds of identity statements (i.e., those containing only proper names and those containing a definite description), not only is it important to understand Kripke's account of rigid designation, but the crucial distinction Kripke draws between necessity and a priority also needs to be recognized. A statement is necessarily true, for Kripke, iff both it is true and its truth-value could not have been otherwise. A statement that is necessarily true is true in all possible worlds (i.e., it remains true under all possible counterfactual descriptions).[24] A statement is then contingently true iff both it is true and its truth-value could have been otherwise. A contingently true, as well as a contingently false, statement is true in some possible worlds (i.e., under some counterfactual descriptions) and is false in other possible worlds (i.e., under other counterfactual descriptions). In contrast, a statement is an a priori truth, for Kripke, iff both it is true and it can be known to be true independent of experience.[25] It does not matter that a

---

24. In the case of weak necessity, the definition is somewhat different: a statement is weakly necessarily true iff both it is true and its truth value could not have been otherwise in any situation in which its terms have reference. In other words, a statement is considered to be weakly necessary iff it is true in all possible worlds in which its terms refer.
25. There are, of course, various sorts of difficulties that attend the notion of knowing a statement to be true independently of experience, but I believe the primary purpose of my exposition of Kripke's distinctions will best be served by following Kripke's lead and ignoring these difficulties.

statement is in fact never known to be true independent of experience; it will be an a priori truth if it *can*, in principle, be known to be true independent of, or prior to, experience. This means that, for a given statement, if its truth can be recognized without consulting one's experience of the world, then that statement is an a priori truth.[26] In contrast, a statement is an a posteriori truth iff both it is true and knowledge of its truth depends on experience. A statement is thus true a posteriori if its truth can only be recognized after, rather than prior to, consulting one's experience of the world. Since, for Kripke, an identity statement's necessity is a matter of the way things must be and an identity statement's a priority is a matter of what can be known prior to experience, in drawing the distinction between necessity and a priority in this manner with regard to identity statements Kripke is in effect making a respective distinction between two types of necessity, metaphysical necessity and epistemological necessity.[27] For Kripke, true identity statements, although weakly necessary, may not be known to be true a priori. Thus, identity statements, when true, are always metaphysically necessary (or at least always express claims about the referents of their terms that are metaphysically necessary) but are not always epistemologically necessary.[28] According to Kripke, it is because philosophers have not recognized the difference between what must be true from a metaphysical standpoint and what must be true from an epistemological standpoint that they have been misled into accepting untenable positions regarding identity.

---

26. Kripke's notion of a priority appears to be more in line with that of Hume rather than with that of Kant. A priori truths for Kripke, I take it, are not what can be transcendentally derived as the preconditions necessary for the very possibility of experience. Instead, a statement expresses an a priori truth for Kripke if the recognition of the truth of the statement does not depend on experience, regardless of whether what is expressed employs concepts that can only be derived from experience.

27. This is a distinction that I have of course alluded to, but not introduced explicitly, in the previous discussion.

28. Kripke thus identifies the weak necessity of any true identity statement as weak metaphysical necessity and regards it as *de re* necessity, while he identifies a priority as epistemological necessity and regards it as *de dicto* necessity.

As will soon emerge in my subsequent discussion of Kripke's proposed resolution of the problem of identity, the success of Kripke's solution requires that necessary truths and a priori truths comprise two separate domains. Kripke contends that neither the class of necessary truths nor the class of a priori truths is wholly contained within the other. Kripke notes that, for instance, there are mathematical statements (such as, for all we know, Goldbach's conjecture) that, if true, are necessarily true (i.e., true in all possible worlds) but are not known to be true (or false) a priori. If to know a priori that a mathematical proposition is true is to be able to prove it in some fixed system, then what Godel showed, in effect, was that there will always be mathematical statements that are necessarily true but are not knowable a priori to be true. Certain statements that are used to make assertions about an object's essential properties provide other examples of necessary a posteriori truths. In Kripke's lectures, he once asked whether or not the table before him could have been made of ice.[29] It is possible, Kripke claimed, that the table is really made of ice in spite of the fact that tables are not normally made of ice and the table does not appear to be made of ice. It requires an empirical investigation to discover what the table is actually made of, but that particular table is necessarily made of that particular substance, whatever it happens to be, out of which the table is in fact made. For Kripke, since an object cannot be made of a substance other than the one of which it is in fact made, the property of being composed of a certain substance (material) is one of the essential properties of an object. According to Kripke, a thing will have all its essential features in every possible world in which the thing occurs.[30]

---

29. Lecture III in *Naming and Necessity*.
30. According to Kripke, if we were to suppose that an object could in some other possible world be made of a different substance than the one of which it is actually made or could possibly have some other essential attribute other than the one it actually has, then we would have to imagine a situation in which a rigid designator referring to that object in the actual world would not be referring to the same thing in another possible world, a world in which, supposedly, a referent for the designator

Kripke therefore regards the statement that a particular object is made of a particular substance as expressing, when true, a metaphysically necessary truth. If the table before him is not made of ice but is, in fact, made of wood, then the statement "This table is made of wood" is one whose truth, although necessary (from a metaphysical standpoint), is not knowable independent of experience. Thus, if the statement is true, it is another example of a necessary truth that is not known a priori.

Kripke contends that there are also a priori truths that are not necessary, although I must confess that I do not find Kripke's alleged example of such a truth to be very convincing. According to Kripke, if "S" is a proper name (rather than a definite description) of the standard meter bar kept in Paris and the reference of the word "meter" is fixed by the nonrigid description "the length of the standard meter bar," then the sentence "S is one meter long" is knowable a priori but is not necessary. The sentence appears to be an a priori truth since the reference of "meter" was defined as the unit of linear measure equal to the length of the standard meter bar and thus the sentence is evidently known to be true (in any world in which S exists) independent of any empirical investigation. The sentence appears not to be necessary, though, since evidently there is no metaphysical reason why the standard meter bar must have the particular length it in fact has, a length that is standardly referred to as one meter. However, one can judge Kripke as having successfully provided an example of a metaphysically contingent a priori truth only if one neglects (or rejects) the distinction between an attributive use and a referential use of the description associated with the word "meter." At the time when the word "meter" was defined, the reference of "meter" was fixed by an attributive use of the nonrigid description "the length of the

<hr />

nevertheless still exists. Since this latter is impossible, an object must have all its essential attributes in every possible world in which it exists. In addition to the substance of which a thing is composed, Kripke also regards the origin of a thing to be one of its essential features. Kripke's essentialism is consequently a substance-origin essentialism.

standard meter bar." This means that, provided that a length of one meter was not defined in any other way at that time (which, I believe, is safe to say), whatever was the length of the standard bar at that time is one meter. However, the declaration that S is one meter long represented at that time a stipulation, rather than an assertion of an a priori truth. Furthermore, when someone today says, "S is one meter long," he is not saying something to the effect of "S has a length of one unit of a measure equivalent to its length." Rather, he has a particular length in mind when talking about a meter (a length he would refer to by a referential use of the description "the length of the standard meter bar"), and that length he may believe is, or was at one time, the length of the standard meter bar in Paris. So, when today someone says, "S is one meter long," his statement is indeed not necessary, but is also not known to be true a priori. Notwithstanding the failure of the sentence "S is one meter long" to express a metaphysically contingent a priori truth, I do agree with Kripke that the category of the a priori may well overlap but not be contained within the category of the metaphysically necessary. However, anyone who advances such a thesis or who contends otherwise will need to give cogent arguments in support of his position, since neither view is obviously true.

Now that I have discussed Kripke's distinction between rigid and nonrigid designators and his distinction between necessary and a priori truths, and the relationship between the latter pair, I can present the account of identity statements that is at the heart of Kripke's solution to the problem of identity. As I indicated above, Kripke's position regarding identity statements containing only proper names is that such statements are, when true, always weakly necessary but are not always known to be true a priori. Since proper names denote rigidly, such identity statements are, when true in the actual world, true in all possible worlds in which there occur the

particulars to which their terms are used to refer in the actual world. As a result, true identity statements containing only proper names are always weakly necessary (and thus are always metaphysically necessary), but they are not always known a priori to be true (and thus are not always epistemologically necessary) because their truth cannot always be recognized without consulting experience. Even though the identity relation necessarily obtains whenever it in fact obtains, the recognition of that relationship, just like the recognition of an essential attribute, may require an empirical investigation. Hence, we can, on the basis of our empirical knowledge of the world, come to accept or reject an identity statement as we would any logically contingent proposition and yet acknowledge that the identity statement is, if true, necessarily true. Furthermore, a statement of the form $\ulcorner \alpha = \alpha \urcorner$ will always be knowable a priori (and thus be of no informational value), whereas a statement of the form $\ulcorner \alpha = \beta \urcorner$, if true, may only be known to be true a posteriori (and thus be of informational value). In this way, Kripke's account of identity statements containing only proper names apparently satisfies the difference in content requirement. Since "**a** = **b**" may be a contingent statement in the sense that, even if true, it may not be knowable a priori to be true and thus may not be epistemologically necessary (in spite of the fact that, when true, such statements express a metaphysically necessary truth), the nonnecessity requirement is also apparently satisfied. If we equate a statement's triviality with its a priority and a statement's nontriviality with its a posteriority, then Kripke's account appears to satisfy the nontriviality requirement, since the truth of "**a** = **b**" may well be knowable only a posteriori, while the truth of "**a** = **a**" is known a priori. Thus, Kripke's position can be succinctly summarized: Kripke contends that identity statements that contain only proper names (which are identity statements proper) are being used to claim that a metaphysically necessary relationship holds

between an object and itself, a relationship which we may nevertheless, in certain particular cases, only be able to recognize as holding after consulting our experience of the world.

Before ending the present section, I do not want to overlook the issue of why, according to Kripke, there is controversy surrounding the status of identity statements (containing only proper names) as necessarily true when true. A discussion of the controversy should help further clarify Kripke's views on identity. The suggestion that any identity statement containing purely denotative terms is in some sense necessarily true, if true at all, appears to many people to be counterintuitive. Kripke sees two reasons why people have thought that most such identity statements, even when true, must be empirically contingent propositions. One reason, according to Kripke, that identity statements have been considered to be contingent is due to the fact that some people have opted for the metalinguistic interpretation of the identity predicate. As a result, they have considered such statements as "Hesperus is Phosphorus" and "Cicero is Tully" as respectively saying, in effect, "The name 'Hesperus' and the name 'Phosphorus' are co-referential" and "The name 'Cicero' and the name 'Tully' are co-referential." Since it is obviously the case that "Hesperus" and "Phosphorus" on the one hand and "Cicero" and "Tully" on the other hand need not necessarily refer to the same thing or to the same person and their co-referentiality can only be established by consulting the empirical facts, people have felt compelled to regard the corresponding identity statements as being empirically contingent. However, the fact that "Hesperus" and "Phosphorus" could have referred to different heavenly bodies (i.e., the fact that people in some counterfactual description use the names to refer to different heavenly bodies) is, for Kripke, quite irrelevant to the truth of the identity statement since he regards the statement "Hesperus is Phosphorus" as

saying something about Hesperus and Phosphorus and not about the names "Hesperus" and "Phosphorus." An identity statement, even when it contains two distinct terms, is always about the reference of its terms and not about the terms themselves. So, the fact that, in some counterfactual descriptions of the world, speakers would use a pair of names other than "Hesperus" and "Phosphorus" in an identity statement to express the fact that Hesperus is Phosphorus (i.e., what is expressed by the sentence "Hesperus is Phosphorus" in the actual world) has no bearing on the fact that Hesperus is Phosphorus and thus no bearing on the possible truth or falsity of the identity statement "Hesperus is Phosphorus" as speakers in the actual world understand that statement.

