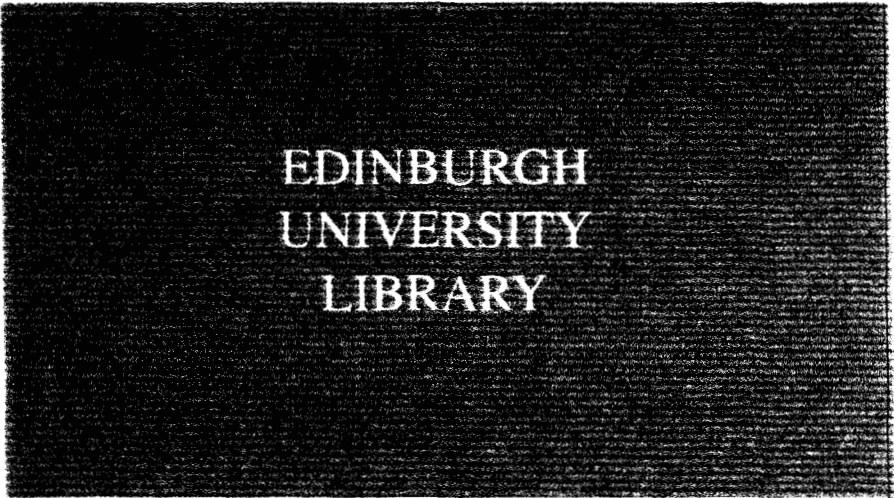




✓



Shelf Mark *Thes Section 2*
 POLLOCK, M. Libb. 2001



The Epistemology of Necessity

William James Pollock

M.Litt., University of Edinburgh, 2001



CONTENTS

Acknowledgements.....	p.3
Abstract.....	p.4
Introduction.....	p.5
Chapter One.....The Direct Reference Theory (Negative Thesis).....	p.8
Chapter Two.....(Positive Thesis).....	p.56
Chapter Three...Putnam on General Terms.....	p.72
Chapter Four.....Identity.....	p.105
Chapter Five.....The Contingent A Priori.....	p.132
Chapter Six.....The Necessary A Posteriori.....	p.163
Conclusion.....	p.195
Bibliography.....	p.199

I, William James Pollock, declare that I have composed this thesis, and that it is my own work

W. Pollock

ACKNOWLEDGEMENTS

I would like to thank the staff and students of the Philosophy Department of The University of Edinburgh for comments on earlier versions of the thesis.

I would also like to thank the Carnegie Trust for the payment of tuition fees during most of my Ph.D. Without such help things would have been much more difficult.

ABSTRACT

The thesis examines the direct reference theory of proper names and natural kind terms as expounded by Saul Kripke, Hilary Putnam and others and finds that it has not succeeded in replacing some kind of description theory of the reference of such terms - although it does concede that the traditional Fregean theory is not quite correct.

It is argued that the direct reference theory is mistaken on several counts. First of all it is question-begging. Secondly, it is guilty of a 'use/mention' confusion. And thirdly, and most importantly, it fails to deal with the notion of understanding. The notion of understanding is crucial to the present thesis - specifically, what is understood by a proper name or natural kind term. It is concluded that sense (expressed in the form of descriptions) is at least necessary for reference, which makes a significant difference to Kripke's claim that there are necessary a posteriori truths as well as contingent a priori truths.

It is also argued that sense could be sufficient for reference, if it is accepted that it is speakers who effect reference. In this sense, sense determines reference. The thesis therefore not only argues against the account of reference given by the direct reference theorists, it also gives an account of how proper names and natural kind terms actually do function in natural language.

As far as the epistemology of necessity is concerned the thesis concludes that Kripke (along with many others) has not succeeded in establishing the existence of the necessary a posteriori nor the contingent a priori from the theory of direct reference. Whether such truths can be established by some other means, or in principle, is not the concern of the thesis; although the point is made that, if a certain view of sense is accepted, then questions of necessity and a priority seem inappropriate.

INTRODUCTION

It has been claimed by Saul Kripke that how proper names and natural kind terms function in natural languages has certain consequences for what is to be regarded as necessary and contingent, and also certain epistemological consequences regarding how necessary truths come to be known. Before Kripke it was more or less taken for granted that necessary truths could only be known a priori and that contingent truths could only be known a posteriori. Kripke challenges this assumption by challenging the semantics of proper names and natural kind terms -specifically, how their reference is determined. Kripke's challenge is to what we might term The Traditional Theory of meaning for names whereby the reference of a name is determined by a sense given by a definite description, or a set thereof. This theory, normally associated with Frege, says that a name 'N' will have a sense, something like 'The F', such that whatever fits the description 'The F' will be the referent of 'N'. Kripke, and various others - mainly Keith Donnellan, David Kaplan and Hilary Putnam - provides several arguments against this description theory and concludes that proper names, and natural kind terms, refer directly and not by means of a sense.

Kripke further concludes that proper names are not only direct designators, but rigid designators; which means that they designate the same thing in all possible worlds (where the appropriate thing exists). Definite descriptions on the other hand are, generally, non-rigid designators - i.e. they can refer to different things in different possible worlds. For example, the name 'Nixon' will refer to the same person in all possible worlds where Nixon exists - namely, Nixon himself. The definite description 'The 37th President of the United states' will refer to different individuals in different possible worlds because it is surely not a necessary truth that Nixon was the 37th president of the United States. There are possible worlds where someone other than Nixon has this distinction. Applying the concept of rigidity to certain empirical discoveries, Kripke overturns our notion of what is to count as necessary. For example, the statement

Hesperus is Phosphorus

may seem obviously contingent, since it was an empirical discovery that Hesperus is Phosphorus, but on Kripke's view this is mistaken because of the rigidity of the names involved. Since both 'Hesperus' and 'Phosphorus' are rigid designators they will refer to the same object in all possible worlds where the object (the planet Venus) exists. Consequently, if it is true in the actual world that Hesperus is Phosphorus then it will be necessarily true because there will be no possible world where Hesperus is not Phosphorus, other than worlds where the planet Venus does not exist. This is what Kripke calls 'weak necessity', or metaphysical necessity, to distinguish it from logical necessity. The main point here for Kripke is that if we regard the names 'Hesperus' and 'Phosphorus' as synonymous with a description which determines reference - something like 'The heavenly body in such-and-such position in the evening/morning sky' - then the statement 'Hesperus is Phosphorus' will be contingent because something other than the planet Venus might have occupied this position in the sky.

In Kripke's scheme the illusion of contingency in such statements as 'Hesperus is Phosphorus' arises from a confusion between descriptions that merely fix the reference of a name and those that are synonymous with the name. He uses this distinction to generate contingent a priori truths in the following way. First of all, a name can be introduced into a language via a description which fixes its reference but is not synonymous with it. For example, the name 'Neptune' could be introduced into the language as a name for whatever is causing certain anomalies in the orbit of the planet Uranus using a description something like 'That which is causing such-and-such perturbations', such that the person who does this would be in a position to know a priori that Neptune is the cause of the perturbations - even though it is contingent that Neptune is the cause of the perturbations.

Thus, Kripke mounts a strong challenge to what is to be regarded as necessary and how necessary truths are discovered, from a theory of semantics. Perhaps the most interesting part of his thesis is the claim that natural kind terms, such as 'water' or 'gold', for example, are rigid designators, which leads to the conclusion

that scientific discoveries such as the discovery that water is H₂O or that gold has atomic number 79, will turn out to be necessary, thus reviving claims that science can discover the real essence of natural kinds. It is the purpose of the present work to challenge Kripke's claims, as well as any supporting arguments due to Putnam, Donnellan etc.

Although I shall grant that some of the arguments of the direct reference theory do force us to make adjustments in the Description Theory of names, I shall nevertheless maintain that the Description Theory has not been refuted. I shall argue that most of the claims made by advocates of the direct reference theory are mistaken. This not only applies to their view of how names function in natural language but also to the claims that there are necessary a posteriori truths and contingent a priori truths; although I cannot rule out such things in principle.

The first two chapters examine the attempt by the direct reference theorists (mainly Kripke) to refute the Description Theory and to replace it with a causal theory of reference, whereby the reference of proper names and natural kind terms is determined by a causal chain of reference preserving links leading from present use all the way back to an initial baptism. Chapter 3 deals with Putnam's attempt to establish a direct reference theory for natural kind terms and also with what I regard as the true theory of how names function in natural language. Chapter 4 deals with the meaning of identity statements, and challenges the claims by Kripke and Ruth Marcus that identity statements involving proper names and natural kind terms are necessary. Finally, chapters 5 and 6 examine the claims made by Kripke (and others) that there exist contingent a priori truths as well as necessary a posteriori truths.

I shall not say a great deal about possible worlds or about the precise meanings of the terms 'a priori' and 'a posteriori'. I shall simply follow Kripke's usage in *Naming and Necessity* where he defines a priori truths as "those which can be known independently of any experience" (1980,p 34). A posteriori truths, on the other hand, are those which have been discovered empirically, such as the discovery that water is H₂O or that gold is the element with atomic number 79. Possible worlds will be regarded non-realistically - "*stipulated*, not discovered".

CHAPTER ONE

Perhaps nothing illustrates the gulf between philosopher and layman better than the philosophical treatment of the semantics of proper names, such as, for example, 'Socrates', or 'London', or 'John Smith'. What to the man in the street are the most mundane and straightforward of linguistic items have presented philosophy with over a century of seemingly very serious problems with far-reaching consequences. In this chapter we will be concerned with the two main contenders for the true theory of what is to count as a genuine name and how they function in language: specifically, natural language. These are the Direct Reference Theory (hereafter, DRT) which holds that names refer to their objects directly, without the mediation of any kind of sense; and what might be called the 'Description Theory', or 'Sense Theory', which is the view that proper names refer indirectly by means of a sense. This sense is usually identified with some definite description or set thereof, and for a long time seemed to give the true picture of how names functioned. Senses, for example, seemed to be indispensable to any attempt to solve the so-called 'Problems of Reference'. As is well known Frege (1892) introduced his theory that names have sense as well as reference in order to solve the problem, as he saw it, of how true identity statements of the form $a=b$ can differ in 'cognitive value' (*Erkenntniswerte*) from those of the form $a=a$. Also in this paper Frege uses the theory of sense to solve the problem of the behaviour of proper names in belief contexts which seems to lead to paradoxical results, as in the following example:

- 1) John believes that Hesperus is a planet
- 2) Hesperus is Phosphorus
- 3) *Ergo*, John believes that Phosphorus is a planet

This looks like a perfectly valid argument if 'Hesperus' and 'Phosphorus' are co-referential; and yet, surely John can consistently deny that he believes that Phosphorus is a planet.

Frege's solution to the problem in terms of the sense of proper names is to deny that names have their customary reference in such contexts. Rather, they refer to their customary sense. This means that in 1) what John believes is different from what he believes in 3). 1) and 3) express different 'thoughts' (in the non-psychological sense) or propositions. The question of what a proper name contributes to propositional content is a major difficulty for DRT, as we shall see.

These two puzzles, along with the problems of negative existential statements and vacuous names, like 'Pegasus', seem to pose serious problems for any theory of naming which does not allow a fundamental role for sense. Nevertheless, in the last few decades the orthodox sense theory has been severely criticised and gradually replaced in the minds of many by the direct reference theory, a revival of the theory of naming advocated by J.S. Mill in his 1843 "A System of Logic", where he argues that proper names have denotation only, and no connotation. Reference is direct and unmediated by anything like a sense.

The chief advocate of DRT in recent times has been Saul Kripke in his works "Identity and Necessity" (1971) and *Naming and Necessity* (1972, 1980) and no doubt the success of DRT in overthrowing the Fregean orthodoxy in spite of the difficulties of the problems of reference is a tribute to the considerable persuasive power of Kripke's arguments which seem to provide a devastating critique of the description theory. Before proceeding to discuss Kripke's views we must be clear about what it is that DRT is opposed to, since it may seem obvious to some people that names have sense in the sense that we associate senses with names, in our minds. The name 'Neil Armstrong', for example, will be associated with the description 'The first man on the moon'. No direct reference theorist could or would deny this. Kripke even concedes that we can *fix the reference* of a name with a definite description, although typically we do not do this, and there is certainly no need to (1980, p.15, 54 - 57). The crucial point for Kripke is that reference-fixing descriptions and descriptions associated with names in the minds of speakers are not

what *determines the reference* of the names. They are therefore not synonymous with the names. Reference is direct and unmediated.

Kripke's starting point is a purely formal one, namely, the model-theoretic treatment of quantified modal logic; in particular the theorem:

$$4) (\forall x) (\forall y) (x=y \rightarrow \Box x=y)$$

This theorem makes it obvious to Kripke that identical objects are necessarily identical. Once we assign values to these variables we get the necessary identity of a particular instance, as in :

$$5) (a=b) \rightarrow \Box (a=b)$$

It seems natural that 'a' and 'b' should be ordinary proper names like, for example, 'Hesperus' and 'Phosphorus'. If proper names refer directly, and therefore rigidly (i.e. they refer to (or designate) the same object in all possible worlds, where the object exists - what Kripke calls 'weak rigidity') we get (5) for ordinary proper names. Assigning values to variables in a formal language does not seem to require anything like a sense; but of course, the situation could be different when we move away from formal languages to natural languages. It is the functioning of ordinary proper names in natural languages that is the concern of *Naming and Necessity*. More precisely, it is the way in which the functioning of ordinary proper names relates to modality and epistemology that is the concern of this famous text. For Kripke, it is a consequence of the direct reference theory, specifically the claim to rigidity, that there are **necessary a posteriori** truths as well as **contingent a priori** truths. When applied to general terms such as natural kind terms Kripke is led to resurrect Aristotelian essentialism - the claim that objects, and kinds of objects, have some of their properties essentially. This is real essence and not just nominal essence arising out of our way of describing an object or a kind. What is more, this essence is discoverable by science. For example, the empirical discovery that water is H₂O is the discovery of a necessary truth, in Kripke's view. Science has

discovered the essence of water, or so it is claimed [Particularly, 1980, p128-134, and 140-146]

Such bold claims require very strong arguments, and this Kripke thinks he has produced in *Naming and Necessity*. In this claim he is supported, in one way or another, by such luminaries as Keith Donnellan (1972), David Kaplan (1968,1975), Hilary Putnam (1973,1975a), and Nathan Salmon who seeks (in his 1986) to deal with the seemingly fatal difficulty of the contribution made by a proper name to a sentence contained in a belief context.

Naming and Necessity consists of two theses - a negative one and a positive one. The negative thesis is Kripke's allegedly devastating critique of what he terms the 'Frege/Russell' view of proper names whereby a proper name 'N' is synonymous with, or abbreviates, to use Russell's term, a definite description which gives the sense of the name, and by means of which the name refers. The positive thesis consists of Kripke's attempt to explain, or give a 'better picture', of how names actually do work and how their semantic properties of direct and rigid designation relate to the epistemological questions of how we come to know necessary truths and contingent truths. It will be the aim of the present chapter to show that Kripke, and his aforementioned confederates, has singularly failed to establish either the negative or the positive thesis. It will also be argued, however, that all is not well with the description theory and that a new theory of the semantics of names is required.

The Negative Thesis

(The Argument Against the Description Theory)

In Lecture I of *Naming and Necessity* Kripke defines the description theory as the theory that "...really a proper name properly used, simply was a definite description abbreviated or disguised." (1980, p27), a view he attributes to Frege and Russell. It is by means of this sense that the name refers. Both Dummett (1973) and Linsky (1977) have claimed that Frege held no such view about the sense of a proper name

being just that of a definite description. Linsky further claims that Russell did not hold the view either if it means that names refer by means of definite descriptions, since definite descriptions do not refer in Russell's scheme; nor do they have a sense. (Op.Cit. p43,44). Both Dummett and Linsky are adamant that Frege never held the view Kripke attributes to him. Although Dummett concedes that Frege expresses possible senses of proper names by means of descriptions, he is quick to point out that these descriptions should be considered as "...merely a device for a brief characterisation of a sense, rather than as a means of conveying the thesis that Kripke ascribes to Frege." (Op.Cit. p 110). The crucial point for Dummett and Linsky is that a name have associated with it a *criterion* for recognising a given object as the referent of the name (Dummett, Op.Cit. p110; and Linsky, Op.Cit. p71) According to Dummett this criterion can sometimes be conveyed by a definite description, sometimes not. Linsky, on the other hand, is reluctant to talk of definite descriptions at all, preferring to characterise this criterion of identification as a *concept* of the thing in question (Op.Cit.p71) For both Dummett and Linsky Frege's theory of proper names is undoubtedly a sense theory but not a description theory - or at least not the description theory attributed to Frege by Kripke.

This denigration, by Dummett and Linsky, of the role of descriptions in giving the sense of proper names is hard to accept. After all, how else is the criterion of identification to be characterised? It is all very well for Linsky to argue that anyone who acquires the respective senses of the names 'Hesperus' and 'Phosphorus' acquires two different concepts or two different criteria of identification of a particular object, but we must surely ask what this difference consists in, and how is the difference to be expressed if not in terms of descriptions? It is easy to agree with Strawson when he writes:

....it is no good using a name....unless we know who or what is referred to by the use of the name. A name is worthless without a backing of descriptions which can be produced on demand to explain the application (1959, p20)

At any rate, what is clear is that Frege held that there is more to a proper name than its reference. It will have a sense which determines its reference. It is equally clear that Russell's treatment of vacuous or non-referring names lead him to conclude that what we ordinarily call names are not genuine names at all. They are

really truncated descriptions. What Kripke may be guilty of is conflating the views of Frege and Russell into the thesis that a name is synonymous with a definite description which determines what the name refers to - a view which perhaps neither Frege nor Russell actually held. It must be remembered that Frege and Russell differ enormously as to what is to count as a genuine name - to such an extent in fact, that any combination of their respective views would lead to a gross distortion of both positions.

However, we do not need to look too closely at the text of *Naming and Necessity* to see that Kripke is attacking a, broadly, Fregean view of naming rather than what he calls the 'Frege/Russell' view. Linsky is surely correct when he says that Russell's theory of proper names does not include the claim that names refer by means of some definite description. For Russell, what we ordinarily regard as proper names are not genuine names at all; and definite descriptions do not have sense.

The view being attacked in *Naming and Necessity* is the view that proper names refer indirectly by means of a sense given by a definite description (or perhaps a set of descriptions) where the description is used attributively, in Donnellan's sense (Kripke, 1980, p26,28). It is very difficult for a defender of Frege to deny that this is, or ought to be, the Fregean thesis. This is the thesis that we shall examine in the present chapter along with numerous variations designed to avoid the criticisms of DRT. These will include the Cluster Theories of Wittgenstein and Searle (and, as it turns out, Linsky); the so-called 'meta-linguistic' theories of Jerrold Katz and Kent Bach ('sense-only' theories); plus other attempts at a compromise between DRT and the Fregean Description Theory.

Kripke has three main arguments against the description theory which we may term '**Epistemological**', '**Modal**', and '**Semantic**'. These are best illustrated by an example:

(A) Aristotle was the teacher of Alexander

or

Aristotle taught Alexander

According to Kripke, if 'Aristotle' just *means* 'the teacher of Alexander', then (A) should be **a priori, necessary and analytic**.

i) (A) is not a priori because it was an empirical discovery that Aristotle taught Alexander

ii) (A) is not necessary because we can surely make sense of a counterfactual situation (possible world) where Aristotle is not a teacher

iii) (A) is not analytic because *it may not even be true*. It is epistemically possible that it is not true. I.e. we could yet discover it to be false.

(A) is therefore **a posteriori, contingent and synthetic**, since it tells us something about the world and not about the meanings of words.

The first point to note here is that Kripke is on rather shaky ground when he claims that (A) is not true by virtue of the meanings of the words, since he does not make it clear what he means by 'meaning'. He does make the celebrated distinction between fixing a reference and giving a synonym (1980, p15,37, 53-60) - definite descriptions may be used for the former purpose but not the latter - but this merely tells us what meaning is **not**. To give a positive account of meaning would involve Kripke in some account of the propositional content of sentences containing proper names, and this, notoriously, is not forthcoming in Kripke's work. This is probably because, if names lack sense given by definite descriptions, their contribution to content can only be in terms of their reference. In the case of co-referential names, this leads Kripke head-on into the problems of reference, especially the 'Identity Problem' and the 'Propositional Attitude' problem. (Actually not all advocates of DRT deny that names have sense - for example, Salmon, (1981) where, following Burge (1977, p 356), he distinguishes different roles for senses - a reference-determining role, a psychological role, and a role whereby co-referential names can contribute different senses to propositional content. However, Salmon later recants this piece of heresy and argues that proper names do not have sense at all).

It seems, however, that DRT can avoid the problems of vacuous names and non-referring names, such as 'Pegasus', which troubled Russell so much, by simply

denying that they are genuine names. They are not genuine names simply because they do not in fact name anything. Consequently, DRT does not owe any account of how these 'names' function in these circumstances. Actually it could be the Fregean who has the problem with non-referring names. He wants to claim that they are genuine names because they have a sense, but he then has to explain how a genuine name does not actually name anything. Frege's ad hoc move of providing an arbitrary reference for these terms, such as the number zero, will not do for any theory of natural language (for which it was never intended). Whilst it may be acceptable formally, in a natural language such as English, it is barely intelligible. What, for instance, does it mean to say that the name 'Pegasus' denotes the number zero? The only possible way in which the Fregean can deal with non-referring names is to make something of the fact that the sense of these terms *contains* the fact that they do not refer. For example, it is an integral part of the sense of 'Pegasus' that Pegasus is a *mythical* creature, and therefore did not exist. Not to include this when giving the sense of 'Pegasus' is to give a sense that is seriously deficient and misleading. To take another example, the sense of 'Sherlock Holmes' is 'the *fictional* detective created by Conan Doyle etc.' This solves the problem of non-referring names in terms of sense. The sense of these names includes the fact that these 'names' do not and never did refer to any actual object. Something similar can be done with theoretical terms, such as 'neutrino', for example, whose sense can be given by the description "The *theoretical* particle believed to have such-and-such properties". Despite having sense, however, these terms are still not genuine names since they do not (or may not) name anything. They are not nothing, however. They have sense and can be used in meaningful discourse. We could call them 'Pseudo-names' to distinguish them from genuine names.

So long as the notion of meaning is unclear, then, the notion of analyticity must also be unclear, since Kripke defines analyticity as truth by virtue of the meanings of the terms involved (1980, p39).

The Epistemological Argument

To deal with the epistemological objection first, (A) certainly does not seem to be the sort of thing we could know a priori. It was undoubtedly an empirical discovery that Aristotle taught Alexander, so (A) would seem to be the very paradigm of a posteriori knowledge; unlike, for example:

(B) Bachelors are unmarried

which seems to be the very paradigm of a statement which is necessary, a priori and analytic. However, we surely only know a priori that bachelors are unmarried once we know what the word 'bachelor' *means* - and this we discover empirically. Why, then, can we not say something similar about (A)? Why not say that we know a priori that Aristotle was a teacher, once we know what the name 'Aristotle' means? This is an empirical discovery, but then so it is with the meaning of any word. When we discover that Aristotle taught Alexander we not only discover something about Aristotle himself, we also discover something about the meaning of the *name* 'Aristotle'. Now this may seem a strange claim to many philosophers because it means that any new discovery about Aristotle, or whatever, leads to a *shift* in the meaning of the name 'Aristotle'. It also leads us to question whether any and all discoveries about Aristotle contribute to the meaning of his name. This is not such a strange claim, however. *If* 'Aristotle' has a sense given by a set of descriptions, then why not add to the sense of the name when we discover new things about Aristotle? After all, we surely have to do something with this new information. As to the question of what exactly is to be included in the sense of the name this will depend on the context in which the name is being used. What seems like a trivial piece of information in one context could be of great significance in a different context - although, ultimately, and strictly speaking, anything and everything that is true of Aristotle will contribute to the sense of his name. This will also apply to general names, as we shall see in chapter 3.

Whether we discover the whole meaning of 'Aristotle' or only part of the meaning may determine the epistemic status of (A). According to Linsky it does. If,

for example, being the teacher of Alexander exhausts our knowledge of Aristotle, in the way in which being the mother of Mary (to use Linsky's example) exhausts our knowledge of the name 'Saint Anne', then it would be a priori that Aristotle taught Alexander, just as it is a priori, in Linsky's view, that St. Anne is a mother because "...Having acquired the name 'St. Anne' by means of the single definite description 'the mother of Mary', I cannot make intelligible to myself the supposition that, though she exists, she is not a mother" (1977, p83) (Philosophers like Kripke and Donnellan would find this perfectly intelligible. These diametrically opposed intuitions lie at the very heart of the debate between DRT and the description theory).

The situation is somewhat different, according to Linsky, in the case of a name where we acquire a whole set of descriptions which we believe express knowledge about a person. Linsky cites Wittgenstein's example of the name 'Moses' in *Philosophical Investigations* (p36,37) where Wittgenstein questions whether 'Moses' has a fixed and unequivocal use in all possible cases because of the variety of descriptions we could substitute for the name. We would not therefore, know a priori that Moses was 'The F', or 'The G', or 'The H', because of this lack of "a fixed and unequivocal sense ...in all possible cases" (Linsky, Op. Cit. p86). Thus Linsky concludes, significantly, that "...under these circumstances...the notion of a priori truth becomes impossible to apply" (Ibid.)

Note that in Linsky's view Wittgenstein's 'Cluster Theory' is very different from that of Searle (1958) who does seem to hold that names have a fixed and unequivocal sense. Searle agrees that to use a single description as giving the sense of a name would commit us to claims about necessity that are clearly wrong - such as the claim that it is a necessary truth that Aristotle taught Alexander. Searle recognises that such a truth is clearly contingent. He does, however, argue that it is a necessary fact that Aristotle has the logical sum, inclusive disjunction, of the properties commonly attributed to him (Searle, 1958, in Caton, p160) Whatever has most, or a weighted most, of these properties, is Aristotle. If most of these properties do not pick out any individual uniquely, then we must conclude that Aristotle did not exist. Similarly, if the Biblical story of Moses turns out to be substantially false,

then we would conclude that Moses did not exist, according to the cluster theory of Searle.

Kripke is quite clear that these claims are just false. As he says in *Naming and Necessity*: “I have already argued (p58 -WP) that the Biblical story does not give necessary properties of Moses, that he might have lived without doing any of these things....The Biblical story might have been a complete legend, or it might have been a substantially false account of a real person” (p66). Likewise at page 62: “Not only each of these singly but the possession of the entire disjunction of these properties is just a contingent fact about Aristotle..”

At first sight at least, it would seem to be Kripke who has the correct intuition here, since it would seem to be straightforwardly the case that Aristotle and Moses could have existed and not done any of the things they actually did, or are believed to have done. That things are not quite so straightforward when deciding whether or not names have a descriptive sense is best illustrated with a thought experiment. Suppose, for instance, that we could build a time machine and travel back to the time (and place) that Moses is supposed to have lived. If we try to identify Moses the only way we can do so is by looking for someone who fits all, or most, of the descriptions we associate with the name. If we did come across such a person (who need not actually be called ‘Moses’) then the Fregean would conclude that we had found Moses. Kripke, on the other hand, would have to say that this person is not necessarily Moses. In fact Kripke is committed to the view that we have no more right to conclude that this person is Moses rather than *any* of the other people we encounter. Moses could be any of these people, or none. The fact that there is an individual who fits all our descriptions does not help us at all.

I think this thought experiment may help even up the score between Kripke and the Fregean as far as intuition goes. It is surely very strange to claim that someone who fits all or most of our descriptions has no more claim to be Moses than anyone else, but that is what Kripke is committed to. (We shall deal with the way in which reference is determined for DRT later).

It is clear, then, that Kripke sees no real difference between the single description and the cluster theory. Linsky agrees with Kripke in his rejection of Searle, but sees a distinction between Searle and Wittgenstein, as already described, which Kripke

does not recognise. Linsky concludes his (1977) by maintaining that the Wittgensteinian cluster theory, which holds that we will give up a name if enough of the ground is cut from beneath us, is to be preferred to any other description theory because “ we cannot say now exactly what that ground is. It becomes fixed as we go along. Or else it does not and we do not know what to conclude from our investigation” (Ibid. p111)

The obvious difficulty with this view is how to determine what descriptions are to be relevant ‘as we go along’. In other words the Linsky/Wittgenstein cluster theory is too loose, too vague to be of any use to us. Any theory of the sense of a proper name must be more rigid than this cluster theory if it is to deal with the problems of reference. Can a name like ‘Moses’, for example, really have one sense in order to deal with Frege’s identity puzzle and another sense in order to deal with the problem of propositional attitudes? It must surely have the same sense in both cases. Note, that this is not the same problem as the problem of different speakers associating different senses with the same name. This is undoubtedly a fact and will have to be dealt with by the description theory. The problem with the Linsky/Wittgenstein view is that the *same* speakers do not have a fixed sense for a name like Moses. Linsky does point out that existence questions such as ‘Did Moses exist?’ are vague but we can remove this vagueness in a given context by fixing on a definite description or cluster for which we will take these names as ‘abbreviations’ (1977, p111). What he fails to take into account, however, is that there are other contexts apart from questions of existence (the only problem he deals with in his 1977), other contexts where we need to fix on a definite description or cluster thereof in order to give a sense to the name. There is no reason to doubt that the descriptions should be the same in each context - i.e. for all the problems of reference. Any theory of sense which expresses that sense in terms of descriptions must be more rigid than the Linsky/Searle/Wittgenstein view.

Of course if we reject this loosening of the sense of a proper name we find ourselves once again faced with Kripke’s charge that the truth of a statement like (A) should be known a priori. As already mentioned, however, this may not be the problem that Kripke takes it to be. Once we discover that Aristotle taught Alexander we do know certain things a priori, for example, that Aristotle was a teacher. This is

because, in discovering that Aristotle is The F, we also discover that it is at least *part* of the *meaning* of 'Aristotle' that Aristotle is The F. This seems a perfectly legitimate way of regarding the question of sense for proper names. We discover their senses empirically. This is no different from the way in which we discover the sense of any word. It is often claimed that we know a priori that bachelors are unmarried or that vixens are female foxes, but this is only the case once we know what these words mean - and this we discover empirically. Kripke's epistemological argument, which initially had great intuitive appeal, is not nearly as obvious or as strong as he claims. At the very least, it is not clear that there is such a difference between statements like

Aristotle taught Alexander

and

Bachelors are unmarried

as far as questions of analyticity and a priority are concerned.