The other reason Kripke gives for why it is that people have mistakenly thought that identity statements are typically contingent propositions is that many identity statements are accepted only on the basis of certain matters of fact and it has always seemed possible for these matters of fact to have been other than what they are. Some would thus argue that the statement "Hesperus is Phosphorus" is only contingently true since it came to be accepted as true only after the orbit of the planet Venus was fully known and it is surely conceivable that Venus could have had some other orbit so that it did not occupy both the position of Hesperus in the evening sky and the position of Phosphorus in the morning sky. It is indeed conceivable that a bright celestial object seen in the early morning sky could have come to be referred to by the name "Phosphorus" and a bright celestial object seen in the early evening sky could have come to be referred to by the name "Hesperus" and yet the celestial objects referred to by the two names be distinct heavenly bodies. We can imagine, for instance, Venus and Mars having different orbits and different orbital velocities so that Venus never made its appearance in the evening sky and Mars took the position

in the sky actually occupied by Hesperus and was just as bright as Hesperus. However, Kripke argues that such a counterfactual situation would not be a situation in which Hesperus is not Phosphorus. It would be a situation, no doubt, where the name "Hesperus," if it were used by people in that situation to refer to a bright object seen in the early evening sky in the position actually occupied by Hesperus, would not refer to the same thing as referred to by the name "Phosphorus," if that name were used by people in that situation to refer to a bright object seen in the early morning sky in the position both actually and counterfactually, in this case anyway, occupied by Phosphorus. In such a counterfactual situation (possible world), it would not be true to say that both names would be used by people in discourse to refer to Venus, since people in that possible world would use "Phosphorus" to refer to Phosphorus and use "Hesperus" to refer not to Hesperus but to Mars. Nevertheless, such a situation would not be a situation where the planet we actually refer to in using the name "Hesperus" is not the same as the planet we actually refer to in using the name "Phosphorus." Kripke acknowledges that "Hesperus is Phosphorus" was not known a priori to be true, but, in discovering the truth of the statement after making careful observations of the planets, astronomers had discovered what is a metaphysically necessary truth.

The case of the identification of Cicero as Tully provides us with yet another example of how people have, in a similar fashion, mistakenly come to view identity statements containing only proper names as contingent. We have come to use the name "Cicero" to refer to the author of certain Latin texts, and we have come to use the name "Tully" to refer to the Roman orator who denounced Cataline. Furthermore, because of our beliefs concerning certain historical facts, we have come to believe that one and the same person both authored these Latin texts and denounced Cataline

in oratory. We thus accept "Cicero is Tully" as a true statement of identity. Since it is in fact true, Kripke takes it to be necessarily true. However, some would contend that the historical facts could have been different and it could have been the case that the Roman orator who denounced Cataline was not the person who authored the texts normally attributed to Cicero. Those who would make such a contention might then go on to argue that the identity statement "Cicero is Tully" is by no means necessary since the persons referred to by the two names need not have been the same. Kripke responds to such critics by claiming that those who would advance such a line of argumentation have misunderstood both what is being said in asserting that Cicero is Tully and how the reference of the names "Cicero" and "Tully" is fixed. The sentence "Cicero is Tully" is not about the names "Cicero" and "Tully" and is not about whoever happens to be, in the actual world or in another possible world, the one who authored certain Latin texts and whoever happens to be, in the actual world or in another possible world, the one who denounced Cataline in oratory. In using the names "Cicero" and "Tully" in the statement "Cicero is Tully," we mean to refer rigidly to a particular person whom we have in mind, a person who is necessarily identical to himself. If tomorrow we were to discover that Cicero did indeed write the texts that have been attributed to him but he did not denounce Cataline in oratory, we would still not reject the identity statement "Cicero is Tully." Instead, under such circumstances, we would only be forced to find some description other than "the Roman orator who denounced Cataline" that would fix the reference of the name "Tully." For Kripke, the names "Cicero" and "Tully" are both rigid designators of the person whom we have identified in the actual world as, respectively, the author of certain Latin texts and the Roman orator who denounced Cataline. Kripke contends that we may sometimes use a nonrigid definite description to fix the reference of a

rigid designator. Thus, the reference of "Cicero" is fixed by the nonrigid description "the Roman orator who denounced Cataline" when we use that description to identify the individual who bore the name "Cicero." However, the name "Cicero" should not be taken to be synonymous with the description, for Kripke rejects the description theory of names. Accordingly, "Cicero" is not being used to denote whoever would have, in some other possible world, denounced Cataline in oratory if Cicero had in fact not been the one to have done so. The name "Cicero" instead refers rigidly to the person we refer to in the actual world when we use the description "the Roman orator who denounced Cataline." The reference of "Tully" is fixed in an analogous manner, and, since both "Tully" and "Cicero" denote rigidly and "Cicero is Tully" is true in the actual world, "Cicero is Tully" is necessarily true because it is true in every possible world in which Cicero and Tully exist. Hence, even though the sentence "Cicero is Tully" is not known a priori to be true, what that sentence says about the world is nonetheless a metaphysical necessity.

4.2   Problems with Kripke's analysis

The account of identity discussed above thus constitutes Kripke's solution to the problem of identity. An identity statement that contains a definite description typically expresses what is a metaphysical contingency. Nevertheless, what a true identity statement says about the referents of its terms is a metaphysical necessity (which is more clearly seen when a true identity statement contains only proper names) in spite of the fact that such a statement is often not epistemologically necessary. Has Kripke solved the problem of identity? I will argue below that Kripke's account of the nature of identity cannot provide us with an understanding of the logical content of identity statements because it does not include an account of the

notion of an essential property, the notion of a purely referential context, and the notion of a rigid designator that are free of the notion of identity. However, before presenting my primary objections to Kripke's account of identity, I need to make some preliminary critical remarks regarding Kripke's analysis.

Appearances to the contrary, Kripke has not actually provided us with an understanding of the identity relation as a metaphysical relation that necessarily holds between every object and itself. For Kripke, what any true identity statement says about the particular objects that are the referents of its terms is metaphysically necessary, even if the identity statement as a whole is not true in every possible world in which its terms refer and is thus not metaphysically necessary. Kripke accounts for the metaphysical necessity of true identity statements containing only rigid designators by the fact that such designators refer to the same thing in every possible world in which they refer at all, which means that such identity statements are true in every possible world in which the referents of their component rigid designators exist. Kripke accounts for the epistemological contingency of some true identity statements containing only rigid designators by their logical contingency or by the fact that they are not known a priori to be true. However, the weak necessity of such identity statements would seem to indicate that the identity predicate, like a rigid designator, does not change its interpretation (or semantic role) across possible worlds. The rigid interpretation of the identity predicate suggests that it is a logical predicate, a suggestion that is reinforced by the epistemological necessity of the identity statement "$a = a$" (where "a" is a rigid designator). As a consequence, I am not sure what it means to say that identity is a metaphysical relation.

In addition, although an examination and critique of Kripke's causal theory of names would go beyond the scope of the present discussion, I should point out here

that much of the success of Kripke's account of identity depends upon the success of his theory of proper names. Without a clear exposition of how it is that rigid designators can have their reference fixed by the use of a nonrigid designator, the metaphysical necessity of identity statements containing only rigid designators can be called into question. According to Kripke, the reference of rigid designators is sometimes fixed by the use of a nonrigid definite description. The reference of the name "Tully," as was previously mentioned, may be fixed by the use of the description "the Roman orator who denounced Cataline," as presumably occurs when someone asserts the identity statement "Tully is the Roman orator who denounced Cataline" in response to a query about the reference of the name "Tully." Kripke needs to clarify the exact manner in which a nonrigid designator fixes, or sets, the reference of a rigid designator, for this will help to explain why any true identity statement containing only rigid designators should be regarded as true in all possible worlds. Definite descriptions are often used to fix the reference of a proper name, and such a use of a definite description is successful only when what is singled out is a particular individual to which the speaker wishes to refer and to designate with a name. According to Kripke, once the reference is set, the proper name, as a rigid designator, continues in all circumstances and under all counterfactual conditions to refer to that particular individual, even if, because of changes in the world or changes in our knowledge of the world, the nonrigid description no longer functions to pick out that individual. Kripke's theory of proper names must account for how it is possible for such a determinate reference to be set. Evidently, the intention to refer to a particular individual in the original act of asserting the appropriate identity statement serves somehow to fix irrevocably the reference of the proper name. However, since, for Kripke, a statement's metaphysical necessity requires its truth in

all possible worlds, without an adequate understanding of how such identity statements do in fact fix reference it is doubtful one would be compelled by Kripke's analysis of identity to accept the metaphysical necessity of identity statements containing only rigid designators. (That there is this dependency seems odd and suggests that there is something not quite right about Kripke's conception of metaphysical necessity; after all, why should the metaphysical necessity of a relation or of a statement of that relation depend upon the manner in which singular referring terms acquire their reference?) Furthermore, if one is not compelled by Kripke's analysis to accept the metaphysical necessity of these identity statements, then one should also not be compelled to accept the component *de re* necessity Kripke contends is present in any identity statement. After all, if simple identity statements containing only proper names or other rigid designators are not metaphysically necessary, then it is hard to see why the identity relation should be regarded as a necessary metaphysical relation holding between any individual and itself. In addition, I should also add that in order for Kripke's causal theory to be a true alternative to a description theory, such as that offered by Searle, Kripke must indicate how the reference of a proper name may be fixed by a description without either the proper name becoming synonymous with the description or the use of the proper name involving the presupposition that the thing referred to satisfies the description.

Since my previous analysis of the appeal to Russell's theory of descriptions in the attempt to solve the problem of the contingent nature of identity statements containing definite descriptions relied on taking seriously the sort of distinction Donnellan made by his referential-attributive distinction, I need to discuss Kripke's rejection of that distinction. In his "Speaker's Reference and Semantic Reference,"

Kripke argues for a unitary approach to interpreting the content of definite descriptions. Kripke rejects the idea of there being a semantic ambiguity in definite descriptions, and Donnellan's distinction requires, Kripke contends, the presence of such an ambiguity in statements containing a definite description. The sort of phenomena that Donnellan cites as evidence of the need to mark the distinction between the referential and the attributive uses can be accounted for, according to Kripke, in a manner that does not amount to ascribing an ambiguity to definite descriptions. Kripke notes that what words mean when uttered in a certain context may differ from what a speaker means in uttering those words. What the words mean is fixed by linguistic conventions, intentions on the part of the speaker, and certain other contextual elements, while what the speaker means is fixed by what his words mean in the context of his utterance and certain special intentions together with more general conventions governing conversational exchanges. There is thus a distinction to be made between what Kripke calls the "semantic reference" of a term and what Kripke calls the "speaker's reference" of a term. With regard to definite descriptions, the former is the particular person or thing the description refers to in virtue of that person or thing possessing the features detailed in the description. In contrast, the speaker's reference of a definite description is the particular person or thing to which the speaker intends to refer in using the description. In cases that Donnellan describes as attributive uses of a definite description, the speaker intends to say something about whatever or whoever is the semantic reference of the description. In such cases, the speaker's reference is the same as the semantic reference. Speakers may, however, have in mind a particular person or thing about which they intend to make their assertion and because of their beliefs about the intended referent use a definite description to single out and refer to that individual. In these latter cases,

which Donnellan describes as referential uses of a definite description, the speaker's reference may, but not necessarily will, differ from the semantic reference, even though the speaker believes that they are one and the same. Kripke calls the first kind of case, wherein the speaker's reference is just the semantic reference the "simple case" of reference. The latter kind of case, wherein the speaker's reference may not be the semantic reference of his words but the speaker believes that his "general linguistic intentions" (those intentions that fix the meaning of his words in the context of their use) also determine his "specific intention" (his intention to refer, by the use of his words, to a certain individual), Kripke calls the "complex case" of reference. Thus, in a simple case of reference in which a speaker says, "The tallest mountain in the world is in Tibet," the sentence the speaker utters and what the speaker says in uttering the sentence will both be true when and only when there is a unique mountain that is taller than all others and is found in Tibet. In contrast, in a complex case of reference involving a speaker saying of a particular party goer drinking from a champagne glass, "The person over there drinking champagne is happy tonight," the sentence the speaker utters is false if and only if there is no unique person over there who is drinking champagne and is happy tonight, but what the speaker says about the intended referent of the definite description is true if that particular person (the speaker's referent) is indeed happy tonight, regardless of what is in his champagne glass (or what is in the champagne glass of someone else who just so happens to fit the description). Since, under Kripke's proposed analysis of the two different senses of the definite description associated with the two different uses of descriptions, no difference in literal meaning is ascribed to the definite descriptions involved, Kripke concludes that there is only one analysis of the content of definite descriptions. If

there is no syntactic or semantic ambiguity present in definite descriptions, then, according to Kripke, they can have only one analysis.