The Modal Argument

Kripke seems to regard his modal argument as his central objection to the description theory; although it has been regarded as his least convincing by most commentators, who mainly accuse him of trading on an illicit shift of scope. (For example, Dummett, Op.Cit.; Linsky, Op. Cit. and Loar, 1976) (More of this below). The modal argument can indeed be regarded as an argument about scope, or it could be seen simply as an argument from counterfactuals. It is also intimately connected with Kripke's most crucial idea - that of the **rigidity** of proper names (and general terms, as it turns out). The modal argument goes something like the following:

(C) (i) If 'N' meant 'The F', then 'N might not have been The F' would be false

(ii) 'N might not have been The F' is true

(iii) *Ergo*, 'N' does not mean 'The F'. (Due to Loar, 1976)

For example, although Nixon was the U.S. president in 1970, we can make sense of the possibility that he might not have been. Kripke's contention is simply that, if 'Nixon' meant 'The U.S. president in 1970', then this would make it a necessary truth that Nixon became president, which violates all our intuitions about necessity and what we can suppose counterfactually. Furthermore, in talking about a counterfactual situation where Nixon does not become president - i.e. a possible world where someone other than Nixon is U.S. president in 1970 - we are talking about a situation where *Nixon - that very man* - is not president. This is the source of the claim to rigidity of ordinary proper names like 'Nixon' or 'Aristotle'. For Kripke, we have a "direct intuition" of the rigidity of proper names (1980, p4, 14, 48). This intuition is supported by something Marcus says about names (1976):

We need a way in language of sorting a thing from its properties so that we may entertain the possibility of its not having this or that property. Proper names serve this purpose.....Proper names are a way of mentioning an object without a commitment to any of its properties excluding trivial ones.....It appears therefore that for the analysis of counterexamples the distinction between proper names and other modes of fixing reference is central. (1976, p134)

In talking about counterfactual situations it is *given*, by stipulation, that we are talking about Nixon (or whoever) - not someone who satisfies certain descriptions; and certainly not what David Lewis calls a "counterpart" of Nixon. Such a counterpart would be of no use to Kripke in talking about counterfactual situations or possible worlds (See 1980, p45n). Our intuitions about counterfactual reasoning only make sense if trans-world identity of individuals is given, by stipulation. Indeed, for Kripke the trans-world identity of individuals is something of a pseudo-problem. As he says on page 44 of *Naming and Necessity* : "A possible world is given by the descriptive conditions we associate with it." These descriptive conditions can contain the fact that Nixon exists in that world (Ibid.)

Another way of looking at the problem of rigidity or non-rigidity is in terms of how we evaluate terms with respect to possible worlds. For Kripke, names and

definite descriptions are to be evaluated differently. In a footnote to his (1972) and (1980) he explains the point thus:

In the formal semantics of modal logic, the sense of *t* is usually taken to be the (possibly partial) function which assigns to each possible world **H** the referent of *t* in **H**. For a rigid designator such a function is constant. (1972, p 346n)

In evaluating a definite description, on the other hand, we must go to a possible world and find the object, if any, that uniquely fits the description. This may vary from world to world. The evaluation function for definite descriptions is not constant. More simply, a rigid designator designates the same object in all possible worlds, whereas a definite description can designate different objects in different possible worlds. Actually some qualification is required here. Kripke feels that he must distinguish between, what he calls ‘**strong**’ and ‘**weak**’ rigidity, and between ‘**De Facto**’ and ‘**De Jure**’ rigidity. A designator is weakly rigid if it designates the same object in all possible worlds, where the object exists. (Ordinary proper names would fall into this category). A strongly rigid designator, on the other hand, designates the same object in all possible worlds, simpliciter. Such a designator would designate necessary existents, such as numbers (1980, p48). The de facto/de jure distinction is introduced in order to deal with the fact that some definite descriptions are rigid - for example, ‘The square root of 4’, or ‘The even prime’. Such designators are merely rigid *de facto*, according Kripke. That is, they merely turn out to be rigid. Proper names, on the other hand, are *de jure* rigid; which means that they are stipulated to designate the same object in all possible worlds: or, as McGinn (1982) puts it, they have constant reference “in virtue of the semantical rules of the language”

Kripke can also appeal to the de facto/de jure distinction in order to deal with the allegation that any definite description can be made rigid by prefixing it with the operator ‘actually’ (See Linsky, Op.Cit. p58, and Plantinga, 1978). Thus, the non-rigid ‘U.S. president in 1970’ becomes the rigid ‘The *actual* U.S. president in 1970’. As Plantinga (1978) puts it, the property of being U.S. president in 1970, has been ‘world-indexed’ - indexed to the actual world, which he calls ‘ α ’. In general, a definite description such as ‘The F’ can always be indexed to the actual world,

whence it becomes 'F-in- α '. Clearly such a description will be rigid since there are no possible worlds where Nixon exists and does not have the property of being the U.S. president in 1970 in the actual world. Such a move, according to Plantinga, makes the thesis that names are rigid designators compatible with the description theory. On this view names do indeed express properties of objects, but not ordinary properties, such as 'The F', but *world-indexed* properties like 'The F-in- α ' (Plantinga, 1978, p131). These properties are *essences* of the objects in question (Ibid.p132). Each object will, according to Plantinga, have several essences (p132) which are logically equivalent but not epistemically equivalent (p133). Co-referential proper names will also express logically equivalent but epistemically inequivalent essences. This would explain how a statement like 'Hesperus is Phosphorus' can be informative, while also explaining how statements like 'Hesperus is a planet' and 'Phosphorus is a planet' can express epistemically inequivalent propositions.

Despite the obvious attraction of a theory of rigid, though indirect, designation which solves the problems of reference, Kripke is unimpressed by the use of an actually operator in this way (See 1980, p59n). For Kripke, and DRT theorists in general, names are thoroughly non-descriptive. They do not express any properties of objects, other than possibly the property of being that very object - what has been termed the 'Haecceity' of the object. The kind of 'essences' posited by Plantinga are not the kind that interest Kripke. Joseph Almog (1986) calls them 'manufactured essences'... "semantic assembly-line essences" that substitute semantical theft for metaphysical toil (1986, p224n). Salmon (1981) also doubts the efficacy of the actually operator in rigidifying definite descriptions. Specifically, he reminds us that the property of being the actual F, or whatever, is not a purely qualitative property of the sort that the orthodox (Fregean) theory demands (Op.Cit. p27). In other words, Salmon wants to know if the actually operator describes the actual world as the only possible world that has the ontological status of being actual; or does it directly and non-descriptively indicate the actual world for semantic evaluation? (Ibid. p44). Plantinga does not consider such subtleties but, then, perhaps he has no need to since his is not an orthodox Fregean theory.

Despite any prima facie plausibility in Plantinga's argument, I shall argue that definite descriptions cannot be rigidified in the way he claims. While not wishing to denigrate the use of the actually operator in general (See, for example, Davies and Humberstone, 1980, where the actually operator is argued to be essential to a correct logic of possibility) it seems clear that it cannot be used to rigidify a definite description. Such a procedure has certain consequences for essentialism, which are, at the very least, seriously counter-intuitive, and at worst, downright absurd. For example, it is a consequence of the type of view advocated by Plantinga that, while Nixon has the property of being the winner of the 1968 presidential election only contingently, he has the property of being the actual winner, or the winner in the actual world, essentially. What, we may ask is the difference between being the winner and being the actual winner? After all, we are not dealing with two events here - Nixon's winning the election, and Nixon's actually winning the election. In the sentence:

(D) Nixon actually won the election

the actually operator is as redundant as the word 'now' in

(E) It is raining now

'Now' is redundant because we are using the present tense. There is no point in having such a thing as the present tense if every use of it has to be augmented by the word 'now', or the expression 'at the present moment'. We are in a similar situation as regards the word 'actually'. Take, for example, the following sentence:

(F) It was an empirical discovery that Aristotle taught Alexander in the actual world

Here the actually operator is clearly redundant, since any empirical discovery could only be of the actual world. We could not possibly be speaking truly about some other possible world here. Plantinga's world-indexed properties look like a prime

candidate for Occam's Razor, if we consider that it is a consequence of indexing properties to the actual world that, for any object that exists in the actual world, for any property it has, it will have the *further* property of *having-that-property-in-the-actual-world*. This just seems totally unnecessary. It may even be perverse.

Perhaps the best argument against the distinction between a property such as The F and The actual F is that it leads to an infinite regress of essential properties as follows. First of all, let us, with Plantinga, call the actual world ' α ', and distinguish between F and F-in- α . Now if F-in- α is an essential property of, say, Aristotle, then he must have this property in all possible worlds where he exists - and this must include the actual world. So he has F-in- α , in α . Now if we go to any arbitrarily chosen world where Aristotle exists and ask the question 'Who in this world has the property F-in- α -in- α ?', then it has to be Aristotle. He has this property in any world where he exists. The key question now is 'Does he have this property [(F-in- α)-in- α] in the actual world?' If he does then we have another essential property on our hands. We are also on our way to an infinite regress. The only way to avoid the regress is to deny that Aristotle has the property (F-in- α)-in- α in the actual world, or to deny that he *has* to have it in the actual world. But then if he need not have (F-in- α)-in- α in the actual world, why must he have F-in- α in the actual world? One reason why he must have (F-in- α)-in- α in the actual world is that it would be a very strange property that something has in all possible worlds *except* the actual world. That surely makes no sense at all. If we accept this then Aristotle must have (F-in- α)-in- α in the actual world and so we have the regress. At any rate, we either have the regress or F-in- α is not an essential property of Aristotle. The alleged distinction between being The F and being The actual F cannot be maintained. It would seem that the case for the rigidification of definite descriptions has not been successful.

However, there are still many problems ahead for the rigidity thesis. Kripke's distinction between weak and strong and de jure and de facto rigidity, with which he sees no real problem, are strongly criticised in Almog (1986). Almog's main argument is that Kripke's claim to be able to establish Aristotelian essentialism from a semantic theory is question-begging. (Salmon's 1981, makes the same charge against Kripke, Putnam and Donnellan regarding natural kinds). For Almog, the crucial notion of rigidity is actually modally oriented. That is, Kripke begins, not

with the *designators*, but with the *designata* and their modal properties. Thus 'Nixon' is a rigid designator since the man Nixon is such that he could not have failed to be Nixon. He has the property of being Nixon essentially. On the other hand Nixon has the property of being the U.S. president in 1970 only contingently, and so the description 'The U.S. president in 1970' is non-rigid. In Almog's view, this subjugation of semantics to metaphysics has adverse consequences for Kripke's attempt to delineate the class of genuine naming devices. First of all, it allows too many designators in as genuine names - for example, 'The offspring of gametes G', which rigidly designates Nixon (according to Kripke); whereas we would not normally regard it as a name. Secondly, as already mentioned, it leads Kripke to the distinction between strong and weak rigidity - strongly rigid designators designating necessary existents like numbers, and weakly rigid ones designating contingent existents, such as people, or in fact most objects. According to Almog such a distinction should not have been drawn for genuine naming devices. If Kripke had begun by attending to the *designators* rather than the *designata* he would have found a very different situation. For instance, from the point of view of language, there is nothing to prevent, say, the numeral '5' from signifying a contingent existent such as a person - in fact, anything at all. Moreover, if Kripke had stuck with the notion of possibility he employed in his early work on modal semantics, namely *logical* possibility, then so-called 'rigid designators' like 'The offspring of gametes G', or even 'The even prime', would no longer be rigid, unless we "sneak in arithmetic essentialism to restrain the range of possibilities" (1986, p217)

As well as the modal orientation, Kripke's crucial notion of rigidity also presupposes a background semantic framework concerned with evaluation in possible worlds. Almog describes this as a one stage theory of evaluation in which there is but one step to go through in order to determine the truth-value of a given sentence - evaluate it in a world. Of course, for Kripke, this would be a *metaphysically* possible world and so, in conflating the two ideas of the one-stage theory of evaluation and the modal orientation, he initiates what Almog calls 'the metaphysical turn' in semantics.

Almog distinguishes between a *mild* one-stage theory, which simply accepts that sentences should be evaluated in a world, and a *strong* one-stage theory, which

holds that contents of sentences are nothing over and above *evaluation functions* - i.e. the functions which assign to possible worlds truth-values in accordance with the 'facts' of these worlds. Now Kripke is a mild one-stage theorist only, in Almog's view -i.e. he does not say *anything* about content. He does not tell us what a sentence like 'Nixon is a Republican' actually *says*. According to Almog, we need a two-stage picture where we, first of all generate a proposition, where the name will contribute the object to which it refers to the propositional content of the sentence which contains it - what is termed (after Kaplan, 1975) a 'singular proposition'. Only then do we go on to evaluate it. This two-stage picture, and not Kripke's, is genuine Millianism. Kripke's theory cannot delineate the class of genuine naming devices.

To summarise Almog's argument, then, Kripke cannot distinguish between rigidity and non-rigidity because of his illicit 'metaphysical turn' - i.e. putting the cart before the horse. Neither can he, in Almog's view, distinguish between de facto and de jure rigidity, because any de facto rigid designator can be made rigid de jure by prefixing the description with the actually operator. Such a move does not have the absurd consequences previously mentioned for Plantinga, since, in Almog's scheme of things, the actually operator does not depend on any metaphysically based essential properties to guarantee its modal stability. For Almog, 'The actual F' is made de jure rigid "in virtue of the semantical rules of the language, in virtue of the truth-conditions stipulated for all formulas containing 'actual'" (Op.Cit. p223) He goes on: "Whereas nothing less than full-blooded arithmetic essentialism must be relied upon to verify the rigidity of 'The smallest prime', nothing more than pure semantics is called upon to demonstrate the modal stability of 'The actual smallest prime' (Ibid. p224)

Again I fail to see the justification for the use of the actually operator in this type of case, or even the motivation for it. The only possible motivation for prefixing a definite description like 'The teacher of Alexander' with 'actually' is to make sense of the fact that there are possible worlds where someone other than Aristotle has this property. (This is Plantinga's motivation) It has already been argued that such a move is superfluous. The move made by Almog is just as pointless. In a way he is in a worse situation than Plantinga since there is no possible motivation for the move

which Almog advocates. Whereas someone may want to make a clear distinction between the actual teacher of Alexander and some possible teacher in some other possible world, there is no need to distinguish between the smallest prime and the actual smallest prime and some other possible candidate for being the smallest prime. There are no other possible candidates other than the number one. ‘The smallest prime’ is rigid simply because it denotes the number one. It cannot denote anything else without changing the meanings of the words involved.

Almog has not succeeded in repudiating the distinction between de facto and de jure rigidity; but what about his claim that the basic rigid/non-rigid distinction is modally oriented? It seems that Almog is probably right about this and that his two-stage picture of evaluation with its complete lack of modal involvement, should be the true Millianism. It also has the distinct advantage over Kripke of giving us an account of content and the contribution made by proper names. However, as we shall see, the problem of content leads DRT into more trouble than it can deal with.

In the meantime we shall deal with one final version of the modal argument - the scope argument. This is the claim that names and definite descriptions behave differently in modal contexts to the extent that definite descriptions induce de re/de dicto scope ambiguities while proper names do not. Take, for example, the following statements:

(G) Possibly, Socrates did not drink the hemlock

(H) Possibly, the philosopher who drank the hemlock did not drink the hemlock

Now if ‘Socrates’ means ‘the philosopher who drank the hemlock’, then (G) and (H) should be equivalent in meaning and we should expect the name and the description to behave in the same way in both statements. But, so the argument goes, this is not the case, since (H) can be formalised with the possibility operator \diamond and Russell’s Theory of Descriptions, in two ways. These are:

(H*) $\diamond (\exists x (\forall y ((Py \ \& \ DHy) \leftrightarrow x=y) \ \& \ \neg Dhx))$

(H**) $(\exists x (\forall y ((Py \ \& \ DHy) \leftrightarrow x=y) \ \& \ \diamond \neg Dhx))$ [McCulloch, 1989]

(H*) is the de dicto reading of (H) and is obviously false. (H**) is the de re reading and is true. The point of this is to show that there are scope choices to be made when definite descriptions are modalized in this way, which do not arise for proper names. This difference in behaviour between the two leads to the conclusion that the name in (G) cannot be synonymous with the description in (H) because (G) and (H) are not synonymous, since one gives rise to scope ambiguity while the other does not.

The first point to make here is that, once again, we seem to be faced with a question-begging argument, because *if* names are synonymous with definite descriptions then they surely would not behave differently in modal contexts. The 'Scope' argument seems to be assuming that names are not synonymous with descriptions in order to show that they behave differently in modal contexts - and this will not do.

Probably the most common way of arguing against the 'Scope' argument is to point out that proper names are normally taken to have wide scope with respect to modal operators (E.g. Loar, 1976, p373); or, at least, we could adopt a convention to this effect (E.g. McCulloch, 1989, p110,111; and Dummett, 1973, p115). Such a scope convention would explain why (G) does not induce scope ambiguity while (H) does. Of course we could question the legitimacy of such a convention which may seem like a merely ad hoc manoeuvre designed to avoid the conclusion that names and descriptions have different semantic properties. In any case, if descriptions are not also subject to the same convention, then they would behave differently in modal contexts.

Actually the convention may not be necessary at all, if we consider that one of the 'choices' (H*) is actually logically absurd. We could reject this as a genuine way of reading (H). If someone wants to object that (H*) is a perfectly legitimate *formal* reading of (H), then we could dispense with formalization and express the alleged scope ambiguity in natural language. We can certainly do this if the formalization depends on our accepting Russell's Theory of Descriptions. Russell's theory can easily be dispensed with. (See Strawson, 1950; Donnellan, 1966). If we take a non-formal view of (H) along the lines of

(h1) The philosopher who did such-and-such might not have done

(h2) It might have been the case that the philosopher who did such-and-such did not do such-and-such

(h1) is clearly true, while (h2) can easily be regarded as false. It need not be regarded as false, but if it is then it is not only false but absurd, so how can it be one of the choices which renders (H) ambiguous? At any rate, I think it false that (H) is any more ambiguous than (G). If it makes sense to say that Socrates might not have drunk the hemlock, it surely makes just as much sense to say that the philosopher who drank the hemlock might not have done so. Marcus is just wrong when she claims that we need a way of sorting a thing from its properties in order to make sense of the possibility of its not having this or that property. It makes perfect sense to say that The F might not have been The F. It is also not true that we need the notion of rigidity in order to reason counterfactually. If we want to say that Aristotle might not have been the teacher of Alexander all we have to say is that if anyone was the teacher of Alexander, or if anyone has ever been anyone's teacher, then such a person might not have been a teacher. We make sense of possibility here without the use of rigid designators - or even proper names come to that.

If this is not accepted then we are back with the problem of scope conventions for proper names. Dummett not only accepts that names can be regarded in this way, but he also makes the claim (Op.Cit. p115) that Frege could be interpreted as holding the view that definite descriptions can also be regarded as taking wide scope with respect to modal operators (although Frege, of course, was not concerned with modality). Dummett further argues that some proper names at least can induce this same de re/de dicto ambiguity (Ibid.). Such a name could be 'St. Anne', for example. Linsky adds the name 'Homer' as a possible candidate for the scope phenomenon. On page 56 of his (1977) Linsky deals with the 'Homer' example by claiming that, since all we know about Homer is that, if he existed, he is the author of The Iliad and The Odyssey, we cannot make sense of the possibility that he might have existed and not been an author. According to Linsky, this would make the statement

(I) Necessarily, Homer was an author

interpreted *de dicto*, compatible with

(J) Possibly, Homer was not an author

interpreted *de re*. If this is accepted then 'Homer' cannot be a rigid designator because, to say that Homer might not have been an author is to say both that there is a possible world where Homer is not an author ((J) taken *de re*), and also where 'Homer is an author' is true ((I) taken *de dicto*). This presents serious problems for Kripke if it can be maintained, since his claim to rigidity is just equivalent, in Linsky's view, to the claim that proper names cannot induce the *de dicto/de re* ambiguity.

Linsky wants to defend the rigidity thesis against Dummett's claim that some proper names induce scope ambiguities by first of all claiming that, formally, there is no reason why definite descriptions cannot be regarded as rigid as well as non-rigid. If this is so, then there will be interpretations of

(K) Necessarily, Homer = Homer

and

(L) Homer \neq Homer

such that (K) is false and (L) true. That is:

(K*) Necessarily, The author of the Iliad (taken rigidly) = The author of The Iliad (taken non-rigidly)

and

(L*) The author of The Iliad (taken rigidly) \neq The author of The Iliad (taken non-rigidly)

(K*) is one interpretation of (K), and is false, making (K) false on one interpretation; while (L*) is an interpretation of (L) which is true, making (L) true on one interpretation. This reduces to absurdity the claim that 'Homer' is equivalent to a definite description.

So Linsky claims to have saved the rigidity thesis from Dummett's charge that some proper names induce scope ambiguities, but only if we accept that we can rigidify descriptions in the way Linsky claims. Actually, Linsky's argument also relies on this special example of the statement 'Homer = Homer'. If this statement says that Homer is self-identical then it cannot induce any kind of ambiguity in modal contexts because it will be unambiguously necessarily true that Homer is self-identical whether we take 'Homer' to abbreviate a description or not. If it is a necessary truth that Homer is self-identical, then it is also a necessary truth that the author of the Iliad is self-identical, whether this description is taken referentially, attributively, rigidly or non-rigidly, or whatever. Linsky's argument seems to rely on distinguishing between the statements 'Homer = Homer' and 'Homer is self-identical', and there is surely no such distinction. Of course Kripke has no need of such manoeuvres since he would simply reject the claim that there is any sense whatsoever in which it is necessary that Homer is an author.

The argument in terms of scope distinction in modal contexts does not impress Kripke's critics. It seems to reduce to an illicit shift of scope between names and descriptions. That is, names are always given large scope with respect to modal operators, while descriptions are given a small scope reading. If descriptions were given a large scope reading as in "The F might not have been The F", they would behave in the same way as proper names.

Now whether Kripke recognises this or not, he seeks to avoid any criticism about scope distinctions by simply denying that his position can be reduced to a scope argument. This he does in the preface of the 1980 edition of *Naming and Necessity*, where he claims that the scope argument is a misrepresentation of his argument (1980, p11). He seeks to illustrate this with the following examples:

- (i) Aristotle was fond of dogs

(ii) The greatest philosopher of antiquity was fond of dogs.

(i) and (ii) are, what Kripke calls 'simple' sentences - i.e. they do not involve modal operators - so there is no room for scope distinctions. The doctrine of rigidity is not one about scope conventions in modal contexts. Rather it is a doctrine about "the truth conditions with respect to counterfactual situations of (the propositions expressed by) *all* sentences, including simple sentences" (Ibid. p12). So once again Kripke appeals to the type of access we have to possible worlds. No-one other than Aristotle himself will do if we are to consider the possibility that Aristotle was fond of dogs. Whereas, if we consider a possible situation where the last great philosopher of antiquity was fond of dogs, we could be talking about someone other than Aristotle. 'Aristotle' is therefore rigid, while 'the last great philosopher of antiquity' is non-rigid. So even if the modal argument in terms of scope distinctions does not work, Kripke rejects it anyway in favour of his intuitions about how names and descriptions give us access to objects in possible worlds; or how names and descriptions are to be evaluated with respect to possible worlds. Against this, however, we could simply object that this is exactly what the scope argument is about - and so it stands.

There are a couple of points worthy of note here which will take us from the modal argument to the semantic argument. First of all, it might be suggested that, if a definite description such as 'The teacher of Alexander' were to be used referentially, in Donnellan's sense, it would be just as rigid as the name 'Aristotle'. Kripke would reject this for two reasons. First of all, a name cannot be synonymous with a referential use of a definite description because, if we ever discovered that, say, Aristotle did not teach Alexander, we would withdraw the description but we would not withdraw the name (1980, p87n). Secondly, Kripke rejects Donnellan's distinction altogether. At least he does not accept that it is a genuine *semantic* distinction. (See 1980, p87n; and Kripke, 1977). If it is any kind of distinction, then it is a purely pragmatic one, in Kripke's view. (He further argues, in his 1977, that proper names can exhibit a similar distinction to the referential/attributive distinction for descriptions). This raises the question of whether Kripke's thesis is

really a semantic one, as he claims, or whether he is trading on pragmatic features of language use which are irrelevant to the semantics of proper names.

The second point we could take issue with is Kripke's qualification of his 'Aristotle' example (1980, p8,9). Kripke wants to be clear that, in this example, he is assuming a certain reading of (i) in order to recognise that there are, or could be, other people called 'Aristotle'. This is a relatively minor qualification for Kripke, not least because, according to Kripke, the classical description theorist makes the same assumption (Ibid.)

The two points just raised are taken up by Bach (1987) where he argues both that Kripke is confusing semantics with pragmatics, and that this 'minor qualification' is far from minor. In fact, according to Bach, Kripke's 'assumed' reading of a statement like (i) is a major qualification which actually exposes the rigidity thesis as illusory (Op.Cit. p167 -171). In Bach's view there is nothing in the semantics of a name like 'Aristotle' which tells us who it refers to - whether it be the philosopher or someone else with the same name. Kripke has, therefore, 'built in' to this example the fact that 'Aristotle' is being used to refer to the philosopher rather than, say, the shipping magnate. This relativizing of the rigidity thesis to uses of a name serves only to trivialise it. On p169 Bach contends that "...the name 'Aristotle' seems rigid only because its use to refer to a certain individual is fixed before the question of its rigidity is raised...(This is what Kripke often calls (1980, e.g. p57) 'using the description to fix the referent')". Furthermore, the rigidity thesis is not even a semantic theory. It is merely a pragmatic theory about the use of proper names to refer.

Bach seems to have a point here. There does seem to be some confusion in Kripke as to whether his thesis is semantic or pragmatic. Although he himself is quite clear that he is dealing with the semantics of ordinary proper names, at least in his negative thesis, his positive thesis, the so-called 'Causal Theory' is more of a pragmatic theory of how proper names acquire their reference and how they are subsequently used to refer. Bach also seems to be on to something when he argues that the use of a name to refer is fixed before the question of its rigidity is raised. This is borne out by something Kripke says on p14 of his (1980). When considering the possibility of introducing a name into a language as a rigid designator he writes:

“Let ‘a’ (rigidly) designate the unique object that actually has property F, when talking about any situation, actual or counterfactual.” The key phrase here is ‘actually has’ because if we take the name ‘Aristotle’, for example, and introduce it into a language as a rigid designator of Aristotle, using the description ‘The teacher of Alexander’ to fix its reference as the object which actually has this property, then we must *already have* access to Aristotle before the name can be introduced in this way; otherwise we would not know who it is that ‘actually has’ this property. When Kripke talks about using a definite description in this way to fix a reference he is not using it attributively (nor referentially). He does not want the name ‘Aristotle’ to designate just whoever happens to have this property, but Aristotle himself. But this is where the use of the name is fixed as a name of a certain person before the question of its rigidity is raised.

Even if this foregoing is accepted, however, Kripke could reply that it only applies to cases where we do in fact introduce names into a language by means of a reference-fixing description. He later argues that there is no need to do this (1980, p59). (How reference is actually fixed is part of Kripke’s positive thesis). However reference is fixed, though, Kripke does seem to have a problem with the apparently innocuous fact that many names have multiple bearers. Kripke’s assumed reading of (i) does seem to have built into it the fact that we are dealing with one ‘Aristotle’ rather than another. Seen in this way, Bach’s diagnosis of the illusion of rigidity seems to be correct. (We shall see, however, that Bach himself has difficulties with the fact that names can have more than one bearer).

It would seem that the only theory that can adequately deal with the problem of names with multiple bearers is the orthodox description theory whereby names have a descriptive sense. Bach would reject such a claim, however. Although he advocates a description theory, it is no orthodox theory. Along with Jerrold Katz (in Boolos, Ed. 1990, and Katz, 1994) and William Kneale (in Nagel, Suppes and Tarski, 1962) Bach is an advocate of what we could call the ‘Meta-linguistic Description Theory’ (MDT) whereby a name ‘N’ has the sense of “The bearer of ‘N’”; or, as Katz puts it, “The contextually definite thing that is a bearer of ‘N’” (1994, p14). According to Bach et.al., if the sense of a proper name is given in this

way, not only can we avoid Kripke's arguments against the orthodox theory, we can also solve the problems of reference. Such a bold claim is worthy of investigation.

Kripke in fact deals with one version of MDT in his (1980) - that of Kneale (p 68-70). According to Kripke, Kneale's theory is circular, as well as being open to the modal objection, since Kneale uses the predicate 'call' rather than 'bear'. If we take, for example, the name 'Socrates', then Kneale would give it the sense of

“The man called ‘Socrates’”

Now obviously Socrates might not have been called 'Socrates', so the modal objection would apply. Furthermore, Kripke accuses Kneale of offering no theory of reference at all. This is the circularity objection which goes as follows.

“Someone uses the name 'Socrates'. How are we supposed to know to whom he refers? By giving the description which gives the sense of it... 'the man called 'Socrates'.... We ask, to whom does he refer by 'Socrates'? And then the answer is given, 'Well he refers to the man to whom he refers'. If this were all there was to the meaning of a proper name, then no reference would get off the ground at all” (1980, p70)

Bach's (and Katz') reply is simply to deny that MDT is a theory of reference at all. It is a theory of the meaning (sense) of proper names and not a theory about their use to refer (Bach, 1987, p160; Katz, in Boolos, 1990, p40). As Katz puts it in his (1994, p9): although sense is necessary for reference, since referential uses begin with knowledge of sense, sense does not determine reference. The problems of reference are thus easily solved. For example, the problem of a negative existential claim such as “Pegasus does not exist”, becomes “The bearer of 'Pegasus' does not exist”. Similarly, we can solve Frege's identity puzzle by pointing out that “Hesperus is Hesperus” expresses the analytic and trivial proposition that the thing which is a bearer of 'Hesperus' is the thing which is a bearer of 'Hesperus'; while “Hesperus is Phosphorus” expresses the informative and synthetic proposition that the thing which is a bearer of 'Hesperus' is the thing which is a bearer of 'Phosphorus'. As for Kripke's modal objection to Kneale, both Bach and Katz, while agreeing that Socrates might not have been called 'Socrates', contend that things are different if

we use the predicate 'bears' rather than 'calls'. This choice of the predicate 'bears' rather than 'called', or 'is known as' etc. is crucial and fundamental to the success of MDT.