Kripke contends that one can also distinguish between simple and complex cases of reference in connection with the use of a proper name, rather than a description, and therefore Donnellan is mistaken in characterizing the referential use of a definite description as the use of the description as a proper name. For instance, someone may misidentify a person seen in the distance as Jones (when in fact it is Smith) and say, "There is Jones," and then go on to say of that person seen in the distance, "And Jones is raking leaves again." In the situation imagined, the speaker's utterance of the first sentence, in which is revealed his misidentification of the person seen in the distance, involves a simple case of reference since the speaker, in using the name "Jones," intends to refer to the name's semantic referent. The speaker's utterance of the second sentence, on the other hand, involves a complex case of reference since the speaker intends to say something about the individual seen in the distance and, because of his mistaken belief that the individual seen is Jones, uses the name "Jones" to refer to that individual. What the speaker means and succeeds in saying in uttering the second sentence, apart from what his words literally mean in uttering that sentence, is true provided that the speaker's referent (the specific person to whom the speaker intends to refer and the speaker believes is named by the name, or believes fits the description) is indeed raking leaves.

Since Donnellan's distinction between the sense of a definite description used referentially and the sense of a definite description used attributively is not required in order to understand what occurs in simple and complex cases of reference, and that distinction does not pertain to cases of mistaken identity involving the use of a proper name instead of a description, Kripke concludes that considerations of simplicity and

completeness count against Donnellan's analysis. Kripke sees Donnellan's appeal to an alleged referential-attributive distinction as an unnecessary extravagance. Kripke maintains that the considerations Donnellan entertains as an argument against Russell are inconclusive and do not in themselves show that Russell was wrong about definite descriptions. The phenomena that Donnellan cites to show that some English sentences presumably have nonRussellian truth conditions would still arise even if the statements containing definite descriptions made by English speakers were considered always to be analyzable in the manner proposed by Russell. Thus, the fact that the phenomena do arise among speakers of English is no proof that Russell's analysis is not correct for English. According to Kripke, no clear misunderstanding of English sentences would result from always interpreting the content of any English sentence containing a definite description in the Russellian fashion. Despite what Donnellan suggests, it is far from obvious that the truth conditions of actual English sentences containing a definite description do indeed differ in the manner implicit in the recognition of the attributive and referential senses of a description. Complex cases of reference, wherein the semantic referent is different from the speaker's referent, always involve the speaker making a statement that is false if interpreted literally but is nevertheless being used to say something that may be true of the intended referent. How this is possible is explained in terms of Kripke's account of how speaker's reference and semantic reference are determined, and this is accomplished without recognizing a special kind of ambiguity associated with the different uses of a definite description. Furthermore, since the use of a proper name to refer to someone who has been misidentified would not normally prompt us to say that proper names can be used ambiguously in the manner Donnellan suggests for

descriptions, Kripke argues that we should not attribute that ambiguity to definite descriptions.

The Kripkean analysis of the hypothetical situation involving the misidentification of Smith as Jones appears, though, to be in no way necessary. Since, according to the story, Smith, not Jones, is seen in the distance and Smith, not Jones, is raking leaves, what the speaker succeeds in saying by uttering each one of the above sentences seems to me to be false. It is only by virtue of a liberal application of a principle of charity that the speaker in such a situation could be interpreted as having said something true. The tendency to suppose that something true has been said in the case imagined stems from thinking of the use of the name "Jones" as succeeding, despite the misidentification, in referring to the person who actually satisfies the definite description "the person seen in the distance." However, if, as Kripke does in fact maintain, a proper name is not equivalent to one or more descriptions, then the name "Jones" does not need, in any sense, to acquire Smith as its referent just because Smith is actually the person seen in the distance. As long as speakers are not employing language merely as a code, the referent of a proper name could always be considered to be what Kripke would identify as its semantic referent. Indeed, I suspect that such a view of proper names may ultimately be required if one accepts the view of proper names as rigid designators. If what speakers refer to in using a particular proper name in normal conversation must be seen as remaining constant when talking about the actual world or any possible world, then a case of mistaken identity would not seem to result in a shift of reference.

In virtue of the distinction between what speakers say literally and what they succeed in telling an audience in uttering certain words, Kripke maintains that it may be the case that what a speaker says literally in uttering a particular statement (i.e.,

what his words mean) is false but what the speaker says in uttering that statement in the context in which it is uttered (i.e., what he means by uttering those words) is true. Thus, as in a case Kripke considers, if a person sees a man and a woman together and says of the man, "Her husband is kind to her," and the two people are not actually married to each other but the man is indeed kind to the woman, then the sentence the speaker uttered is literally false even though what the speaker has succeeded in saying about the man is true. However, just what is it that (1) the speaker has accomplished to say about the man and (2) is true? If truth is a property of statements or assertions, then what satisfies the two conditions must be a sentence or must be expressible by a sentence. The sentence spoken, though, cannot be, or cannot express, what is said truly about the man since that sentence (according to Kripke) is false. Nevertheless, the sufficient and necessary conditions for the speaker having said something true are, by hypothesis, satisfied since the speaker believes that the semantic referent of "her husband" is indeed the person to whom he intended to refer and that individual is kind to the woman. Consequently, it is very clear what sentence (i.e., what words) the speaker uttered, but it is not clear what true statement the speaker made.[31] The most

---

31. Searle, in Chapter 6 of his *Expression and Meaning* (137-161), offers an alternative analysis that handles much better the problem of specifying what exactly is the statement that gets made in what is called a referential use of a description. Searle also rejects Donnellan's referential-attributive distinction and presents a view that is, by his own admission, similar to that of Kripke. According to Searle, the speech act of referring to an individual can be accomplished in most cases by a speaker using any number of definite descriptions to pick out the individual in terms of features it alone is alleged to possess. The unique feature (or aspect of the individual) the speaker mentions when using a description whose satisfaction is not a part of the truth conditions of the statement made is what Searle calls the "secondary aspect." If a speaker knows what he or she is talking about in the sense that there actually is some individual the speaker has in mind and about which he wishes to say something, then, according to Searle, there must be some aspect of the individual that is unique to that individual and can be appealed to in using a description. The aspect that is actually unique to the individual and can be mentioned in the use of a definite description to refer to the individual is what Searle calls the "primary aspect." For Searle, a definite description used to mention a primary aspect of an individual is such that (1) the individual the speaker intends to refer to actually satisfies the description and the speaker is aware of this, (2) the description's satisfaction is a component of the truth conditions of the statement actually asserted, and (3) the content of the description is part of the content of the statement made, but not necessarily a part of the meaning of the sentence uttered. Searle

plausible way to resolve the difficulty in terms of Kripke's speaker's reference-semantic reference distinction would be to view the sentence "Her husband is kind to her" as false if the referent of "her husband" is taken to be its semantic referent and as true if the referent of "her husband" is taken to be the speaker's referent. However, to adopt this view would amount to seeing the sentence as, in effect, ambiguous and to reintroducing, albeit in different terminology, Donnellan's referential-attributive distinction.

More importantly, despite what Kripke suggests to the contrary, there is evidence (in addition to what I have already provided) for the need to recognize the sort of distinction that Donnellan makes. Noting Donnellan type distinctions in the content of definite descriptions can reveal the implausibility of certain inferences. The faulty nature of the traditional theistic rejoinder to the classical dilemma proposed to show that neither God nor any other being could be omnipotent can be made apparent if consideration is given to the content of a key descriptive phrase. Those who pose the dilemma argue that, since God either can or cannot create a situation in the universe in which He cannot act and either hypothesis entails that God

---

claims that each of Donnellan's alleged examples of a referential use of a description involves a case where the satisfaction of the description has to do with a secondary aspect, whereas each of Donnellan's alleged examples of an attributive use of a description involves a case where the satisfaction of the description has to do with only the primary aspect. Thus, for Searle, the statement that actually gets made in a so-called referential use of a description is always expressible by a sentence containing a description that characterizes the individual referred to in terms of the primary aspect alone. The trouble with Searle's analysis is that it is hard to see how the satisfaction of a definite description describing an individual in terms of the primary aspect could actually enter into the truth conditions of a statement in a case where the sentence uttered does not contain that description. Since the primary aspect may vary across different counterfactual situations and the definite description the speaker would need to employ to refer to the individual in terms of the primary aspect so as to make the same claim about the same individual may therefore vary in different counterfactual situations, the satisfaction of such a description is not what really counts when determining the content of the statement made. What is important is just reference to the individual, which suggests that the definite description functions like a proper name.

is not omnipotent, God must not be omnipotent. The sort of reasoning involved in reaching the first horn of the dilemma can be explicated as follows.

1. Either God can create a situation in the universe in which He cannot act or God cannot create a situation in the universe in which He cannot act.

2. Suppose God can create a situation in the universe in which He cannot act.

3. Thus, there is some possible situation in the universe that God can create and in which He cannot act.

4. So, suppose that the situation in the universe in which God cannot act is a possible situation in the universe that God can create and in which He cannot act.

5. Thus, the situation in the universe in which God cannot act is a possible situation in the universe in which God cannot act.

6. Thus, God cannot act in every possible situation in the universe.

7. Thus, God is not omnipotent.

Perhaps the most common theistic response to the problem ("the paradox of omnipotence") is to admit that God can conceivably create the situation in which He cannot act but argue that God's acting in such a situation is not logically possible and thus God's inability to act in such a situation is no infringement on His omnipotence. A theist who took such a position would thus reject the above line of reasoning as fallacious. (The criticism would focus on the inference from line 6 to line 7.) The theist would argue that the sentence "God acts in the situation in the universe in which God cannot act" is self-contradictory and thus any description of God acting in such a situation is a description of a "logically impossible action" and is therefore really a description of no possible action at all. As a consequence, God can be

thought to have the ability to create the problematic situation, and be unable to act in that situation, without compromising His omnipotence. This refutation, though, relies on treating the description "the situation in the universe in which God cannot act" as an attributive description, for the above-mentioned sentence is self-contradictory only if the content of the description is part of the content of the sentence and fixes its truth conditions. However, the description enters line 4 in the above derivations that give rise to the first horn of the dilemma as a designator for a particular thing the existence of which is guaranteed by the truth of the existential generalization at line 3. The description must therefore be interpreted as a referential description, but, if it is so understood, the contradiction the theistic critic has in mind cannot be derived. In the supposition at line 4, one legitimately could, after all, have instead used the designator "the Ω situation" in stating the particular instance of the generalization, but the sentence "God acts in the Ω situation" is not self-contradictory.

What is of even more concern than the above case, given the main topic under discussion, is the case of problematic inferences involving identity statements. The following example was motivated by a similar example provided by John Paulos.[32]

The temperature outside at 12:17 a.m. is the ambient temperature of 65$^{\text{O}}$F.

The temperature outside at 12:17 a.m. is never recorded by the weather bureau.

Therefore, the ambient temperature of 65$^{\text{O}}$F is never recorded by the weather bureau.

What is wrong with the above argument can be accounted for in terms of the referential-attributive distinction. (The inference does not appear to involve an

---

32. On page 25 of his amusing little book, *I Think, Therefore I Laugh: The Flip Side of Philosophy.*

opaque context or to turn on an ambiguity in the scope of a generalization.) The description "the ambient temperature outside at 12:17 a.m." in the first premise should be interpreted referentially as a device for referring to a particular temperature at a particular time. The same description in the second premise should be interpreted attributively so that the premise is understood to state that the weather bureau never records the ambient temperature outside at 12:17 a.m., whatever that temperature is. The conclusion that the weather bureau never records the ambient temperature outside at 12:17 a.m. (i.e., the ambient temperature of 65°F) does not then follow. The weather bureau, after all, may well not have a vested interest in recording whatever is the ambient temperature at precisely 12:17 a.m. but may well have no qualms about recording the specific ambient temperature of 65°F whenever it occurs at a recording time.

As I previously indicated, the above remarks were intended to constitute a preliminary criticism of Kripke's position. The difficulties of understanding how identity can be a necessary metaphysical relation, how nonrigid designators can fix the reference of rigid designators, and how best to analyze definite descriptions are all very important but are ancillary to other more pressing problems. I will now turn my attention to my main argument against Kripke's proposals as an analysis of identity. In the long run, I do not find Kripke's account of the nature of identity statements very helpful in unraveling the content of such statements and resolving the problem of identity. This is the case because, as I shall argue, Kripke's analysis cannot ultimately be employed in a full explication of the epistemological and metaphysical necessity (and thus the content of) identity claims since his analysis presupposes a prior understanding of the nature of identity. I will defend the latter claim by citing three reasons why Kripke can be charged with providing an analysis of identity that is

not independent of a prior notion of identity. First of all, Kripke cannot rely upon his account of essential predication in order to explain the metaphysical necessity of true identity statements since the ability to recognize an essential property requires one already to have a notion of what it means for terms to be co-referential. Secondly, in order to extend Kripke's analysis to account for the logical features of identity statements, a notion of a purely referential context free of a notion of identity is required, but such a notion is not provided by Kripke's analysis. Thirdly, the notion of rigid designation cannot be exploited as a means of accounting for the epistemological and metaphysical necessity of identity statements because the notion of a rigid designator carries with it the notion of identity.