According to Katz (Op.Cit. p39) "Only a theory formulated in terms of the bare name relation does not buy into the description theory". To use a predicate like 'call', is, according to Katz, to make a further property, that of being called by the name in question by members of a certain community, part of the condition for something being the bearer of a name (Ibid.). If Bach and Katz are correct about this distinction between the predicates 'bear' and 'call' then there would seem to be at least one version of MDT which would survive Kripke's criticisms of Kneale. It is far from clear, however, that MDT has made this distinction. Katz gives the example of orthodox Jews who are not allowed to pronounce the name of God. Nevertheless, God is the bearer of this name. He is the bearer of the name even though no-one actually calls Him by the name. This will surely not stand up. If God is *never* known by the name (presumably, 'Jehovah'), he surely does not *bear* the name. Or if He does bear the name, it is only because there are people in the world who *actually use* the name. These people keep the name alive, as it were. We could certainly make a case for saying that an object does not bear a name unless the name is actually used from time to time. The fact is that MDT simply does not say what 'the bare name relation' consists in. It relies on this crucial distinction between the predicate 'bear' and other predicates such as 'called' or 'is known as', but it does not make clear what the difference is between them. This is a major flaw in the theory. (Even if the 'Jehovah' example did work it would not show that we can make the relevant distinction, *in general*. Bearing a name would seem to be every bit as contingent a property of objects as being called by a name).

One obvious criticism of MDT is that it cannot deal with names with multiple bearers; a charge made against DRT by Bach and Katz, as we have seen. After all, if the name 'John Smith' has many bearers, how are we to know who is being referred to if the only sense the name can supply is "The bearer of 'John Smith'"? Even if the sense is "The contextually definite thing which is a bearer of 'John Smith'" it is difficult to see how this context does not give a sense to the name in terms of definite descriptions, since the context is *indispensable* to the actual functioning of

the names in these sorts of cases. It is a curious feature of MDT that it wants to make a rigid distinction between semantics and pragmatics - something that could be seriously questioned - especially if we regard meaning as use.

Bach says something very strange about names with multiple bearers in belief contexts, in (1987, p166). He says that when 'N' has many bearers, the description "the bearer of 'N'" is *incomplete*, but can be completed in the following way:

S believes that the bearer of 'N' WHO/WHICH is F is G

In such cases MDT seems to require something more than a meta-linguistic sense when dealing with names with multiple bearers, which is as we should expect. Bach really gives the game away here when he admits that the meta-linguistic sense is incomplete in these cases and must be augmented in some way. How can he be giving a theory of sense for proper names if the senses are very often incomplete? Furthermore, on p160 of his (1987), Bach seems to again appeal to information other than the meta-linguistic information that 'N' has the sense of "the bearer of 'N'". In dealing with Kripke's circularity objection he considers the example of someone uttering 'Socrates drank hemlock'. The circularity objection would contend that the speaker is stating nothing more than that the man he is referring to drank hemlock. (Actually he is not even saying that, since Socrates could be a woman for all we know). Bach replies that MDT allows that the speaker is doing more than that. On p160 -161 he says:

"After all, there is information on which the speaker expects the hearer to rely in order to identify the individual being referred to. If the sentence is semantically equivalent to 'The bearer of "Socrates" drank hemlock'...the hearer can infer the name of the intended referent (It is mutually believed that the relevant bearer of this name is an ancient Greek philosopher). Thus the speaker intends, and can reasonably expect the hearer to infer that he intends, to be stating that the ancient Greek philosopher bearing the name 'Socrates' drank hemlock. Of course not all the relevant information is built into the meaning of the name, but MDT does not require that it should be."

Now it seems to me that it is simply not the case, or not necessarily the case, that the relevant bearer of the name 'Socrates' is a Greek philosopher, rather than a Brazilian

footballer. If the intended bearer of the name is determined by the context of the discussion, then it is this context, which would be given in terms of descriptions, which would give the sense of the name and not simply “the bearer of ‘Socrates’”. Bach and Katz would deny this of course. They would contend that any contextual information which determines reference is purely a pragmatic feature of particular uses of tokens of names. MDT is a semantic theory only, and does not give any explanation of uses to refer.

Be that as it may, it seems that MDT is useless in dealing with names with multiple bearers, as well as the circularity objection, without these allegedly pragmatic features. MDT is really irrelevant as a theory of the sense of names if it wants to divorce itself altogether from the determining of reference. When Katz (1994, p10) claims that a meta-linguistic sense is necessary (though not sufficient) for reference, it could be that it is only necessary in the sense that it is simply taken for granted.

MDT is no rival to the orthodox Fregean theory, or the cluster theory, because of its refusal to deal with reference. It attempts to win the game by not really taking part, or by creating a game within a game. It claims to be able to solve the problems of reference, but any theory that deals only in sense can only solve the problems of reference by denying that they are problems of reference at all. This would not seem to be the case. MDT cannot make the crucial distinction between bearing a name and being actually referred to by the name. Neither can it deal with the problem of names with multiple bearers without smuggling in something like a non meta-linguistic sense without which the name could not function.

As well as the foregoing there is one important fault with MDT, and that is that it does not give the sense of a name in terms of what speakers *understand* by the name. Now there may be problems with what speakers understand by a name; the main one being that different speakers may understand different things by the same name. Nevertheless, it seems perfectly reasonable to suggest that the sense of a proper name is just what speakers understand by it; or what they ‘grasp’ when they understand the name - or a sentence which contains the name. Seen in this light, MDT is hopelessly inadequate as a theory of sense because no-one would ever understand a name like ‘Socrates’ as “The bearer of ‘Socrates’”. That is just not the

sense of that name for speakers of the language. Something more is needed otherwise the name could not be used by anyone. This is certainly true in the general case.

The Semantic Argument

We come now to the last of Kripke's main arguments against the Description Theory - the Semantic Argument; also known as 'The Argument from Ignorance and Error' (Bach, 1987, Op. Cit., p157). The Semantic Argument is regarded by many as the most persuasive of the arguments proffered by DRT, and is by no means to be attributed solely to Kripke. One of the best known examples of the argument comes from Donnellan (1972). It concerns the sense of the name 'Thales', which is usually taken to be something like 'The Greek philosopher who held that all is water'. Now, the argument goes, suppose there never was a Greek philosopher who held this view. Alternatively, suppose our use of the name 'Thales' stems from a well-digger called 'Thales' who was taken out of context and never actually held such a view. Suppose further that there was a philosopher living at the same time as the well-digger who did hold this view, although he had no connection whatsoever with the well-digger. The crucial question for the description theory is "To whom would we be referring when we use the name 'Thales'?" In the former case where no-one held the view that all is water, the Fregean seems to be committed to saying that the name 'Thales' does not refer to anyone; while in the latter case the Fregean seems committed to the view that 'Thales' refers to the philosopher with whom we have no connection, and not to the well-digger from whom our use of the name derives. The case against the description theory is compounded by Kripke's famous 'Godel/Schmidt' example from *Naming and Necessity*. Kripke asks us to imagine that the Incompleteness Theorem was in fact discovered by the unknown Schmidt, and not Godel, who simply took the credit for the discovery. Now if the Fregean is correct in saying that the name 'Godel' refers to whoever satisfies the description which gives the name its sense, then, in this case, 'Godel' should refer to Schmidt. When we say that Godel proved Incompleteness, we are not saying something false of Godel, but something true of Schmidt. Now this is just not the case, according to

Kripke et. al. When we use a name like 'Thales' or 'Godel' we refer directly to the person with whom our use of the name has some connection. The name 'Godel' always refers to Godel, never to Schmidt, or to anyone who happens to fit a certain description (Kripke, 1980, p84: also Salmon, 1981, p29,30). The same goes for the name 'Thales'. It refers to the well-digger, if that is the source of our use of the name, and not to whoever happens to fit the description. This represents a *Reductio ad Absurdum* of the Description Theory.

Now DRT seems to have a point here. A name like 'Godel' surely does not refer to whoever happens to have done a certain thing. If it is to count as a genuine name it must refer to Godel - that very person. And this must be the case for any theory of naming - any theory which is a theory of the reference of proper names and not just their sense. So what, then, must the Fregean concede to DRT at this point? *It seems that he must concede that the sense of a proper name is not given by a definite description used attributively.* The obvious question now is "Can a proper name have a sense given by a *referential* use of a definite description, which determines reference regardless of actual fit?" Kripke does consider such a possibility (1980, p87n;) but, as we have already seen, he rejects it on the grounds that, if we discovered that the description did not fit the object, we would withdraw the description, but not the name (1980, Ibid.). The name is therefore not synonymous with the description.

Of course Kripke is quite correct when he claims that we would withdraw this false description. What he fails to notice, however, is that, as soon as we have done this, we would *immediately replace it with another description* - something like 'The man who *pretended* to prove Incompleteness'; or 'The man who *claimed* to have etc...'; or, 'The man who was *formerly believed* to have etc...' It could be that the 'Godel' argument merely shows is that we could be mistaken about the meaning of a proper name; not that names do not have *any* meaning. Once we discover our mistake we have the true sense of the name . So Kripke cannot shake off descriptions so easily. (We shall see later that he cannot shake them off at all. Names simply cannot function without descriptions). He may have shown that names are not synonymous with attributively used descriptions, but he has not succeeded in refuting the referential use. Of course, if we switch from the attributive

to the referential sense of definite descriptions, the sense of proper names will be very different from that posited by Frege.

Not all commentators, however, accept that the attributive use should be abandoned. As we have already seen, Linsky (1977) has the opposite intuition from Kripke; at least regarding names whose sense can only be given by a single description - e.g. 'St. Anne', or 'Homer' (Op. Cit. p56). Linsky makes the point that if Homer exists, then he can only be the author of *The Odyssey* and *The Iliad*. This description exhausts our information about Homer, and we can make no sense of the claim that Homer might have existed and not been an author. This of course, is just the sort of claim that DRT would reject. Kripke would claim that it makes perfect sense *de re* to say that Homer might have existed and not been an author; although we could question this claim on the grounds that one of things at stake in this debate is whether it makes sense to talk or think *de re* at all. DRT simply takes it for granted.

Clearly we have a clash of intuitions here, a clash which takes us to the heart of the question of whether names refer directly or indirectly. If Kripke's intuition seems stronger than Linsky's we can strengthen Linsky's case by reformulating his 'Homer' example in the same way as the 'Thales' and 'Godel' examples. That is, we can ask whether we could be mistaken in thinking that Homer was an author of certain works, or whether we could actually discover that Homer did not do these things. Now it seems to me that the only way in which we could discover that Homer did not do such and such would be if we had access to Homer which was independent of the description 'The man who did such and such', and this is something we do not have in this case; or indeed in many other cases of names which have the sense of a single description. (We could of course discover that the author of the *Iliad* was never called 'Homer' by his contemporaries, but that would be irrelevant). So it seems that some names, at least, are immune to the Semantic argument, and could be said to have the sense of a definite description used attributively.

Of course there are examples of single description names where we could easily discover that the description did not fit the object it is supposed to fit. One obvious example is the name 'Godel', but this is because we know who Godel is

independently of the description 'The man who proved Incompleteness', or indeed of any one description. The situation is different with a name like 'Thales' because if no-one held the view that all is water, then we could not discover that this description did not fit *Thales*. If this description is the only thing we know about Thales, then to discover that it fits no-one would lead us to conclude that Thales did not exist.

This contrasts strongly with an argument put forward by Grayling (1982), where he says the following:

Now suppose that Thales was a well-digger and not a philosopher...Does the fact that he fails to fit the description 'The philosopher who held that the arche is water' mean that we are not talking about Thales? On the contrary. It only makes sense to say that a description fits or fails to fit ...something if we can refer independently of descriptions; otherwise we should be obliged to say that if Thales does not fit the descriptions of him, he would not have existed. (p182)

For Grayling, this way of arguing is a consequence of Donnellan's work on the referential/attributive distinction and is a neat summary of the Semantic Argument. It also serves to highlight everything that is wrong with the Semantic Argument.

To begin with, it is just not true that we must be able to refer independently of *any* descriptions in order to determine whether a particular description, say, 'The F', fits an object. We only need to refer independently of 'The F'. We could refer using 'The G', or 'The H'. Secondly, suppose the name 'Moses', for example, is taken to abbreviate the description 'The leader of the Exodus', then we could discover that there never was an Exodus, and hence discover that Moses was not the leader of the Exodus, without referring to Moses in any way. This would be the type of situation where the Description theory would conclude that Moses did not exist.

The third point involves a denial of this existential conclusion by DRT. Grayling is quite clear (and so is Kripke and Donnellan) that we should not conclude that Thales did not exist just because he does not fit a certain description. Thales after all could have been a well-digger about whom a false tradition arose. Now this may seem obviously true, but actually it is far from obvious. After all, who are we talking about when we say that Thales could have been this or that? Kripke would say that it

is the person we are connected to by a chain of communication. This is all very well, but how do we know that we *are* connected to anyone? If none of our descriptions fit anyone then it would become doubtful whether we were connected to anyone by a certain name. Thales may have been a well-digger or he may not have existed at all. We simply do not know. The way in which Grayling constructs his example simply begs the question against the Description theory because it is built into the example that the name Thales refers to a certain person. *If* we suppose that *Thales* was a well-digger then of course we are referring to Thales when we use the name (even though he fails to fit the description in question) but the point is that it is far from clear that we can do this. Similarly, when Grayling says that we must be able to refer independently of descriptions otherwise we would have to conclude that Thales did not exist because he fails to fit a certain description, he is again begging the question by more or less stipulating that we are referring to Thales, when we may not be referring to anyone. Of course if we discover that *Thales* did not do such and such then the name refers to Thales. But this is not what we do in these cases where we only have one description. The only way we can discover that *Thales* is not 'The F' is if we can refer using another description, such as 'The G', which would then give the sense of 'Thales'.

Kripke is guilty of the same question-begging confusion in his 'Moses' example (1980, p58-59) where he is dealing with a whole set of descriptions rather than a single one. Kripke wants to argue that, not only can we reasonably suppose that there is some possible world where Moses does not do *any* of the things commonly attributed to him, but also that we could be mistaken about these things, and could discover that we are mistaken. On p58 he is again concerned with the distinction between descriptions which give a synonym and those which merely fix a reference:

"If 'Moses' *means* 'the man who did such and such, then if no one did such and such, Moses didn't exist;...But if the description is used to fix a reference rigidly, then...that is not what is meant by 'Moses didn't exist', because we can ask... of a counterfactual case where no one did indeed do such and such...does it follow in such a situation Moses wouldn't have existed"

Kripke answers his question in the negative because "...surely Moses might have just decided to spend his days more pleasantly in the Egyptian courts"(Ibid.) Now again this may seem obviously true, and so the name 'Moses' cannot be synonymous with *any* descriptions, but, again, it is not so obvious upon closer examination. If Kripke is correct when he says that Moses might have spent his days in the Egyptian courts and not done any of the things commonly attributed to him, then at least *part* of the Moses story must be true - the part which tells of his upbringing in the Egyptian court. So we have at least one non-trivial description which could give the sense of 'Moses'. Of course we could be mistaken in believing that Moses (i.e. the man who was brought up as a prince in the Egyptian court) did most of the things commonly attributed to him, but in order to be able to make sense of this possibility it seems that we require at least one description which gives us 'access' to Moses. The key question now is 'Could we discover that even this description is false of Moses?'. For the Description theorist the answer must be a clear 'No!' because we cannot make sense of the fact that Moses existed and did none of the things commonly attributed to him. DRT would of course answer affirmatively.

Once again we have the clash of intuitions, and once again DRT seems to have the obviously true one. However, it is actually the Description Theory that is nearer the truth here; although it is not quite correct if it holds that Moses definitely did not exist if it turns out that none of these descriptions fit anyone. What the Description Theory should say is that, if we could be mistaken about all of these descriptions, or if we actually discovered that no-one fits any of them, then the question of the existence or non-existence of Moses becomes *meaningless*. The onus is then on DRT to explain how the question could be meaningful.

Despite the apparent intuitive correctness of the Semantic Argument against the Description Theory it is far from convincing in the type of examples just considered. In fact it is confused over the existence question. This confusion can be traced to Kripke's original example of the name 'Godel', where he argues that we obviously would still use the name 'Godel' as a name of Godel even if the description we associate with the name fit someone else, or no one. Now this is undoubtedly true, but this is because with Godel we are in a completely different situation from the

'Moses' or 'Thales' examples. The difference is that we know that Godel existed independently of the sort of descriptions normally taken to be the sense of the name. We can simply point to Godel, or to a photograph of him, and say "That is Godel. That is the man we believe to have done such and such, but who might not have done, and about whom we could yet discover that we were mistaken." The same goes for Kripke's other example of Nixon. We have a completely different access to these people than we do with Thales, or Homer, or Moses, or a whole host of others. So of course if we discover that Godel did not do such and such, the name 'Godel' still refers to Godel, and not to whoever fits the description. (Godel is *that man*, or the man in the photograph). And of course we would not conclude that Godel did not exist if it turned out that no-one did such and such because we have an independent access to Godel. (Kripke probably knew him personally!) Kripke seems to take this one example and generalise to **all** names - which is premature, to say the least. It may even be downright reckless.

So we have names of objects whose existence is not in doubt (Godel, Nixon etc.); names of objects whose existence is doubtful (Thales, Homer etc.) and 'names' which purport to name objects which definitely do not exist (Pegasus, Zeus etc.) These latter are easily dealt with, as already mentioned. They are simply not genuine names since they do not name anything. Names which purport to name objects of doubtful existence are synonymous with definite descriptions used attributively. At least, they can survive the Semantic Argument. It makes no sense to say that we could be mistaken in believing that Thales did not do such and such since we have no idea who Thales is independently of a certain description. If we want to argue that we could be mistaken in thinking that the man Aristotle was referring to with the name 'Thales' did not do such and such, then our access to Thales is still via a description (even if Aristotle had a different access to Thales). However, if we do conclude that in these cases names are synonymous with a definite description, it seems that we cannot conclude that this description determines reference if it is doubtful whether there is a reference. In the case of a name like 'Thales' we cannot talk about referring, either directly or indirectly. In the case of names like 'Nixon' which refer to objects whose existence is not in doubt, we do seem to have a direct access to these objects. In answer to the question 'Who is Nixon?' we need not give

a description. We could simply point to Nixon, or to a photograph of him, and say 'That is Nixon'. But note, it is not the *name* 'Nixon' that gives us this direct access. Rather, it is the *demonstrative* 'that' which does all the work. Instead of saying 'Nixon did such and such' we could say 'That man did such and such'. Now demonstratives are undoubtedly rigid because they are direct designators, but their reference is also context-dependent, and so they would not be of any real use to Kripke in trying to establish the existence of necessary a posteriori truths.

The Semantic Argument has not succeeded in showing that proper names are not synonymous with definite descriptions. Kripke has not succeeded in refuting the referential use of descriptions; or even the attributive use, if the description is formulated in the correct way. 'Godel' is not 'the man who...', but 'the man *commonly believed to...*' It is impossible for *the linguistic community as a whole* to be mistaken about a sense like that (although an individual speaker could be). Thus, Kripke's 'Godel' argument has been defeated; although we are forced to qualify our senses in the way just mentioned. If we actually found out that Godel did not do this thing then the sense would change to 'the man formerly believed to have...'; or 'the man who pretended to...' DRT does not allow for this change in meaning; but it should be remembered that the meaning of proper names is discovered empirically, and so there is always the possibility of mistakes. This possibility does not count against the Description Theory, especially if we are careful to formulate the description in the right way. And of course there is no possibility of being mistaken if the name has the sense of only one description.

The bottom line for the Semantic Argument is that a statement like

(*) Godel proved Incompleteness

cannot be analytic because we can always be mistaken about this 'fact'. Now of course, if this turns out to be false, then 'Godel' cannot be synonymous with the description 'the man who...' But why can we not have a qualified notion of analyticity, just as Kripke has a qualified notion of necessity? I.e. (*) is analytic, **if true**, just as Kripke claims that "Water is H₂O" is necessary, if true?

The main criticism of the Semantic Argument, indeed the whole basic claim that names are not synonymous with descriptions, is that DRT, and Kripke in particular, does not say what is to be understood by 'meaning'. If, for example, meaning was simply to be identified with *use*; and we could not use a name without associating it in some way with a description (which seems to be the case), then that name would be synonymous with the description in question - in some sense of 'synonymous'. Until DRT can tell us what it means by 'meaning' it is on very shaky ground when denying that names have a meaning given by a definite description.

We end the examination of Kripke's negative thesis with a brief look at his minor arguments against the Description Theory. These are:

- 1) In many cases there is *no single description* which gives the sense of the name (1980, p30)
- 2) Sometimes the speaker cannot give *any* description which gives the sense of the name (Ibid.)
- 3) The Inextricability Objection (1980, p81)

The first objection arises from a passage in Frege (1892) where he highlights a looseness in our language whereby a name can have one sense for one group of speakers and a different sense for another group. Frege gives the example of the name 'Aristotle', which could have the sense of 'The teacher of Alexander' for one group and the sense of 'Plato's most famous pupil' for another. Now from a purely semantic point of view, this difference is not important, so long as the reference is the same. The only real problem would be if the two groups were to meet in order to discuss Aristotle. When one group voices the opinion that Aristotle was a great teacher, this may cause confusion among the other group of speakers who know nothing of Aristotle's teaching experience. This breakdown in communication is easily resolved, however.

Objection 2) is simply not true. Anyone who can actually use a name to say something meaningful must be able to give *some* kind of sense to the name, in

terms of descriptions. He must be able to answer the question “Who is ‘N’?”, in a non-trivial way. I.e. he must not simply say “N is the person who bears the name ‘N’”; or “N is the person about whom so-and-so was talking”. In Kripke’s ‘Feynman’ example (1980,p91) the person who simply hears the name ‘Feynman’ in the market place is in no position to use the name. He is not entitled, for example, say something like “Feynman is the man about whom the people in the market were talking”, because he does not know who Feynman is - and of course there could be more than one Feynman.

Gareth Evans tries to improve on the ‘Feynman’ example in his (1973) where he offers the example of a speaker who overhears a conversation in a pub about a man named ‘Louis’. Evans thinks it obvious that the speaker could simply join in the conversation and denote Louis without having a sense for the name ‘Louis’, saying things like “What did Louis do then?” or “Louis was quite right to do that” (1973, in Moore Ed.(1993) p212,213); but this is just a mistake, because the name ‘Louis’ would take on the sense of “The man who did that” (whatever ‘that’ may be) for the speaker. If the speaker has no sense at all for the name ‘Louis’ he will *mention it rather than use it*, after the manner of “This ‘Louis’ was quite right to do that”. Even if we grant Evans his claim here (and not forgetting his insistence on contextual factors in determining reference) could the conversation take place if *none* of the participants had a sense for the name ‘Louis’? Surely not, since they would literally not know who they were talking about. Given this, then, if our speaker does denote Louis without having a sense for the name, then his act of denoting would be parasitic on the senses associated with the name ‘Louis’ by the other participants. At least someone involved in the conversation must have a sense for the name.

The ‘Feynman’ and ‘Louis’ examples expose a clear use/mention confusion in DRT, at least concerning proper names. This is a serious mistake and arises from the failure of DRT, especially Kripke, to say what is understood by a name (More on this later).

Finally, the so-called Inextricability Objection (McCulloch, 1989). This is the objection which states that in order to avoid circularity, a name must go on giving

senses for all the names appearing in descriptions which are taken to be the sense of proper names. For example, if we posit 'The teacher of Alexander' as the sense of 'Aristotle', we still have a name as part of the sense; and so we can ask for the sense of 'Alexander'. Now we cannot give this as 'The pupil of Aristotle' without circularity, so we would have to give some other sense - for example, 'The son of Phillip of Macedon'. But now we can ask for the sense of 'Phillip of Macedon', and so on. We cannot get rid of proper names from our senses.

The first point to make here in defence of the description theory, is that the same phenomenon arises for Kripke's reference-fixing descriptions. If we fix the reference of 'Aristotle' using the description 'The teacher of Alexander', we can ask how the reference of 'Alexander' is to be fixed, and so on. The solution of the problem is the same in both cases - viz, the descriptions will be grounded in some kind of *ostensive* definition. For example, if the description 'The first man on the moon' is not sufficient to give the sense of the name 'Neil Armstrong' because it contains the name 'moon', we end the regress by simply pointing to the moon. Neil Armstrong is the first man to set foot on *that*. Other examples will no doubt be more complicated, but the principle will be the same. So the Inextricability Objection fails; although once again we are relying on an indexical to effect reference.

SUMMARY OF NEGATIVE THESIS

What have we seen so far, then? We have seen that DRT wants to challenge, at a fundamental level, what we have called 'The Description Theory of Proper Names' whereby a name is synonymous with a definite description, or set thereof, used attributively, which determines the reference of the name. For DRT this view has absurd consequences, especially in connection with what we regard as a priori, necessary, and analytic. According to DRT reference is direct, unmediated by anything like a sense given by a definite description or a set thereof. In the hands of Kripke, proper names not only refer directly, but rigidly, which allows him to conclude that there are necessary a posteriori truths as well as contingent a priori ones.

We have seen, however, that the case for DRT against the description theory is nowhere near as convincing as its various advocates claim. First of all, the Modal Argument, which Kripke regards as a very important argument, trades on an illicit shift of scope. Secondly, the Epistemological Argument, while immune to Linsky's criticisms, does not work in the way Kripke claims. Although we do indeed discover empirically that Aristotle, for example, taught Alexander, in doing so we also discover something about the *meaning* of the name 'Aristotle'- or at least we discover information which can become part of the sense of the name. Once we know the meaning of the name we are in a position to know certain things a priori - for example, that Aristotle was a teacher. Actually we discover the meaning of all names empirically; and so, if we know a priori that bachelors are unmarried, it is only because we first discovered the meaning of 'Bachelor' empirically. Without knowing the meaning of the word 'bachelor', we do not know anything a priori about bachelors. The same could be said for ordinary proper names.

Prima facie, the Semantic argument is more successful than the aforementioned. It seems to show that the sense of a proper name cannot be that of a definite description used attributively, because if we use a name like 'Godel', we surely do not use it as a name for whoever happens to fit a certain description. We use it as a name of Godel himself. However, even this argument does not show what DRT claims. DRT fails to distinguish between names which purport to refer to objects whose existence is doubtful, e.g. 'Thales'; and names of those whose existence is not in doubt. In the 'Thales' example, we, as a linguistic community, could not be mistaken in thinking that Thales is the philosopher who held that all is water unless we have another way of gaining access to Thales (apart from that description). This we do not have. Grayling's example simply begs the question against the description theory by building into the example that we have an independent means of referring to Thales. It makes no sense to say that 'Thales' refers to Thales himself, rather than to whoever satisfies a certain description, because we do not know who Thales is independently of that description.

The situation is different with names of objects we know to exist, such as Godel. As already mentioned, the name 'Godel' does not refer to whoever did such and such but to Godel himself. We can of course be mistaken in thinking that Godel is

the man who fits this description, unless we change the description to ‘The man *commonly believed* etc.’ If the sense of ‘Godel’ is ‘The man commonly believed to have done such and such’, then we, as a linguistic community, cannot be mistaken about this, since it is the linguistic community who commonly believe this thing. The community can of course be mistaken in thinking that Godel is ‘The F’; but it cannot be mistaken in thinking that it *believes* that Godel is ‘The F’. Once we discover the truth, the sense of the name would simply be corrected.

Now Kripke does deal with such a theory of the senses of proper names in *Naming and Necessity* (p88 - 90) and rejects it on the grounds that in giving the sense of a name in terms of what is believed to be true of the object named we would be violating a non-circularity condition of reference. As he says on page 89:

All of us in the community are trying to determine the reference by saying ‘Godel is to be the man to whom the incompleteness of arithmetic is commonly attributed’. None of us will get started with any attribution unless there is some independent criterion for the reference of the name other than ‘the man to whom the incompleteness of arithmetic is commonly attributed’. Otherwise all we will be saying is ‘We attribute this achievement to the man to whom we attribute it’, without saying who that man is, without giving any independent criterion of the reference, and so the determination will be circular.

Now Kripke certainly seems to have a point here and I am well aware of the need to specify exactly what is going on in such a situation, but I will deal with it in chapter 3 when we deal with the general theory of how it is that proper names and natural kind terms actually refer.

So the Semantic Argument does not work in the way that DRT claims. Nevertheless we get the feeling that something about the pure Fregean Description Theory must change. If we change the sense of a name from ‘The F’ to ‘That which is commonly believed to be The F’, then we are relying on other speakers in order to effect reference. (In chapter 3 we shall see that other speakers are crucial to the way in which proper names refer).

In addition to the foregoing Kripke claims that we have an intuition of the rigidity of proper names. His original intuition seems sound enough. It is that, when talking about counterfactual situations where a certain person does not do the things he does in the actual world, we are talking, by definition, about that very person, and not someone who satisfies a certain description. If, for example, we accept that Nixon might not have been a politician, then this only makes sense if we are talking about Nixon himself. This is what Kripke means when he says that the trans-world identity of individuals is given. However, even if we agree with this, it is not obvious that this tells us anything at all about how names refer. In fact this argument from counterfactuals may show the opposite of what Kripke wants. For example, the counterfactual “Nixon might not have done such and such”, if it is a genuine counterfactual, only makes sense if Nixon is *in fact* the man who did such and such. I.e. if it makes sense to talk counterfactually, then we must assume that we are dealing with facts. So, for the counterfactual statement to make any sense, it must be at least *part* of the sense of the name ‘Nixon’ that Nixon is the man who did such and such. Furthermore, someone who did not know the sense of the name ‘Nixon’ would grasp the sense, or part of the sense, from this counterfactual. So the counterfactual argument could be turned against Kripke in order to argue for the opposite position. At any rate, as we have seen, the notion of rigidity is in no way essential to the ability to reason counterfactually. Descriptions with wide scope will do perfectly well.