If an object's self-identity is considered to be one of its necessary properties and the necessary properties are the essential properties, then it may be thought that an appeal to Kripke's essentialism could be made in an explication of an identity statement's metaphysical necessity. According to Kripke, a thing's essential properties are just those properties whose attribution is necessary in order for a certain name to refer in any possible world (counterfactual situation) to the same thing that it does in the actual world. Kripke maintains that in some counterfactual descriptions of the world a proper name may not succeed in referring to anything because it cannot refer to what it refers to in the actual world due to certain features of the counterfactual description. A counterfactual situation prevents a name from having the same referring role that it has in the actual world whenever the counterfactual description specifically precludes the referent of the name in that possible world from possessing some property its referent in the actual world necessarily possesses (i.e., possesses in every possible world).

For example, consider the sentence "If Nixon had been born Chinese in the sixteenth century, he would not have become a corrupt politician." The referring role played by the name "Nixon" in this sentence is very much in doubt since it does not seem possible for the name to refer to the same individual man both in the actual world and in the counterfactual situation. Unless, perhaps, there is a clear notion of a soul that constitutes a person's identity regardless of the person's physical attributes and his or her time and place of birth, nobody who was born (in the actual world or any possible world) of Chinese parents in sixteenth century China could be identified as the same man who has in actuality been referred to in using the name "Nixon." Thus, the attribute of not having been born Chinese in the sixteen century was evidently an essential attribute of Nixon.[33] However, consider the sentence "Had Nixon not resigned he would have been removed from office." Presumably, there is no problem posed by the use of the name "Nixon" in this sentence since it is possible to imagine one and the same man as having resigned in the actual world and not having resigned in some other conceivable world. The attribute of having resigned thus appears not to have been an essential attribute of Nixon.

The trouble with appealing to the notion of essential attribution in accounting for the metaphysical necessity of identity statements should be obvious. In order to identify a property as a thing's essential property, one needs to know whether a rigid designator referring to that thing could be used to refer to what is the same thing but lacks the property in some counterfactual situation. Thus, a notion of co-

---

33. Admittedly though, Kripke presumably does not consider one's nationality or time of birth to be a person's essential properties. Nevertheless, he does consider a thing's particular origin to be an essential feature of the thing, and a person's origin seems to me to be more commonly understood in terms of an origin within a particular genealogy. (Kripke has given consideration to a person's origin in terms of a particular sperm and egg, but most people, including most reproductive physiologists, are completely unfamiliar with the particular sperm and egg that unite and give rise to a particular person.) I take it that anyone who was born ethnically Chinese in the sixteenth century could not have had the lineage that Nixon in fact had.

referentiality, which in turns requires a prior notion of identity, is antecedently required in order to talk about the essential attributes of the referent of a term in any statement of identity. The notion of essential predication therefore cannot be usefully employed in a Kripkean analysis of the nature of identity statements.

As has already been mentioned, any totally successful analysis of identity will account for the validity of the basic patterns of inference that get expressed in a formal language as the inference rules of identity elimination and identity introduction. If identity is a relation that only holds between every object and itself (and necessarily holds when it in fact holds), then it would seem plausible to say that, if **a** is identical to **b**, then anything true of **a** is also true of **b**, since **a** and **b** are the same thing. (Unfortunately, though, this latter does not seem to be an altogether helpful way of putting the matter since saying that **a** and **b** are identical appears exactly equivalent to saying that **a** and **b** are the same thing.) If this is a plausible, and significant, thing to say, then the validity of the basic patterns of inference associated with identity statements is evidently secured. It would appear then that Kripke's objectual analysis of identity might reasonably be extended to give an account of the logical nature of identity statements. However, it is not always correct to say that, when **a** is the same as **b**, whatever is true of **a** is also true of **b**. As was discussed earlier, when what we say of **a** involves an intensional context, we may be saying something about **a** that we could not properly say of **b**. The inferences judged to be legitimate according to the analysis of identity must be limited to those that involve only purely referential (i.e., referentially transparent) contexts. Those contexts that are purely referential are extensional contexts. If an analysis of identity can only provide an account of a purely referential context that is based fundamentally on a previous explication of the validity and invalidity of inferences

involving the substitution of co-referring terms, then the limitations that must be placed on a principle of intersubstitutivity cannot be accounted for and justified on the analysis (for in such a case the analysis would be circular). If the analysis of identity cannot make plausible a principle of intersubstitutivity in an appropriate manner, the analysis will not give an account of the validity of the basic inferences associated with identity statements. A purely referential context, as has been mentioned before, is traditionally recognized as one in which the substitution of co-referring terms always preserves truth-value. A principle of intersubstitutivity, of course, governs the substitution of co-referring terms and is not to be understood apart from a notion of identity. (After all, the notion of co-referring terms carries with it the notion of identity.) It is thus difficult to give an account of a purely referential context without making essential reference to a principle of intersubstitutivity and a prior notion of identity. Anyone who considers it feasible to extend Kripke's analysis to give an account of the basic inferences associated with identity statements must believe that his analysis can provide, or can be provided with, an account of a purely referential context that is free of a notion of identity and is not based fundamentally on a prior understanding of the validity and invalidity of inferences involving the substitution of co-referring terms. As a consequence, those who take for granted the former consideration presuppose the latter belief; however, since Kripke does not provide such an account of a purely referential context, those who wish to extend Kripke's analysis must evidently rely upon the traditional account of a purely referential context. Therefore, those who claim that a Kripkean analysis of identity can account for the validity of the basic patterns of inference associated with identity statements are assuming as a given what a successful analysis of identity is designed ultimately to provide.

As I have explained above, Kripke relies upon his account of rigid and nonrigid designation in order to make credible his position concerning the epistemological and metaphysical necessity of identity statements. However, the notion of a rigid designator is the notion of a term that refers to the *same* thing in every possible world in which its referent (in the actual world) exists. The notion of a rigid designator carries with it, as a part of its meaning, a notion of identity. Thus, if one does not understand the idea of **a** and **b** being the same thing, one cannot reasonably expect the idea of rigid designation to be, in the final analysis, very enlightening. If such is the case, then Kripke cannot legitimately account for the necessity of identity statements by appealing to the distinction between rigid and nonrigid designation. Therefore, when Kripke accounts for the epistemological and metaphysical necessity of identity statements, as part of his analysis of the nature of identity claims, he actually presupposes a previous understanding of the very notion he attempts to elucidate.

In the final analysis, Kripke does not succeed in elucidating the concept of identity, and thus the logical content of identity statements. Instead, Kripke provides us with only an explanation of how it is possible for identity statements to be in some sense contingent (i.e., not necessary) even though the relation of identity holds necessarily whenever it holds. Kripke must be understood as taking identity to be a primitive relationship that necessarily holds between every object and itself in spite of the fact that we do not always employ a trivial and logically necessary statement to assert that relationship.

In the next chapter, I will focus on an attempt to solve the problem of identity by, among other things, developing a metalinguistic account of identity. Such

accounts of identity are appealing in certain respects, but they encounter their own sorts of difficulties that do not arise in connection with objectual analyses of identity.

Chapter 5: A Metalinguistic Solution

5.0   Introduction

Both Frege's and Kripke's approaches to clarifying the nature of identity statements are not completely successful at resolving the problem of identity because they do not succeed in overcoming the difficulties that typically attend objectual analyses of identity. Furthermore, both accounts require for their completion a prior notion of identity that it should be the purpose of their analyses to provide. Of the two strategies previously mentioned at the beginning of the last chapter, I will now give some consideration to the second strategy that has been pursued in order to resolve the various difficulties associated with the objectual and the metalinguistic analyses of identity. In the present chapter, I shall examine a couple of metalinguistic accounts of identity statements that employ notions from the theory of speech acts. I will begin in the first section below with a discussion of an account of identity statements offered by Michael Lockwood. I will then critique that account in the subsequent section, and then in the final section I will examine and critique a possible modification of Lockwood's account. In the process of evaluating these views, I will uncover additional problems that I believe inevitably beset any metalinguistic account of identity.

5.1   Lockwood's account of identity

The approach pursued by Michael Lockwood, a modified version of which was subsequently endorsed by Thomas Morris, involves accounting for the differences in informational value of different identity statements by redefining the

notion of a purely referential context and its associated principle of substitution.[1] For Lockwood, the problem of there being contingent statements of identity is easily solved. All substantival expressions (definite descriptions and proper names) do not, in general, have the same reference in every possible counterfactual situation. (Lockwood evidently does not subscribe to the idea of rigid designation.) Thus, it is possible for two substantival expressions contingently to denote the same thing and for the identity statement containing those two expressions to be contingently true. The real problem, according to Lockwood, is generated when we allow co-referring terms to be freely substituted for one another in all contexts that are traditionally recognized as being purely referential. According to the standard account of a purely referential context, the singular terms occurring in such a context serve only the purpose of singling out a unique individual and enabling, in a declarative utterance, a speaker to say something about that individual. Consequently, co-referring terms are considered to function in an equivalent fashion in a purely referential context to enable a speaker to refer to the particular individual about which the speaker wishes to talk. Because singular terms occurring in a purely referential context are being used only as devices to refer and co-referring terms serve the same purpose in such a context, it is customarily thought that any two co-referential terms can be substituted for one another in purely referential contexts. Such substitutions, when made in the context of declarative sentences, should yield declarative sentences that are standardly used to make the same assertion. Since "Gaurisankar" and "Everest" refer (supposedly) to the same mountain, it ought to be legitimate to substitute the name "Everest" for the name "Gaurisankar" in the sentence "Gaurisankar is Everest."

---

1. See Lockwood's essay "Identity and Reference," reprinted on pages 199-211 in *Identity and Individuation*. For the account provided by Morris, see his "functional account" of identity in Chapter 4 of *Understanding Identity Statements*.

However, the result of such a substitution, the sentence "Everest is Everest," expresses a trivial truth and lacks the force and informational value of the statement ordinarily expressed by the original sentence. The temptation at this point, says Lockwood, is not to challenge the notion of a purely referential context and the principle of substitution but instead to reject the idea that proper names and descriptions used referentially serve only the purpose of referring and play no attributive role. The problem, after all, does not arise in the case of descriptions used attributively since the substitution principle never warrants their being substituted one for the other.[2] Nevertheless, Lockwood considers it more reasonable to maintain that proper names and some definite descriptions do serve merely to refer and to place the blame for the problem on the standard account of a purely referential context and the substitution principle, both of which he attempts to modify in order to solve the problem. Lockwood does not find fault with the very idea of a purely referential context but rejects the more traditional interpretation of how the preservation of truth-value under substitutions of co-referring terms delimits or defines exactly what are the contexts that are purely referential.

Lockwood considers the difficulties to be resolved once we understand just how it is that a speaker's utterances of declarative sentences used to make not only assertions of identity but also other assertions about one or more individuals can be

---

2. Indeed, the rendering of the distinction between the referential and the attributive uses of a definite description provides a simple solution to Lockwood's problem, at least for many such substitutions involving a definite description. If the description "the tallest mountain seen from Tibet" is used attributively in an utterance of the sentence "The tallest mountain seen from Tibet is Everest," then clearly the substitution of the name "Everest" for the description will yield a sentence (namely, "Everest is Everest") that is not synonymous to the original. Thus, it is easy to see in this case why the sentence that results from substitution does not have the same informational content as the original sentence. Lockwood's problem would therefore be totally resolved (but most, if not all, substitutions of co-referring terms would be viewed as illegitimate) if it were reasonable to hold that all uses of names amount to uses of disguised definite descriptions understood attributively and no descriptions are ever used purely referentially.

informative or not with respect to a given audience. When a speaker intends to inform an audience by uttering a declarative sentence in order to make an assertion in a standard conversational setting, the speaker assumes that the audience understands the literal meaning of his words and any other meaning attached to those words due to the context of their utterance. Also associated with the referential use of any substantival expression (i.e., the use of an expression only to refer to something and not to attribute to something) in a declarative utterance is what Lockwood calls an "assumption of knowledge" and an "assumption of ignorance" regarding the extralinguistic information the audience has at its disposal. The speaker assumes that his audience is familiar enough with the object he is referring to and the features of that object so that the audience can identify exactly what the speaker is referring to in the use of the substantival expression. The speaker also assumes that his audience is either unaware of, or can be usefully reminded of, whatever, as the primary substance of his assertion, he is ascribing to (predicating of) the individual or individuals to which he is referring. The former assumptions comprise the assumption of knowledge; the latter assumptions, the assumption of ignorance.