According to Almog, Kripke’s argument tells us nothing about names because he starts with the designata rather than the designators which could be designated by any name at all. Kripke is begging the question once again. Bach would agree, pointing out that the claim to rigidity is illusory. This is because, in the case of names with many bearers Kripke has to build into his example the fact that he is using the name to refer to one bearer rather than another. This seems correct, although Bach’s meta-linguistic thesis fares no better on this point, or indeed on many other points, and so is no alternative to DRT.

So Kripke, and DRT in general, has not succeeded in eliminating descriptions from the sense of proper names, or in showing that names are not synonymous with

definite descriptions. Neither has he shown that names are rigid designators (although they may well be). He has not even succeeded in making the distinction between descriptions which give a synonym and those that merely fix reference. Actually Kripke may be confused about the very nature of naming if he thinks that we fix the reference of a **name** - however he wants to do it. In truth, we **never** fix the reference of a name. If something is a name for something then *its reference must already be fixed* - otherwise it could not be a name for anything. The reference could be fixed ostensively or using a definite description, but something must fix the reference of a word if it is to function as a name. This distinction between a word and a name is never made in the literature, either by DRT or by Fregeans, but actually it is crucial to the understanding of what a name actually is. What we do when we name something is take a *word* or *sign* and fix its reference in some way. This *makes* the word or sign into a name. It will now have certain semantic properties in virtue of its being a name, but the important point here is that this name would have remained a mere sign if we had not made it into a name by fixing its reference in some way. This means that Kripke can no longer talk about a definite description *merely* fixing the reference of a name, because if the description performs a reference-fixing role, then it is the description that makes the name into a name. Without the description to perform this role, and in the absence of an ostensive definition, the so-called 'name' would remain a mere sign, devoid of any semantic properties whatsoever. Given this, it is very difficult to see how the description is not synonymous with the name since the name would not be a name at all without the description. At the very least, the foregoing blurs the distinction between descriptions that are synonymous with names and those that merely fix a reference. If Kripke cannot maintain this distinction then his claim that there are contingent a priori truths collapses.

It may seem that we can fix the reference of a name in the sense that we can take an already existing name, such as 'Aristotle' for example, and use it as a name for something else - a dog, say. We could fix the reference of 'Aristotle' either ostensively or using a description such as 'our new dog' or 'the dog we bought this morning'. Either way 'Aristotle' is not a name *for the dog* until we do this. As far as the dog is concerned 'Aristotle' is not a name until *we* make it into one. The fact that

it is already used as a name for something else is irrelevant - and of course we could have used any arbitrarily chosen sign as a name for our dog.

Kripke definitely states at various points in (1980) that we can fix the reference of a *name* (E.g. p14,57,59). What we actually do is fix the reference of a *word*, and in so doing, *make it into a name*. We could have made it into a verb or an adjective or a preposition. It is entirely up to us. This failure to distinguish between a word and a name not only undermines Kripke's distinction between descriptions that give a synonym and those that merely fix reference, it may also indicate that philosophers of *all* persuasions do not really know what a name is.

DRT has fallen woefully short of its aims as far as the negative thesis is concerned, although this may be in part due to the fact that its aims are too ambitious. But perhaps all is not lost for DRT. It could be that we will have to look again at the sense of proper names given by attributive descriptions; as well as the role of indexicals in fixing reference. For example, if we are forced to modify the description 'The first man on the moon' as 'The first man to set foot on *that*', then we introduce indexicals into the picture. If we modify the sense to read 'The man *commonly believed* to have been the first etc...' then we are relying on other speakers in order to effect reference. This would not be a purely Fregean theory. It does, however, lead us to consider Kripke's positive thesis - the so-called 'Causal Theory'.

CHAPTER TWO

THE POSITIVE THESIS

DRT is convinced that proper names do not refer by means of senses given by definite descriptions. How then, do they refer? This is the concern of Kripke's positive thesis in *Naming and Necessity* - the so-called 'Causal Theory of Reference'. According to the causal theory, reference is effected by means of a causal chain of communication which leads from our current use of a name all the way back to an initial baptism of an actual object. Thus, language is causally connected to the world. Kripke gives an outline of the theory on page 91 of *Naming and Necessity*:

“Someone, let's say, a baby, is born; his parents call him by a certain name. They talk about him to their friends. Other people meet him. Through various sorts of talk the name is spread from link to link as if by a chain. A speaker who is on the far end of this chain, who has heard about, say Richard Feynman, in the market place or elsewhere, may be referring to Richard Feynman even though he can't remember from whom he first heard of Feynman or from whom he ever heard of Feynman.”

The main point about the causal theory is that, not only is reference passed from speaker to speaker in a causal chain, thereby giving the speaker a causal connection with the referent, but reference is effected *by means of this chain*, and the causal chain alone. No discriminating information is required to be able to refer to, say, Feynman. All that is required is that we be exposed to the name. This bare exposure is necessary and sufficient for the speaker to be referring to the referent of the name.

Not only does this theory apply to proper names, it also, rather surprisingly, applies to general terms, such as names for natural kinds. (More of this in chapter 3 where we shall also look at Putnam's version of the causal theory). Thus the causal theory provides a natural mechanism connecting language to the world.

The two main features of the causal theory are the **initial baptism**, where the name is first introduced, and the notion of **reference borrowing**, where the reference of the name is passed from one speaker to another. On p96 of (1980) Kripke further explains the theory thus:

“An initial ‘baptism’ takes place. Here the object may be named by ostension, or the reference of the name may be fixed by a description. When the name is passed from link to link, the receiver of the name must, I think, intend when he learns it to use it with the same reference as the man from whom he heard it.”

Now there is no doubt that something like this happens with most of the names we know. That is, we *acquire* names in something like the way Kripke claims. There are, however, serious problems with this hypothesis as a way of explaining how reference is effected. The main difficulty is with the notion of reference borrowing and the intention to refer to whatever the other speaker is referring to. How exactly is this accomplished? I have never *consciously* intended to refer to anything, and I wouldn't know how to anyway. Actually Kripke does say that the intention to refer is a *given* (Ibid. p97); but if it is just given that in acquiring a name we refer to the object that the other person is referring to, why do we need to talk about intending to refer at all?

Searle (1983) takes issue with this point, accusing Kripke of introducing an Intentional element into the causal theory - something that should not be there in a purely externalist theory. Searle then gives counter-examples to show that the causal chain is neither necessary nor sufficient for names to refer (1983, p237, 240). Although the alleged counter-examples are ultimately not very satisfying Searle is surely correct in claiming that the causal theory should not require this internalist Intentionalist element in order to effect reference. In support of this claim we could argue that any intention to refer is actually irrelevant to a causal theory. For example, if a name does not refer, then we can *intend* to refer until we are blue in the face, but we will not refer. On the other hand, suppose we intend *not* to refer using a name which actually has a referent, then we will surely refer, due to the causal chain. The intention to refer is therefore *irrelevant* to whether or not we actually refer, and yet it is an essential component of the causal theory.

Another major flaw in Kripke's theory is the way in which names are actually passed from speaker to speaker. Kripke seems to think that it is enough simply to mention the name and it has been passed on. But this is not what actually happens. We pass on names in certain contexts, along with descriptive information which

could give the sense of the name, or could merely have a reference fixing role. Whatever the role of this descriptive information it is necessary to the functioning of the name. Kripke's 'Feynman' example will simply not work. It is not enough to hear the name 'Feynman' mentioned in the market place in order to be able to use the name to refer to Feynman. This is because it is not enough to refer to someone simply to mention the person's name. We must be able to *use* the name *in a sentence*, to *say something about that person*. Now in Kripke's example, what could anyone say about Feynman if all they have is the name? Nothing, surely. The speakers in question could not even say that Feynman was the person being talked about in the market place because there could be many people called 'Feynman'. All they are entitled to say is that the people in the market place were talking about someone called 'Feynman'. They can mention the name but not actually use it to refer, because they cannot use it at all. Similarly, if the speaker in question goes to the library and asks the librarian if she has any books by or about Feynman the name is merely being mentioned rather than used. Kripke's version of the causal theory is too simplistic to be of any use as a theory of how names refer.

Gareth Evans (1973;1982) recognises the inadequacies of Kripke's rather naive version of the causal theory while retaining its basic principle. In his (1973) Evans appeals to the body of information the speaker associates with the name, as well as causal origin, in order to explain how reference is effected. This causal origin is a necessary but not sufficient condition for an object to be the intended referent of the use of a name (Ibid. p218). Intention to refer is another necessary condition, although for Evans, intentions must be 'manifest', which means manifest in the community. As he says on p222 of (1973):

“Intentions alone don't bring it about that a name gets a denotation. Without the intentions being manifest there cannot be the common knowledge required for the practice.”

Evans' version of the causal theory is that

“ 'NN' is a name of x if there is a community C

in which it is common knowledge that members of C have in their repertoire the procedure of using 'NN' to refer to x (with the intention of referring to x)"

(In Moore, Ed. 1993, p222)

This development of Kripke's basic theory, which Evans regards as simplistic, even 'magical', has impressed many. Whether we should be impressed, however, is another matter. For instance, what exactly is this 'procedure of using 'NN' to refer to x'? What does it consist in, and how was it established? How does the community's procedure differ from that of the individual? Unless these questions can be answered it seems that all Evans has done is shift the problem from how the individual refers to how the community refers. And there is still the Intentional element which really should not belong in a causal theory.

Donnellan also attempts to defend a version of the causal theory, but in a different way from Evans. He sets out his theory in a paper, "Speaking of Nothing", in the following way:

"...when a speaker uses a name intending to refer to an individual and predicate something of it, successful reference will occur when there is an individual that enters into the historically correct explanation of who it is that the speaker intended to predicate something of" (1974, p16)

In order to explicate the idea of the historically correct explanation Donnellan introduces the notion of an omniscient observer who can see the relation between the individual and the explanation. For example, our use of 'Socrates' will refer to Socrates if an omniscient observer sees an individual related to the author of certain dialogues; that one of the characters in the dialogues was modelled on that individual, etc., and that the speaker is talking about this individual is explained by his having read these dialogues.

Clearly we cannot know if the omniscient observer is correct about any particular name. We cannot know if the historical explanation is correct or, in many cases, whether our use of a name really does go all the way back to an actual individual. Presumably DRT would argue that we do not need to know these things. I disagree, for the following reasons. First of all, any theory of naming is only applicable to names; but no theory of naming ever tells us that there *are* any names. That is, they

purport to tell us *how* names refer but not *that* they refer. Now we must know this. We must know if any putative name really is a name - i.e. really denotes - otherwise we do not know whether it is a genuine name or not. If we don't know whether any putative name really is a name there is no point in debating which theory of naming is most appropriate to explaining how this 'name' functions. The best that any theory of naming could do would be to say that , **if** 'N' is a name, then it will function in a certain way. Or, **if** there are any names at all, then they will function in a certain way. Now this is not good enough if we want to actually use names: especially if we want to be able to conclude anything about the necessary a posteriori or the contingent a priori from a theory of naming. Quite simply, we need to know of any putative name whether or not it is a genuine naming device - i.e. whether or not it actually names anything.

Now some of our 'names' do not refer to anything - for example, names of fictional characters or mythical creatures; but most of our names do in fact refer to some actual object and we have various ways of establishing this fact. Most of the names we use are of people we know personally or have seen on television or in photographs. In these cases we can *see* that our putative names really do refer. They refer to *that person there* - either in the flesh or in a photographic image. It is as though we have our own baptism.

Now the ironic thing about this way of establishing that our names really do refer is that the means by which we establish this actually renders the causal chain *redundant* as a means of explaining *how* these names refer. Once we are in a position to point to the actual referent, or a photograph of one, we no longer require the causal chain as a means of determining reference (although we are causally connected to the referent). As for 'names' like 'Thales' or 'Homer' or 'Robin Hood' we simply do not know if they refer, and so it may be that we cannot talk about reference here.

There are many other names, however, where things are not quite so straightforward. These are names of objects we cannot point to but nevertheless there is little doubt as to their existence. A prime example would be the name 'Julius Caesar'. We cannot point to Julius Caesar, or to a photograph of him, and yet we are in a much stronger position regarding his existence than we are with the likes of

Thales or Robin Hood. This is because the existence of Julius Caesar is *well documented*. It is this documentation, given in terms of descriptions, that tells us that Julius Caesar existed and that therefore 'Julius Caesar' is a genuine name. Now of course this documentation could be passed to us through some kind of chain; although a chain as such is not necessary. We could for instance discover a written account of Caesar's life written while he was still alive, and so we go straight to the source, as it were. We do not need any long chain of communication. Of course we still seem to be causally connected to Caesar in some way, but the point is that it is the descriptive information in the document that tells us of Caesar's existence and that effects reference for us - if anything does - but this would not be a causal chain since the documentation could pass through the hands of people who did not understand a single word of it. The documentation therefore not only tells us that 'Julius Caesar' refers, it is what effects reference for us. If this documentation is not sufficient to effect reference then we will simply have to suspend talk of reference in cases of this type.

The causal chain is redundant as far as referring is concerned. Even if it was not, it is very difficult to see how it would actually work. Kripke's basic theory is simplistic, even naive, and Evans and Donnellan are of no help. The notion of intending to refer is obscure, unnecessary to a genuine causal theory, and irrelevant as far as effecting reference is concerned.

Kripke may, however, have shown that a causal connection to the referent of a name is necessary for reference to take place, if only because a causal connection is inevitable; albeit that this connection is established by demonstratives rather than ordinary proper names. He has not, of course, shown that this causal connection is sufficient. We shall deal with the problem of how reference is actually effected in chapter 3.

In the meantime, the notion of a causal connection between name and referent once again raises the problem of what a proper name contributes to the content of a sentence which contains it; and it is to this that we finally turn.

DRT: A REDUCTIO

As previously mentioned DRT is on shaky ground when it claims that proper names are not synonymous with definite descriptions, unless it can be clear about what is to count as meaning. It seems that DRT has only two choices here. Either proper names have no meaning at all, or their meaning is given by what they refer to. To argue that names refer directly but still have a sense which determines their propositional content has not been regarded as an attractive possibility by most philosophers. (An notable exception is Evans in his (1982) where he argues that direct reference need not rule out sense for proper names). For most direct reference theorists, however, there seem to be serious problems when we come to consider what a proper name contributes to propositional content. In what follows we shall consider the contribution made to content of co-referential proper names in epistemic and doxastic contexts, which will lead to a *reductio ad absurdum* of DRT.

We begin with an example. Take the following sentences:

(iii) Hesperus is a planet

(iv) Phosphorus is a planet

Now if the names 'Hesperus' and 'Phosphorus' contribute their referents to the content of these sentences then (iii) and (iv) should express the same proposition. It seems to follow from this that anyone who knows or believes (iii) also and thereby knows or believes (iv), regardless of what they may say to the contrary. At the very least this is counterintuitive, given the way we speak. It might even be regarded as some kind of *reductio ad absurdum* of DRT.