In the referential (nonattributive) use of a definite description in the utterance of a sentence, the particular assumption of knowledge involves the speaker's assumption that the audience knows enough about the attributes of a thing and its relations to other things to be able to recognize it as being singled out by the description. The particular assumption of ignorance in this case involves the speaker's assumption that the attribute or relation ascribed to the object or objects referred to in the statement is not a part of what is known by the audience that would enable its members to recognize the particular individual the speaker intends to refer to in using the description.

In regard to the referential use of a name (i.e., the use of a name as a logically proper name and not as a disguised description) in the utterance of a sentence, the matter of what constitutes specifically the assumption of knowledge and the assumption of ignorance is a bit more complicated. In order for an audience to be informed by the assertion made by a speaker in uttering a sentence containing a name, the name used must mean something to the members of the audience. Lockwood's sense of a name "meaning something" to an audience is the sense in which the name gives access to a body of knowledge, possessed by each member of the audience, concerning the referent of the name.[3] The notion of a name giving access to a body of knowledge concerning its referent is, I believe, the idea of the name's occurrence in each sentence of a complex of sentences that manifests, at least in part, a person's knowledge regarding the referent of the name. The body of knowledge, about an individual, to which a name gives access is what Lockwood calls the "mental file" associated with the name. The particular assumption of knowledge involved in the referential use of a name in a declarative sentence is that the person to whom the utterance of the sentence is directed has a mental file to which the name can give access and the information in this file about the individual referred to is compatible with the use of the name in that context. The particular assumption of ignorance involved in this case is that the information contained in the mental file to which the name will give access, on the part of the person to whom the utterance of the sentence is directed, does not contain the information that the assertion on the whole is intended to impart.[4]

3. If I understand Lockwood correctly, what he means by a name "meaning something" to someone can also be stated more informally as what a person would mean in asking another person a question like "Does the name 'Michael Lockwood' mean anything to you?"

4. The information the assertion of the sentence is on the whole intended to convey is, if I follow Lockwood correctly, the information about the subject of the sentence that is provided by the content of the grammatical predicate of the sentence, the subject of the sentence in this case being not

In the specific case of a speaker's utterance of a statement of identity that is intended to inform an audience, the speaker assumes, when the identity statement contains only names and those names do not function as disguised descriptions, that each member of the audience has more than one mental file (on the individual referred to) to which the names will give access. The speaker also assumes in uttering an informative identity statement containing only names used purely referentially that, for each name occurring in the identity statement, the mental file to which the name gives access does not contain the information that the individual referred to in the use of the name is also referred to in the use of the other name in the identity statement. When a speaker, for instance, utters the sentence "Gaurisankar is Everest" with the intention of informing an audience, the speaker assumes that the names "Gaurisankar" and "Everest" each give access to a mental file on the referent of the name. Furthermore, the speaker assumes that the information within the mental file to which the name "Gaurisankar" gives access does not include the idea that the individual on which that mental file exists is also the individual on which exists the mental file to which the name "Everest" gives access. Similarly, the speaker assumes that the information within the mental file to which the name "Everest" gives access does not include the idea that the individual on which that mental file exists is also the individual on which exists the mental file to which the name "Gaurisankar" gives access. Thus, the assumptions of knowledge and of ignorance involved in a speaker's utterance of an intentionally informative identity statement in which the identity predicate is flanked only by proper names are that the names in the identity statement give separate access to mental files and that these

_____

the grammatical subject of the sentence but rather the referent of the substantival expression occurring in the grammatical subject position in the sentence.

mental files are, for the audience, distinct mental files. The speaker's purpose in uttering an informative identity statement, Lockwood declares, is to encourage the audience to merge these distinct mental files into one.

Since different uses of names and descriptions in differing contexts involve different assumptions of knowledge and of ignorance and these assumptions on the part of the speaker are key to understanding the informational value that an utterance has for an audience, it is clear for Lockwood why the separate roles of the substantives occurring in an identity statement are not fully recognized when all that is known is their intended reference. Therefore, we should not expect, according to Lockwood, the substitution of co-referring terms one for the other in a sentence always to yield alternative formulations of the same statement. For this reason, Lockwood proposes the adoption of an appropriately modified version of the principle of substitutivity. For Lockwood, the purely referential use of terms must comply with a condition placed upon the interchangeability of co-referring terms that is somewhat stricter than mere sameness of reference. According to the new version of the principle, the substitution of one term for another co-referring term in a sentence is legitimate only when the substitution preserves the informational content of the sentence, and the substitution of co-referring terms in a sentence will yield another sentence that expresses the same statement as expressed by the original provided that two conditions are met.

The first condition that must be met is that the assumptions of knowledge and of ignorance associated with the utterance of a sentence into which such a substitution has been made must not conflict. For Lockwood, the assumption of knowledge conflicts with the assumption of ignorance whenever the identification of the particular referred to in the speaker's use of a term in a sentence requires the audience

already to possess the information the utterance of the sentence is overall intended to convey. The sentence "Everest is Everest" cannot be standardly employed as a means of informing an audience since its assertion would typically involve a conflict between the assumption of knowledge and the assumption of ignorance. In uttering the sentence, the speaker would ordinarily assume (as part of the assumption of knowledge) that each member of the audience has a mental file to which the name "Everest" gives access and would also, paradoxically, assume (as part of the assumption of ignorance) that the mental file to which the name "Everest" gives access does not contain the information that the particular on which there is this file is the particular on which there is a file to which the name "Everest" gives access. However, the information assumed to be lacking is essentially the information that the name "Everest" gives access to a certain mental file, which is something that needs to be recognized by the members of the audience in order for them to identify the particular referred to in the speaker's utterance of the name "Everest." The fact that the combined assumptions normally associated with an assertion of "Everest is Everest" do not make sense indicates for Lockwood that its assertion is not generally equivalent in content to the assertion of "Gaurisankar is Everest." Hence, the substitution of "Everest" for "Gaurisanker" in the sentence "Gaurisanker is Everest" may not preserve informational content and thus may represent an illegitimate substitution.

The second condition that must be met if co-referring terms are to be used interchangeably in a legitimate manner is that the speaker who uses the terms must be able to assume that his listeners or readers do not have distinct mental files to which each term gives its own separate access. The sentence "Gaurisankar is the tallest mountain seen from Nepal" and the sentence "Everest is the tallest mountain seen

from Nepal" cannot be used to make the same assertion whenever the persons to whom the utterance of either sentence would be directed possessed a mental file to which the name "Gaurisankar" gave access that was distinct from the mental file to which the name "Everest" gave access. Only when it is reasonable for the speaker to assume that the members of his audience have only one mental file to which co-referring terms each give access can those terms be used interchangeably to pick out the same mental file and thus to say what amounts to the same thing about one and the same individual.

With a revision of the substitution principle comes a change in the understanding of a purely referential context. For Lockwood, the purely referential use of a term is the use of a term under conditions that generally involve more than the mere fact that the speaker could possibly have instead used an alternative co-referring term in an utterance having the same truth-value. A purely referential context, for Lockwood, is a context in which a term is being used purely referentially. A term is being used purely referentially, for Lockwood, if and only if the use of the term serves only the purpose of enabling the speaker to indicate to his audience a particular mental file on the referent of the term and, when the speaker can assume that his audience has distinct mental files to which the term and other terms having the same reference give separate access, the substitution of the term for one of its co-referring terms in the context of a sentence does not produce a sentence with the same informational content as the original. Thus, according to Lockwood, we would normally expect substitutions of co-referring terms in a purely referential context not to preserve informational content, and possibly also truth-value, under certain circumstances and to preserve informational content, and hence also truth-value, under other circumstances. Substitutions that preserve informational content always

preserve truth-value, but substitutions that preserve truth-value do not always preserve informational content. For Lockwood, then, a purely referential context does not, oddly enough, guarantee the preservation of informational content when co-referring terms are substituted one for the other. Consequently, for Lockwood, more contexts are purely referential than are traditionally recognized.

So, in the final analysis, how exactly, and to what extent, has Lockwood solved the problem of identity? The nonnecessity requirement is easily satisfied: an identity statement of the form $\ulcorner \alpha = \beta \urcorner$ is not necessary since it is logically contingent, and, since it expresses a fact about language use, is presumably a statement that expresses a contingent matter of fact. The contingent nature of statements of identity of the form $\ulcorner \alpha = \beta \urcorner$ should also bestow upon them their nontriviality so that the nontriviality requirement is also apparently satisfied. Lockwood considers an identity statement of the form $\ulcorner \alpha = \beta \urcorner$ to have the same content as an identity statement of the form $\ulcorner \alpha = \alpha \urcorner$ when the assumptions of knowledge and of ignorance associated with the use of the terms α and β conflict. However, when these assumptions do not conflict in the case of statements of the form $\ulcorner \alpha = \beta \urcorner$, identity statements with a content different from the content of identity statements of the form $\ulcorner \alpha = \alpha \urcorner$ result. The difference in content requirement is thus apparently satisfied.

5.2   Objections to Lockwood's account

Nevertheless, there are several puzzling aspects to Lockwood's account of identity, and it is susceptible to a host of possible objections. It is questionable whether, in the final analysis, Lockwood actually provides an account of the content of identity statements, as opposed to merely proposing an explanation of how it is possible for there to be failures of the traditional principle of substitutivity in contexts

that are nevertheless purely referential. The essentials of Lockwood's basic position on identity can be succinctly summarized as follows. The purpose served by assertions of identity is to get an audience to regard as a single body of knowledge what was previously regarded as separate bodies of knowledge. The statement "$a = a$" is always uninformative since its assertion always fails to satisfy what is required for statements in general to be informative. The statement "$a = b$" can sometimes be used by a speaker to inform an audience, and this occurs only when (1) the audience has some idea of what "a" and "b" refer to but does not know that "a" and "b" are co-referring and (2) a recognition of what "a" refers to does not depend on a recognition of what "b" refers to and vice versa. (When this latter condition obtains for a particular assertion of identity, there is no conflict between the assumptions of knowledge and of ignorance involved.) Furthermore, a speaker can legitimately use co-referring terms "a" and "b" interchangeably in making a particular assertion provided that his audience is aware that the terms are co-referring and has some knowledge of the referents of the terms but does not already acknowledge the content of the assertion and does not need to acknowledge that content prior to recognizing what the terms are being used to refer to in the speaker's utterance. Thus, Lockwood attempts to provide an account of identity that centers upon solving the difference in content requirement. The assertion of "$a = a$" is never informative, while the assertion of "$a = b$" can be informative when certain conditions are met. The conditions that must be met are conditions under which it is not legitimate to substitute "a" for "b" (or vice versa) in an identity statement or any other statement. A specification of a workable principle of substitutivity, or of a purely referential context (or minimally of a context that is not purely referential), is therefore crucial to Lockwood's account of identity.

The main trouble with Lockwood's analysis is that the manner in which he resolves the differences in content among identity statements ultimately does not shed very much light on the issue of what exactly is being claimed, if anything, when a speaker makes a true assertion of identity. What Lockwood primarily provides us with is, in effect, a simple account of when the substitution of one co-referring term for another in a statement should result in a statement with the same informational content. Lockwood appeals to some rather commonsense considerations in developing a revised account of a purely referential context, but, unfortunately, too much of what he says is left at the level of an analogy and not explained in more straightforward language. What precisely is a "mental file"? What does it mean for a term to give access to a mental file? What does it mean for a file to contain information on an individual? What does it mean to merge mental files into one? These questions never get fully answered, and yet they are important since they seem to concern matters of knowledge, reference, and linguistic aboutness and meaning. (Presumably, Lockwood could answer these questions by appealing to notions of such things as mental contents, linguistic predispositions, and capacities to recognize concepts, but, unfortunately, Lockwood never does so.) Moreover, Lockwood advances a modified principle of substitutivity according to which substitutions of co-referring terms allegedly preserve sameness of informational content, but Lockwood offers no account of informational content and thus no well-founded support for his proposed revisions to the principle of substitutivity. The modifications he makes to that principle are evidently motivated by a desire to circumvent certain classical problems of substitutivity. However, without an analysis of informational content, his emendations seem ad hoc, the only reason for adopting them being their

usefulness in avoiding some of the difficulties in interpreting some of the apparent failures of the principle of substitutivity.

In addition, it is clear that substitutions made in accordance with Lockwood's revised principle of substitutivity (i.e., substitutions in situations where the two conditions required, according to the principle, for the substitution to be legitimate are met) do not always succeed in preserving truth-value, and it is arguably the case that some such substitutions do not succeed in preserving informational content even if they do succeed in not altering truth-value. That the former is the case can easily be seen by considering a simple example. Suppose, borrowing from an example provided in Chapter 3, Joe believes that Mark Twain wrote *The Adventures of Tom Sawyer and Huckleberry Finn*. However, suppose it is not true to say, because of Joe's lack of knowledge about Mark Twain's real name, that Joe believes Samuel Clemens wrote *The Adventures of Tom Sawyer and Huckleberry Finn*. Suppose now that a speaker, who is someone distinct from Joe, wishes to inform an audience about Joe's belief concerning the author of *The Adventures of Tom Sawyer and Huckleberry Finn*, and suppose each member of the audience knows who Mark Twain was and knows that Mark Twain was Samuel Clemens and is ignorant of Joe's beliefs about the matter. According to Lockwood's principle of substitutivity, the speaker could utter either one of the following two sentences as a means of informing the audience about Joe's beliefs.

Joe believes that Mark Twain wrote *The Adventures of Tom Sawyer and Huckleberry Finn*.

Joe believes that Samuel Clemens wrote *The Adventures of Tom Sawyer and Huckleberry Finn*.

The latter sentence results from the former by substituting the term "Samuel Clemens" for the term "Mark Twain." The substitution that generates the latter sentence is in conformity with Lockwood's revised principle since in the case imagined the utterance of the latter sentence does not involve a conflict between the assumptions of knowledge and of ignorance (since the name "Samuel Clemens" means something to the audience and the audience is ignorant of Joe's beliefs) and the audience does not have distinct mental files to which the names "Mark Twain" and "Samuel Clemens" each give separate access (since the audience knows that "Mark Twain" and "Samuel Clemens" are co-referring terms). According to the suppositions that have been made, however, the former sentence is true while the latter sentence is false. Truth-value, and thus informational content, has therefore not been preserved under substitution in this case.

Lockwood might respond to this situation by pointing out that in the case imagined only the former sentence could be used to report accurately the content of Joe's belief, so the latter sentence could not be used to inform the audience about Joe's belief, the revelation of the content of that belief supposedly being the speaker's purpose for uttering either sentence. The problem, he might claim, is not with the revised substitution principle but with the supposed facts of the matter. However, this would mean that further emendation of the substitution principle is necessary. In order to be legitimate, not only would a substitution of one co-referring term for another in a sentence need to involve no conflict between the assumptions of knowledge and of ignorance when the utterance of the sentence is directed at an audience who recognizes that the terms are co-referring, but the substitution would also need to preserve the truth or falsity of the original sentence. An appeal to a new,

appropriately emended, substitution principle could then only, at best, account for the conditions under which some substitutions of co-referring terms (i.e., those substitutions that preserve truth-value) produce different sentences the utterance of which is, for a certain sort of audience, equally informative. (Notice that the relevant notion of sameness of informational content is too weak to be a general notion of synonymy since it presupposes an understanding of, rather than accounts for, the conditions under which different sentences will exhibit a sameness of truth-value.) Alternatively, Lockwood might respond to my alleged counterexample by claiming that sentences about propositional attitudes, as well as those involving indirect discourse, should always be given a transparent reading. Since both the former and the latter sentences are true given a transparent reading, such a move would eliminate the difficulty. Nevertheless, given this response to the problem (as is also the case with the former response), it hard to see the value of Lockwood's principle of substitutivity, inasmuch as it is of no help in resolving the most serious problems of interpretation that arise in connection with the standard notion of intersubstitutivity. Rather than aiding our understanding of why the substitution of co-referring terms fails to preserve truth-value in certain troublesome cases involving grammatically complex sentences, an appeal to Lockwood's principle of substitutivity can evidently aid our understanding of why substitutivity fails when substitutions of co-referring terms produce only trivial statements of identity or statements containing terms with which an audience is unfamiliar. To be truly useful, Lockwood's principle needs to be further refined so as not to legitimize substitutions of co-referring terms in connection with the utterances of sentences that involve conflicts between the assumptions of knowledge and of ignorance having to do not only with what the speaker and his audience know and do not know but also with what other people

(such as people whose beliefs are reported on) know and do not know. Only by making such refinements will the application of the principle not be questionable in the case I have imagined and in a myriad of other possible cases displaying various degrees of additional complexity. Unfortunately, the necessary refinements of the principle may get quite complicated, and I see no easy way to state a general principle of substitutivity along the lines of Lockwood's principle that is both plausible and significant.

I have characterized Lockwood's position on identity as a metalinguistic view of identity, even though Lockwood never explicitly declares that identity is a relation that holds between the terms in an identity statement and not a relation between the referents of those terms. In fact, since Lockwood makes it clear that he does not wish to challenge the traditional idea of the referential use of proper names and descriptions as serving only the purpose of singling out their referents, it would appear that he views the referential use of terms in identity statements as enabling speakers to refer to the terms' objects of reference instead of to the terms themselves. Nevertheless, if what he proposes is to be taken as a serious account of identity, then his views on identity should be seen as providing for a metalinguistic account of identity. Since it is doubtful that Lockwood subscribes to the peculiar view that assertions of identity are primarily about mental files rather than the terms that give access to mental files or the individuals on whom there are these files, the items that are claimed to be related in making assertions of identity are for Lockwood either the terms or the individuals. What Lockwood states explicitly about the content of identity statements strongly suggests that he interprets the terms to be the subjects of identity statements. For Lockwood, the identity statement "$\mathbf{a} = \mathbf{b}$" means (or the content of the identity statement "$\mathbf{a} = \mathbf{b}$" is) that the terms "$\mathbf{a}$" and "$\mathbf{b}$" can and should

give access to the same mental file. The identity statement "**a** = **a**" means essentially that the term "**a**" gives access to a mental file, which is something that a speaker who utters such a statement presupposes his audience already recognizes. The terms in a true identity statement therefore stand in a certain relation, the relation of giving access to a single mental file on a single individual. For Lockwood, then, the primary content of identity statements is evidently about their terms, not about the referents of their terms. Speakers who assert identity statements are only indirectly concerned with mental files and the particulars on which there are these files. Lockwood's account of identity is thus a metalinguistic account of identity.

Admittedly, though, Lockwood does seem occasionally to depart from a metalinguistic analysis, which suggests that Lockwood's view of identity is not entirely consistent. For instance, when explaining why an assertion of the sentence "Everest is Everest" involves a conflict between the assumptions of knowledge and of ignorance, Lockwood states that the speaker who utters the sentence would be assuming, among other things, that the mental file to which the name "Everest" gives access does not contain the information that the particular on which there is this file is the particular on which there is a file to which the name "Everest" gives access. However, this assumption is, according to Lockwood's notion of conflict between the assumption of knowledge and the assumption of ignorance, the information the utterance of the sentence "Everest is Everest" is overall intended to convey. Thus, Lockwood is apparently committed to the view that the information provided by (i.e., the content of?) the assertion of "Everest is Everest" is about the particulars to which the terms flanking the identity predicate give access. This view is an unfortunate result of Lockwood's particular attempt at accounting for how the two different kinds of identity claims can differ in informative value and is at odds with Lockwood's

fundamental position regarding the content of assertions of identity, which is, despite this anomaly, a metalinguistic account of identity.

Being a metalinguistic analysis of identity, Lockwood's account of identity is thus susceptible to all the difficulties that customarily attend such a view. It is difficult, for instance, to see how the assertion of an allegedly informative identity statement can really be used to inform an audience about the world, since all it conveys is the idea that certain terms are to be regarded in a certain manner (i.e., as terms that give access to the same mental file). Furthermore, Lockwood's stated position on identity concerns only identity statements that contain terms that are being used purely referentially. Lockwood does not deny that an identity statement may contain terms that are being used attributively. (The problem of identity that concerns Lockwood arises, though, only in the case of an assertion of identity that involves the referential use of terms.) However, it is not clear, given Lockwood's conception of identity, what information is imparted to an audience by the utterance of an identity statement that contains a term that is being used attributively. Given Lockwood's notion of identity, the assertions of identity statements containing terms used attributively are in many cases somehow about both terms and the referents of terms. (The values of variables in the component assertions of uniqueness will be objects, but the values of the variables that appear in the contained ascriptions of identity will be both objects and terms.) Lockwood's metalinguistic account of identity thus seems to necessitate that the variables of quantification involved in identity statements somehow be understood in two different ways in order for the account to be applied more generally to cases involving either the referential or the attributive use of terms in an assertion of identity.

In addition, the attributive use of terms in assertions of identity does not appear in most cases to involve simply encouraging an audience to merge mental files. Indeed, unless there is some legitimate sense in which people can be thought to construct at will new mental files on the basis of previously existing mental files, speakers who use complex singular terms usually cannot assume that members of an audience have mental files to which the terms give access. For most people, it is doubtful, for instance, that complicated mathematical terms containing numerous functor expressions give immediate access to mental files on the numbers to which the terms refer. Furthermore, assertions of identity corresponding to mathematical equations containing such complicated terms clearly do not serve the purpose of getting an audience to merge mental files. Rather, such assertions of identity most likely serve the practical purpose of informing an audience about the results of certain calculations. Similarly, the attributive use of more mundane complex terms in an assertion of identity seems in the typical case to serve the purpose of conveying information about the reference of the terms instead of promoting the merger of mental files. For example, the grocery shopper who tells the clerk in the checkout lane, "The most expensive item on my grocery list is the item I want to buy the least," may be attempting to convey explicitly or implicitly several different bits of information but is certainly not trying to get the store clerk to merge mental files. (Unless the situation is unusual, the store clerk does not even have mental files on the most expensive item on the list and on the item the shopper wants to buy the least.) Thus, even if Lockwood's analysis had succeeded in solving the problem of identity in the case of a referential use of singular terms in an identity statement, his analysis still could not be extended to solve the problem of identity in the case of an attributive use of singular terms in an identity statement.

I should also point out that there are some rather obvious problems of circularity with Lockwood's account of identity. First of all, the purposes served by assertions of identity are, for Lockwood, to be accounted for in terms of encouraging an audience to merge mental files into one, or the same, file, but, if we do not understand sameness as a relation between terms, then we are not in a position to understand sameness as a relation between mental files. Lockwood's account of identity, like Frege and Kripke's, requires a prior understanding of identity in the metalanguage. Secondly, separate mental files can be, or should be, merged because the subjects on which those files exist are the same, and thus Lockwood's metalinguistic analysis of identity relies upon a prior notion of identity as a relation between objects. Presumably, if a speaker's identity claims are not always trivial, there should be a reason why the members of an audience can follow the speaker's recommendation to merge mental files. That general reason must reside in the speaker's belief in the identity of the subjects of those files. However, if the identity of the subjects is considered to be just the mergeability of their associated mental files, then a circular truism ("Mental files can be merged because they can be merged") is the result. Thus, Lockwood needs to hold that mental files can be merged because their subjects are the same and that subjects are not the same merely because their mental files can be merged. Hence, the idea of why mental files are the same, and consequently Lockwood's proposed analysis of identity, depends upon a prior notion of identity, a notion of identity as a relation between objects.

5.3   A modified metalinguistic account

Perhaps, though, an identity theorist could offer the following alternative account of identity that is explicitly metalinguistic and is well within at least the spirit

of Lockwood's analysis of identity.[5]   The singular terms used in any assertion of identity are being used to refer only to the terms themselves, and thus there is no distinction between the referential use of a singular term and the attributive use of a singular term at least when it comes to assertions of identity. Furthermore, when a speaker uses an identity statement in making an assertion, the only thing the speaker says literally is that the terms that flank the identity predicate can be or should be used interchangeably. An ascription of identity has no other inherent content other than that associated with a declaration of the interchangeability of its terms. Whether the interchangeability of terms is considered to be due to a stipulation or to the nature of the world is revealed by certain contextual elements and by the features of the utterance of the identity statement that determine its illocutionary force. How then can statements of identity be used to inform an audience? Although the utterance of an identity statement is rarely informative all by itself, an assertion of identity is often made within a context where claims about the intersubstitutivity and co-referentiality of terms have important ramifications that can be recognized by both the speaker and his audience. It is this background that provides identity statements with an indirect informational content and enables them to have an added significance that they normally do not possess. For instance, in telling his students that two terms can be substituted one for the other, an algebra instructor may in effect be informing his students of something significant due to the fact that, given what he has already told them about the subject, they are then able to recognize how to make a substitution of

---

5. This alternative account seems to come close to Morris's "functional account" of identity presented in his *Understanding Identity Statements*. However, if I understand the latter account correctly, Morris's position is even more radical. Morris seems to think that the utterance of an identity statement is not really an assertion of any kind of fact at all. For Morris, as far as content or meaning is concerned, identity statements should be understood only in terms of the function or purpose they serve in getting an audience to merge bodies of knowledge. Wittgenstein also seems to have endorsed a similar more radical view of identity at certain places in the *Tractatus* (e.g., 4.241, 4.242, 5.5303, 6.2, 6.21, 6.23, 6.2322, and 6.24).

terms that will enable an equation to be reduced algebraically. In a similar vein, because it is generally recognized that certain astronomical conditions must be fulfilled in order for us to be able to use two designations for a heavenly body interchangeably, the statement "The morning is the evening star" can be used to inform an audience indirectly about the satisfaction of those astronomical conditions. Identity statements are always trivial in the sense that what is conveyed to an audience immediately by any utterance of one concerns only the interchangeability of terms, but identity statements can have a significant indirect, or connotative, content due to the context of their utterance. In addition, whereas a single term is always in effect used interchangeably with itself, two distinct terms are not necessarily used interchangeably. Thus, the statement "$\mathbf{a} = \mathbf{a}$" always appears trivial and linguistically necessary, but the statement "$\mathbf{a} = \mathbf{b}$" can appear both nontrivial and nonnecesssary. Where "$\mathbf{a}$" and "$\mathbf{b}$" are distinct singular terms, what is said literally and what is conversationally implied by an utterance of "$\mathbf{a} = \mathbf{b}$" is typically different than what is conveyed literally and what is suggested (if anything) by an utterance of "$\mathbf{a} = \mathbf{a}$".

This alternative analysis represents what is perhaps one of the most satisfactory metalinguistic accounts of identity possible. It avoids the problem of giving a clearly circular characterization of identity that ruins the attempt to define identity in terms of co-referentiality (i.e., sameness of reference). It seems to account nicely for the nonnecessity and nontriviality of identity statements, while at the same time preserving at least the indirect significance of such statements. In spite of this, however appealing this view of identity may be, I believe it suffers from two very major defects and should in the end be rejected. First of all, just as is true with Lockwood's metalinguistic account of identity, there is the problem of interpreting generalizations containing the identity predicate. What flanks the identity predicate

in a sentence may not be singular terms but may in effect be variables, as in the generalizations "Something is identical to itself" and "Things both identical to a thing are identical to each other." (The former sentence is sometimes used, perhaps inappropriately, to assert that a domain of discourse is not empty, and the latter sentence may be used as a way of expressing the notion of the transitivity of identity.) Of course, identity statements containing terms that are being used attributively also amount to generalizations in which variables in effect flank the identity predicate. If identity is taken to be a relation that, properly speaking, only holds between terms, then there is a difficulty in interpreting what statements containing the identity predicate without singular terms, or with singular terms used only attributively, are about. The sorts of generalizations mentioned above appear to be about what amount to the values of their variables of quantification. Indeed, in the case of identity statements the utterance of which involves the attributive use of singular terms, the singular terms themselves drop out when these statements are put into their proper quantificational form. However, these sorts of generalizations cannot be exclusively about the referents of singular terms if identity is a relation holding only between terms. These kinds of generalizations would in some fashion be about language and about things in the world at the same time, a highly implausible situation. So, unless a good case can be made for there being two different kinds of quantifiers, one whose variables range over objects and another whose variables range over singular terms, both metalinguistic accounts discussed above, and generally all metalinguistic accounts of identity, face the difficulty of relying upon an inconsistent notion of quantification.[6]

---

6. A proposal to introduce two such kinds of quantifiers into an analysis of identity statements has been entertained by James B. Freeman in "Quantification, Identity, and Opacity in Relevant Logic" (in *Directions in Relevant Logic*).

The other major defect in the proposed analysis of identity has to do with the relationship between the primary or literal content of an identity statement and what is taken to be part of the secondary or implied content of an identity statement. According to the account of identity proposed above, the literal or denotative meaning of an assertion of identity consists of only an assertion of the intersubstitutivity of terms, while a claim about the co-referentiality of terms is only a part of the suggested or connotative meaning of an assertion of identity. Thus, in asserting "**a** = **b**," a speaker says literally only that the terms "**a**" and "**b**" can be used interchangeably in suitable contexts without altering the truth or falsity of what is asserted. Nonetheless, since we would not normally claim that two singular terms can be used interchangeably unless there was something that they both referred to, given certain considerations concerning the nature of reference and the nature of singular terms, the assertion of the interchangeability of "**a**" and "**b**" seems to suggest, or to conversationally imply, that there is some x such that "**a**" refers to x and "**b**" refers to x. This latter fact in turn seems to suggest, given certain higher level metalinguistic assumptions, that the terms "the referent of '**a**'" and the referent of '**b**'" can also be used interchangeably. If these latter two terms can be used interchangeably, then, according to the metalinguistic notion of identity, the referent of "**a**" is the same as the referent of "**b**," which means that "**a**" and "**b**" are co-referring terms. In this fashion, the co-referentiality of "**a**" and "**b**" may be seen as a part of the connotative meaning of asserting that **a** is equal to **b**. However, the metalinguistic assumptions about singular terms and reference must be about both objects and language in order for what is stated about language to imply something about objects. These assumptions must be about the uniqueness of the reference of singular terms and the "aboutness" of statements or assertions and the

interchangeability of their terms. The notion of unique reference involved will contain notions of identity, at least one of which cannot be a notion of identity as a relation between terms. In order for singular terms to be connected with their referents, at least one component notion of identity must be a notion of identity as a relation between objects. In order for a speaker to imply the co-referentiality of terms by asserting the intersubstitutivity of terms, the speaker must therefore rely upon assumptions some of which involve treating identity as a relation between objects. The identity theorist who advocates the proposed analysis has thus failed to realize that, if "$a = b$" is true, then "a" and "b" can be substituted one for the other because they are co-referring terms. If singular terms are intersubstitutable in appropriate contexts, then they are so intersubstitutable because they are co-referential; they are not co-referential merely because they are intersubstitutable.

Therefore, from the examination of Lockwood's account of identity and the alternative analysis, there emerge two other sorts of problems with which the metalinguistic identity theorist must deal. There is a problem of interpreting the quantification involved in identity statements: the metalinguistic analysis appears to force an inconsistent interpretation of quantification. There is also the problem of the reliance upon a prior notion of identity as a relation between singular terms and objects: a metalinguistic account of identity seems to depend for its completion upon an objectual account of identity. Furthermore, the appeal to speech act theory does not alleviate these problems but instead may actually make them more apparent.

Chapter 6: Toward a Solution to the Problem

6.0   Some conclusions regarding attempts to solve the problem

The difficulties that attend the typical sort of objectual analysis of identity emerge from the consideration of Frege's and Kripke's accounts of identity. Associated quite naturally with the view of identity as a relation between objects is the understanding of identity as a metaphysically necessary relation that always holds between anything and itself. With regard to at least metaphysical necessity, the objectual identity theorist thus comes to reject the idea that there are contingently true (i.e., true but not necessary) statements of identity. All identity statements are at least in one sense necessarily true when true. The identity theorist who advances an objectual account then provides for the sense in which some identity statements are not necessary and not trivial by a special theory of the expressions or terms that flank the identity predicate in an identity statement. One consequence of the special theory of expressions may be that not all statements that appear on the surface to be statements of identity are actually, properly speaking, identity statements. The fact that some statements that appear to be identity statements are informative is then interpreted to be a result of the fact that not all such statements are in some sense both necessary and trivial. Unfortunately, the sense in which some apparent identity statements turn out to be neither necessary nor trivial does not seem to capture in its entirety our ordinary idea of contingently true identity statements. Some statements of identity seem to be contingently true because the identification of the referents of

the expressions flanking the identity predicate as the same object is itself a contingent matter of fact quite apart from the nature of the expressions used to refer to the objects. Another trouble for the objectual identity theorist is that his particular strategy for solving the problem of identity does not enable him to provide a noncircular account of identity that is thereby free of a prior notion of identity. The identity predicate remains in an important sense primitive and unanalyzed on an objectual analysis of identity.

A metalinguistic account of identity, such as the one Lockwood offers, is plagued by another set of difficulties. If identity is a relation among the terms in an identity statement, the terms in an identity statement are being mentioned instead of being used, and what, in asserting a statement of identity, a speaker is claiming is essentially that the terms are being used or should be used interchangeably. As a consequence, if identity is a relation among terms, identity statements are statements primarily about language use and only have an indirect connection to the things in the world about which we wish to speak. The untoward consequences of adopting a metalinguistic approach to identity thus seem more obvious than they are for an objectual approach, for the metalinguistic interpretation of identity statements appears immediately counterintuitive. There is also the problem of making sense of the quantification involved when the identity predicate occurs in generalizations, particularly in generalizations where it is clear that a statement of identity constitutes only a part of the generalization. In addition, the identity theorist who defends a metalinguistic analysis will ultimately have to rely on an objectual understanding of

identity statements in order to account for how it is possible for identity claims to be significant claims that are at least indirectly about things in the world. An objectual understanding of identity statements will also be required if such an identity theorist wishes to characterize a purely referential context, in which substitutions of co-referring terms are always legitimate, so that his specification of the truth conditions for identity statements on the metalinguistic reading is meaningful in that it avoids vicious problems of circularity. Thus, one who offers a metalinguistic account of identity is actually presupposing that all the difficulties that attend an objectual account of identity have been resolved. As a result, there seems to be little reason for pursuing the metalinguistic strategy for developing an account of sentence meaning aimed strictly at identity statements.

In the remaining sections to follow, I want to give some consideration to the idea of identity as a logical relation and to the idea of identity as indiscernibility. Several distinct points will emerge from my brief discussion, and I will conclude by weaving these different lines of thought together in order to suggest a strategy that may in the end be fruitful in solving the problem of identity.

6.1   Identity and truth as logical notions

Identity is often treated as an irreducible logical relation that holds between any object and itself. What sense, though, can be made of the notion of a strictly logical relation that holds true of objects? Since the logical relations are normally thought of as the inferential relations that obtain among sentences, this is certainly not

the exact sense we should attach to the identity predicate. In formal systems of logic, the sign for identity is treated as a binary predicate except that it has a standard interpretation, and thus in this sense has a fixed meaning, and for this reason it is deemed a logical predicate. However, any formal language sentence of the form $\ulcorner \alpha = \beta \urcorner$, where α and β are distinct singular terms, is a contingency, just as it is the case with any nonlogical binary predicate, and the differences in truth-value among different identity statements of this form are due solely to differences in the denotations (assignments or interpretations) of the singular terms.

In formal systems, we do have precise rules governing the relations among identity statements and certain nonidentity statements in proofs. In spite of the differences in the interpretation of the nature of the identity relation provided by the objectual and metalinguistic accounts of identity, the inference rules of identity introduction and identity elimination are standard. Furthermore, there is an important sense in which to understand the meaning of a logical particle is just to understand the inferences that can be made from sentences containing that particle, and I think this is also true in the case of identity (which is understandable given that, as I have indicated, the logical problem of identity will not be solved unless and until there is provided an understanding of identity that warrants its associated inference principles). Thus, to understand the meaning of the identity predicate, one needs to take into consideration its associated inference rules, especially the rule that gets expressed in formal systems as the rule of identity elimination, since the truth conditions for identity claims can be stated (and are probably more accurately stated)

in terms of this rule. However, these rules are specified in terms of syntax only, and identity is not a relation merely among words. It is also not merely a relation among objects, for the construing of it as such a relation has led to the perplexities previously mentioned. Nonetheless, we can specify the truth conditions of an identity statement in terms of both syntax and semantics by saying that a statement of the form $\lceil \alpha = \beta \rceil$ is true whenever we can replace the name α with the name β (and vice versa) in the transparent context of any sentence and preserve the truth-value of the sentence. Obviously, though, such a specification is plausible only if we can provide an appropriate notion of a transparent context that can be understood independently of the permitted substitutions.

There is an interesting parallel between the identity predicate and the truth predicate. Both the predicate "is identical to" and "is true" have logical and semantical properties. Just as an identity statement has certain logical consequences, so does a statement that ascribes truth to another statement. From the statement $\lceil p$ is true$\rceil$, p can be inferred, and, from the statement p, the statement $\lceil p$ is true$\rceil$ can be inferred. Questions analogous to the questions of meaning, reference, and informational value, which are asked about identity statements, can also be asked about statements that ascribe truth. What precisely is a speaker indicating when he asserts that something is true? What is truth a property of (i.e., what kinds of things have the property of being true)? How, given just the semantical features of the truth predicate, is it possible for statements that ascribe truth to be significant and informative? The distinction between matters pertaining to the logical content of a

statement and matters pertaining to the extralogical conditions under which a statement is judged to be true can also be maintained with regard to statements that ascribe truth. Thus, there are the issues concerning the meaning or content of statements that ascribe truth on the one hand and issues concerning the conditions under which an ascription of truth is itself accepted as true (or the content of a statement that ascribes truth, whatever that content happens to be, is accepted or believed) on the other hand. The similarities between the questions and issues surrounding the identity predicate and those surrounding the truth predicate at least suggest that there is some connection between the predicate "is identical to" and the predicate "is true."

6.2   Identity as indiscernibility

Identity is sometimes characterized in terms of indiscernibility, as is allegedly suggested by Leibniz. Some take the principle of the indiscernibility of identicals and the principle of the identity of the indiscernibles to in effect give the logical meaning of identity statements. According to the former principle (the relatively uncontroversial one of the two), things that are identical are indiscernible (have all their properties in common). According to the latter principle (the more controversial one), things that are indiscernible (have all their properties in common) are identical. The combination of both of these principles is sometimes called "Leibniz's Law." However, Leibniz never clearly advances either one of these principles as a logical

principle.[1] Perhaps the clearest statement of the logical nature of identity Leibniz

made occurs in his discussion of his proposed "universal calculus." In that

discussion, Leibniz had the following to say about identity.

Def. 1.   Two terms are the *same* (*eadem*) if one can be substituted for the other
without altering the truth of any statement (*salva veritate*). If we have A and B, and
A enters into some true proposition, and the substitution of B for A wherever it
appears, results in a new proposition which is likewise true, and if this can be done
for every such proposition, then A and B are said to be the *same*; and conversely, if A
and B are the same, they can be substituted for one another as I have said. Terms
which are the same are also called *coincident* (*coincidentia*); A and A are, of course,
said to be the same, but if A and B are the same, they are called *coincident*.[2]

The above definition of the identity relation unfortunately describes identity as a

relation among terms, but the definiens is interesting, though, since it seems to

represent, in essence, the specification of the truth conditions for identity statements I

mentioned above in the previous section. However, since the definition does not

characterize identity, as a relation between objects, in terms of indiscernibility, those

who see the definition as advancing the view of identity as strict indiscernibility

(according to the two above-mentioned indiscernibility principles) must be appealing

to additional matters (including, perhaps, matters Leibniz never entertained).

Some, evidently prompted by a consideration of the two indiscernibility

principles (or something akin to those principles), have suggested that the identity

relation be defined as a relation that holds between object **a** and object **b** when and

---

1. See, for instance, Ishiguro's discussion of these matters in his *Leibniz's Philosophy of Logic and Language* and Simons' entry in *Routledge Encyclopedia of Philosophy* (s.v. "Identity of Indiscernibles).
2. Translated as section XIX by C.I. Lewis, *A Survey of Symbolic Logic*, 373.

only when all properties of **a** are properties of **b** and vice versa. In a formal system, this definition gets expressed in standard second-order logic as the following formal statement.

$$(x)(y)(x = y \equiv (X)(Xx \equiv Xy)$$

This approach does seem to be promising for several reasons. The questions of reference and meaning seem to be answered in a straightforward manner: the terms in an identity statement refer to objects, and those objects stand in the identity relation just in case they have all their properties in common. Since an identity statement amounts to a logically contingent statement, the assertion of such statements can in principle inform us about the world and for that reason can be significant. Also, a complete grasp of the definition does not appear, on the surface anyway, to rely on a prior understanding of identity. However, depending upon how the quantification over properties is interpreted, individual properties may need to be clearly identifiable, which would require at best a notion of the sameness of properties and at worst a prior notion of the sameness of objects. Furthermore, the domain over which the second-order variables range, or the class of the substituends of those variables, must be specified in such a way as to avoid modal and other sorts of paradoxes, without incurring problems of circularity, and thus be successful in specifying the contexts in which substitutions of co-referring terms are plausible.

Perhaps, though, the strategy of defining the identity relation in terms of some sense of indiscernibility can be employed in some other way so as to yield a feasible account of identity. I will now end my discussion with a proposal for such an amended account. Rather than taking identity to be a relation that holds when things have all their properties in common, one could conceivably construe identity as a relation holding between **a** and **b** just in case everything true of **a** is also true of **b** and vice versa. In order to make this precise, however, one will interpret identity statements to be statements rendered in the metalanguage and will construe the metalanguage as containing both expressions used to refer to sentences in the object language and terms occurring in the object language used to refer to objects. Metalinguistic expressions used to refer to object language sentences will be constructible from sentences in the object language together with an apparatus in the metalanguage that functions like quotation. The metalanguage will also contain the predicates of the object language together with a special predicate, having logical properties, that does not occur in the object language and that designates the relation of a sentence being true of an object. The predicate that designates this relation between sentences and objects will be characterized semantically in such a fashion that the substitution of one term for another in object language sentences will preserve truth-value among such sentences provided that the relevant metalanguage sentence representing the appropriate identity statement is true. In addition, we will take the quantification over sentences in the object language to be substitutional quantification and characterize syntactically the values of the variables occurring only

in the metalanguage as object language sentences that represent just those contexts in which the substitution of co-referring terms is always legitimate. The development of such a metalanguage will obviously require a metalanguage to be thought of in an entirely different manner than is customary. Nevertheless, the approach I have just outlined seems to be worth pursuing since it may well result in an account of identity that has all the benefits of the above-mentioned account, while at the same time circumventing its pitfalls. If my suggested approach to the strategy of defining identity in terms of indiscernibility is ultimately successful, it should have the added bonus of providing an account of identity that takes identity to be a logical notion free of metaphysical notions and consequent metaphysical entanglements.

Bibliography

Blackburn, Simon. *Spreading the Word: Groundings in the Philosophy of Language*. Oxford: Oxford University Press, 1984.

Brody, Baruch A. *Identity and Essence*. Princeton, New Jersey: Princeton University Press, 1980.

Butchvarov, Panayot. *Being Qua Being: A Theory of Identity, Existence, and Predication*. Bloomington, Indiana: Indiana University Press, 1979.

Donnellan, Keith. "Reference and Definite Descriptions." Reprinted in *The Philosophy of Language*, 2nd ed., edited by A.P. Martinich, 235-247. New York: Oxford University Press, 1990.

Evans, Gareth. "The Causal Theory of Names." *Aristotelian Society: Supplementary Volume 47* (1973): 187-208.

Freeman, James B. "Quantification, Identity, and Opacity in Relevant Logic." In *Directions in Relevant Logic*. Kluwer: Norwell, 1989.

Frege, Gottlob. *Begriffsschrift*. In *Frege and Godel: Two Fundamental Texts in Mathematical Logic*, edited by Jean van Heijenoort, 1-82. Cambridge, Massachusetts: Harvard University Press, 1970.

---------. *Begriffsschrift und Andere Aufsatze*. Translated and edited by Ignacio Angelelli. Hildesheim: Georg Olms, 1964.

---------. *Grundgesetze*. In Peter Geach and Max Black, *Translations from the Philosophical Writings of Gottlob Frege*, 137-244.

---------. "On Sense and Reference." In Peter Geach and Max Black, *Translations from the Philosophical Writings of Gottlob Frege*, 56-78.

---------. "Review of Husserl's *Philosophie der Arithmetik*." In Peter Geach and Max Black, *Translations from the Philosophical Writings of Gottlob Frege*, 79-85.

Geach, Peter, and Max Black. *Translations from the Philosophical Writings of Gottlob Frege*. Oxford: Basil Blackwell, 1952.

Ishiguro, Hideo. *Leibniz's Philosophy of Logic and Language*. 2nd ed. New York: Cambridge University Press, 1990.

Kaplan, David, "Dthat." In *Syntax and Semantics*, Vol. 9, edited by Peter Cole, 221-253. New York: Academic Press, 1975.

Kripke, Saul. "Identity and Necessity." In Munitz, *Identity and Individuation*, 135-164.

Kripke, Saul. *Naming and Necessity*. Cambridge, Massachusetts: Harvard University Press, 1980.

Kuhn, Steven T. "Logical Expressions, Constants, and Operator Logic." *Journal of Philosophy* 78, no. 9 (July 1981): 487-499.

Lewis, C. I. *A Survey of Symbolic Logic*. Berkeley: University of California Press, 1918.

Lewis, David, "Counterpart Theory and Quantified Modal Logic." *Journal of Philosophy* 65 (1968): 113ff.

Lockwood, Michael. "Identity and Reference." Reprinted in Munitz, *Identity and Individuation*, 199-211.

Marcus, Ruth Barcan. "Modalities and Intensional Languages." In *Boston Studies in the Philosophy of Science*. Vol. 1. New York: Humanities Press, 1963.

Mason, Richard. *Before Logic*. Albany, New York: SUNY Press, 2000.

McCarthy. "The Idea of a Logical Constant." *Journal of Philosophy* 78 (1981): 449-523.

Morris, Thomas V. *Understanding Identity Statements*. Atlantic Highlands, New Jersey: Aberdeen University Press, 1984.

Munitz, Milton K., ed. *Identity and Individuation*. New York: New York University Press, 1971.

Paulos, John Allen. *I Think, Therefore I Laugh: The Flip Side of Philosophy*. New York: Columbia University Press, 2000.

Peacocke, Christopher. "What is a Logical Constant?" *Journal of Philosophy* 73 (1976): 221-240.

Quine, W.V.O. "On What There Is." Chap. 1 in *From a Logical Point of View*. Rev. 2nd ed. Cambridge, Massachusetts: Harvard University Press, 1980.

---------. "Propositional Objects." Chap. 6 in *Ontological Relativity and Other Essays*. New York: Columbia University Press, 1969.

---------. "Reference and Modality." Chap. 8 in *From a Logical Point of View*. Rev. 2nd ed. Cambridge, Massachusetts: Harvard University Press, 1980.

---------. *Word and Object*. Cambridge, Massachusetts: M.I.T. Press, 1960.

Russell, Bertrand, "On Denoting." *Mind* 14 (1905). Reprinted in *Contemporary Readings in Logical Theory*, edited by Irving Copi and James Gould. New York: The Macmillan Company, 1967.

Searle, John R. *Expression and Meaning: Studies in the Theory of Speech Acts*. New York: Cambridge University Press, 1986.

---------. "Proper Names." *Mind* 67 (1958): 166-173.

Simons, Peter. S.v. "Identity of Indiscernibles." In Vol. 4 of *Routledge Encyclopedia of Philosophy*, edited by Edward Craig. New York: Routledge, 1998.

Smook, Roger. "Logical and Extralogical Constants." *Informal Logic* 10, no. 3 (1988): 195-199.

Weiner, Joan. *Frege*. Past Masters Series. Oxford: Oxford University Press, 1999.

Williams, Christopher J. F. *What is Identity?* New York: Oxford University Press, 1989.

Williamson, Timothy. S.v. "Identity." In Vol. 4 of *Routledge Encyclopedia of Philosophy*, edited by Edward Craig. New York: Routledge, 1998.

Wittgenstein, Ludwig. *Philosophical Investigations*. Translated by G.E.M. Anscombe. 3rd ed. New York: Macmillan Publishing Co., 1968.

---------. *Tractatus Logico-Philosophicus*. Translated by C.K. Ogden. 1922. Reprint, New York: Routledge, 1990.

Ziff, Paul. *Semantic Analysis*. Ithaca, New York: Cornell University Press, 1960.