There are two points to be made here. First of all, the foregoing is not the *reductio* we are aiming at. Secondly, it will be argued by some that not all advocates of DRT are committed to the view that anyone who knows or believes (i) knows or believes (ii). In a footnote to his (1979) Kripke says that DRT is not committed to the

substitution of co-referential terms in epistemic contexts, even if it actually does so in modal contexts (1979, Footnote, 10). This is highly questionable, especially since Kripke gives us no account whatsoever of content. His “Puzzle About Belief”, however, indicates that even he is aware of the potential for serious problems in this area.

Some advocates of DRT see no problem in accepting the foregoing consequences. Fitch (1976) and Salmon (1986) notably defend the claim that (iii) and (iv) express the same proposition, which leads them to conclude that we know a priori truths of the form $a=b$ - contrary to Kripke’s claim. It also leads Salmon to embrace the seemingly counterintuitive claim that anyone who knows or believes (iii) also knows or believes (iv). The Description Theory of course concludes the opposite since (iii) and (iv) will express different propositions due to the different senses of the two names.

The reductio of DRT stems from the claim that (iii) and (iv) express the same proposition. Assuming this to be so, and assuming that we are acquainted in some sense with these propositions, then we would know that the names ‘Hesperus’ and ‘Phosphorus’ must make the same contribution to propositional content. The only way in which they can make the same contribution is if they are coreferential; which means that their referents would be identical. In other words, DRT is committed to the view that we would know that Hesperus is Phosphorus because we know that Hesperus is a planet and Phosphorus is a planet - or by reason alone, which is absurd. This is a reductio ad absurdum of DRT, so long as we accept that we really do know these propositions.

One way in which DRT might reply is to deny that we are acquainted with these propositions - or indeed any propositions containing proper names. This is the view of John Tienson in his (1984) where he attempts a reductio of what he calls ‘The Sophisticated View’, (the view that anyone who knows or believes (iii) also and thereby knows or believes (iv)), while preserving the basic tenets of DRT. For Tienson, it is a consequence of DRT that ordinary proper names are to be regarded as something like Russell’s logically proper names because they are direct designators. But if this is so then surely we must be acquainted, in some sense, with their bearers, and hence the appropriate propositions. Whether we accept that

reference is direct or indirect we surely are acquainted, in some sense of the word, with the propositions expressed by sentences containing most of the proper names we know (Otherwise, how do we know that they do actually express propositions?). So the *reductio* still stands.

One possible response to such an alleged *reductio* is to argue that although we are acquainted with the singular propositions expressed by sentences such as (iii) and (iv) we actually grasp the proposition in two different ways; or under two different ‘guises’, as Salmon puts it. For Salmon we need a whole new analysis of the belief relation in order to make sense of the Sophisticated View. The traditional binary relation in which an agent simply grasps a proposition must be replaced by a ternary relation in terms of an agent, a proposition, and a guise under which the proposition is grasped (1986, p111). Unfortunately for Salmon he is unable to tell us what kind of things these guises are. He does admit, surprisingly, that they must be something like Fregean senses, although guises would not be part of the semantics of the sentences in question. We could certainly dispute this last point on the grounds that Salmon seems to be committed to the consequence that no-one actually grasps *any propositions at all*. All that is grasped is the way-of-being-given, the mysterious third relatum. If this is correct then we are really dealing with a theory which is difficult to distinguish from a Fregean one. The notion of a proposition distinct from these ways of being given would be redundant. Even if we do not take this line, we can surely ask why it is that these guises are sometimes successful, sometimes not. Why is it, for example, that when Lois Lane (to use Salmon’s example) is presented with the proposition that Clark Kent can fly through the sentence “Clark Kent can fly” she does not recognise the proposition and so does not assent to the sentence, even though she actually grasps the proposition? How can we grasp propositions yet not recognise them?

Yet another flaw in Salmon’s theory concerns the possibility of false belief. For example, Lois surely believes the proposition that Clark Kent cannot fly since she would assent to the appropriate sentence. The Sophisticated View is going to have to accuse Lois of knowingly holding logically incompatible beliefs (the belief that Clark Kent can fly and the belief that Clark Kent cannot fly), or else maintain that

there is something incoherent about speakers holding false beliefs. The Sophisticated View can maintain neither option.

As if the foregoing were not enough we can deal a fatal blow to Salmon's defence of the Sophisticated View in terms of the behaviour of the agent. For example, why does Lois behave in a totally different way when she is with the person she calls 'Superman' from the way in which she behaves in the presence of the person she calls 'Clark Kent', if she believes the same things of them? Even if we accept Salmon's guises we cannot accept that Lois believes that Clark Kent can fly if she does not behave appropriately. What we normally call knowledge and belief have the potential, at least, to affect people's behaviour. They can actually affect behaviour in the most profound way. Salmon's so-called 'De Re' knowledge and belief do not even have the *potential* to affect behaviour. Consequently, they are no kind of knowledge or belief.

Obviously our reductio requires a premise to the effect that DRT is committed to the Sophisticated view. Recently, however, some defenders of DRT, notably John Perry and Mark Crimmins, have come to reject such a consequence. In their (1989) they make a sophisticated and comprehensive attempt to defend DRT while rejecting the Sophisticated View. On the Perry/Crimmins view a speaker can believe that Hesperus is a planet while failing to believe that Phosphorus is a planet, even though the two sentences involved express, or encode, the same proposition. In short, we can get two beliefs out of one proposition. This notable feat is accomplished by, first of all, distinguishing between the content of a belief, viz, a proposition, and its *constituents*, which on the Perry/Crimmins view, are concrete particulars in the mind of the individual. On this account, beliefs are concrete cognitive structures - "things in the head". They are "particulars that belong to an agent, come into existence, endure and go out of existence" (1989, p688). Furthermore, a belief comes into existence when an agent forms it; "it is not the sort of thing that is around for the agent to adopt"(Ibid.) The constituents of beliefs Perry and Crimmins call 'notions' and 'ideas'. Notions are associated with ideas in forming beliefs and are things in the mind that stand for things in the world (p692).

To illustrate how Perry and Crimmins get two beliefs from one proposition let us consider their example of the beliefs of the fictional character Miles Hendon from the book *The Prince and the Pauper*. In the story Miles comes across a boy dressed in rags claiming to be the Prince of Wales. Unbeknownst to Miles the boy really is Edward Tudor, Prince of Wales. Even so, on the Perry/Crimmins view we are entitled to conclude that

(v) Miles believed that he (the boy in rags) was not of royal blood

and also

(vi) Miles believed that Edward Tudor was of royal blood

We get two beliefs from the one proposition:

<< Being of royal blood; Edward Tudor >>

because Miles' beliefs involve two different notions of Edward Tudor. One is a visual notion formed when Miles saw the boy in rags, and is associated by Miles with ideas of being a pauper, not being royal, and being insane (since the boy claims to be a prince). The other notion need not be a visual one. It is associated with ideas of being a prince, being rich, not being a pauper etc. Moreover, when we report beliefs such as (v) and (vi), these notions are unarticulated constituents of the content of the report. They are propositional constituents that are not explicitly mentioned (p697).

Clearly Perry and Crimmins have developed a much more complex analysis of belief ascription than Salmon. Is it any real improvement on Salmon, though? The answer has to be in the negative, for the following reasons. First of all Perry and Crimmins fail to make an adequate distinction between notions and ideas. They do say that the difference is that between ways of thinking of individuals versus properties (p690), but surely a way of thinking about an individual must involve properties -which blurs the distinction. Whether it does or not is surely one of the

things at stake in the debate between DRT and the description theory. Secondly, the failure to say exactly what a notion is makes it virtually impossible to understand how a notion, which is supposed to be a concrete particular in someone's mind, can be a constituent of an abstract object like a proposition, albeit an unarticulated constituent.

Indeed the whole idea of unarticulated constituency could be seriously questioned. It is all very well to say that in a statement like "It is raining" there is an unarticulated constituent concerning a certain time and place provided by the context; but surely another thing altogether to claim that in reporting a belief like (v) the proposition involved has an unarticulated constituent which seems to allow Miles to hold a belief which is the exact opposite of its propositional content!

It is certainly very strange to make a distinction in the first place between constituents and contents of propositions, or anything else. Surely the content of anything has to be the *sum* of its constituents, not something actually distinct from them. We can also question how beliefs can be concrete when their contents are abstract.

Yet another problem for the Perry/Crimmins view concerns what happens when a belief ceases to exist. Presumably the proposition does not cease to exist, and the two notions could certainly persist. The only way a belief can cease to exist is if the relevant notion ceases to be associated with the appropriate ideas. But this requires a rigid distinction between notion and idea, which has not been made by Perry and Crimmins.

One final point against Perry and Crimmins, and a crucial one. Again it involves the issue of false belief, which troubled Salmon. Quite simply, it seems to be a consequence of the Perry/Crimmins view that we cannot hold false beliefs. If we take, for example, Miles' belief that the boy in rags is not of royal blood, it can be seen that this is a false belief involving the proposition:

<< Being of royal blood; Edward Tudor >>

Now if Miles has this belief, and he knows anything about propositions, he must also believe that the content of his belief is the proposition:

<< Being a pauper; the boy in rags >>

He cannot believe that the content of his belief is << Being of royal blood; Edward Tudor >> because then he would have to give up his belief that the boy in rags is a pauper. He can only believe that the content of his belief is the proposition << Being a pauper; the boy in rags >>. Now this is a false belief, so Miles has a false belief about his belief, and this would appear to be incoherent in the present context. We can have false beliefs about the world, but we cannot have false beliefs about our beliefs - at least, if we are rational beings and characterise beliefs in the way that Crimmins, Perry and Salmon have done. The view that we can get two beliefs from one proposition leads to the consequence that we cannot hold false beliefs about the world without holding false beliefs about the content of those beliefs. Which means that we cannot hold false beliefs about the world. But we surely *can* hold false beliefs about the world - therefore we cannot get two beliefs from one proposition.

The Perry/Crimmins view fails in its attempt to account for the behaviour of coreferential terms in belief contexts. It is therefore no threat to our reductio.

There is perhaps one last chance for DRT to save itself, and that is Kripke's (1979) "Puzzle About Belief". In this paper Kripke comes the closest he has ever come to saying anything about what proper names contribute to content. It is not, however, the purpose of his paper to say what names actually contribute, but, rather to argue that the situation is more complicated than many philosophers seem to think.

Kripke generates his puzzle in the following way. First of all, he introduces what he calls 'The Disquotational Principle', which he formulates thus:

"If a normal English speaker, on reflection, sincerely assents to 'p', then he believes that p" (In Salmon and Soames, 1988, p113)

This principle is a necessary condition for the puzzle to arise and is regarded by Kripke as something of a self-evident truth (Ibid.) Suppose we now consider an

example of a rational, competent speaker of English - call him 'Peter' - who sincerely assents to the sentence

(vii) Paderewski had musical talent

where Paderewski is a famous pianist. Unfortunately Peter also assents to

(viii) Paderewski had no musical talent

believing that he is here dealing with a Polish statesman, distinct from the pianist, who had no musical talent. Now by the Disquotational Principle we seem to be committed to ascribing contradictory beliefs to Peter. But we surely cannot do this since Peter would never knowingly hold contradictory beliefs. Hence the puzzle.

Another way of posing the problem is to say that we cannot answer the question "Does Peter, or does he not, believe that Paderewski had musical talent" - a question we should be able to answer. According to Kripke we have a situation here where our ability to ascribe belief comes under severe strain, and may even break down altogether. If this is so then we should not be too quick to come to conclusions about what names contribute to content because it is as much a problem for the Description Theory as it is for DRT. This is because we cannot take what might seem the obvious way out of the difficulty and appeal to a difference in sense between the two occurrences of the name 'Paderewski'. No substitution of coreferential names is involved here. The two occurrences of 'Paderewski' are two tokens of the *same name*, and so there is no appeal to a difference in sense.

Nevertheless, attempts at solving the puzzle usually take the form of appeals to different senses for the two name tokens (E.g. Bach, 1987; Katz, 1990; and Schiffer, 1977). Peter himself would certainly associate different senses with what he regards as different names, but Kripke would not accept this way out because we are simply not dealing with two different names, regardless of what Peter believes. There is no need to do this, however. The 'puzzle' is quite simply not a genuine puzzle. What breaks down in these type of cases is not our ability to ascribe belief but the Disquotational Principle. We cannot apply the Principle in the cases where the

puzzle is supposed to arise because it requires us to ascribe contradictory beliefs. If the Principle cannot be applied in such cases then there is no puzzle because the Principle is the necessary condition for the puzzle to arise. So the alleged puzzle has been dissolved rather than solved. The irony here is that Kripke has actually made too strong a case in the sense that, if it is *impossible* to answer the question about what Peter believes, then it is just not a proper question.

Now some people will not accept such an easy way out. They will point out that we really could find ourselves in similar situations to the 'Paderewski' case. It is not some outlandish thought experiment, but a very real possibility, and one which we will have to deal with.

Even if we accept this, though, there is still no real problem. Even if we accept that we cannot ascribe belief in this type of case, where the agent has the mistaken belief that he is dealing with two people rather than one, what does this tell us about our ability to ascribe belief, **in general**? And what does it tell us about how names function, **in general**; or even in these type of examples? Nothing at all. If someone asks us how we are to *characterise* Peter's beliefs in such cases we can answer that there is no need to, although we can *explain* why he has them, or seems to have them. He holds these beliefs simply because he has the mistaken belief that he is dealing with two people rather than one. Kripke's alleged puzzle is a complete failure, as are the arguments advanced by Salmon, Perry and Crimmins. The *reductio* still stands.

In conclusion, then, we have seen that DRT has failed to show that ordinary proper names are not synonymous with definite descriptions. The best it can do is to show that names are not synonymous with the attributive use of descriptions; but even this can be challenged if we formulate the description in a certain way. In particular, Kripke has not shown that names are rigid designators; although it would seem to be the case that demonstratives are rigid. He has not succeeded in making a distinction between fixing reference and giving a synonym, especially where only one description is available. This has serious consequences for his attempt to establish the existence of contingent a priori truths.

Kripke's positive thesis fares no better. It simply does not work - although he may have shown that a causal connection to the world is necessary for reference to take place.

DRT can in fact be reduced to absurdity on the question of what a proper name contributes to the propositional content of a sentence which contains it - especially in epistemic and doxastic contexts.

The failure of DRT to distinguish between a name and a mere sign or utterance suggests that Kripke et. al. literally do not know what a name is. This is also a problem for the Description Theory. In fact it could be that all theories of naming are superfluous since they never tell us whether there are any names or not. We need to know this if names are to be used in any way - particularly if we are to conclude anything about the necessary a posteriori or the contingent a priori.

In chapter 3 we shall examine the Direct Reference Theory as applied to general terms - specifically, terms for natural kinds. This may seem superfluous if it is accepted that DRT has already been reduced to absurdity as regards proper names; but this is not so. Because of the totally different nature of general terms and the different kinds of properties involved in any alleged sense of a general term, I feel that it is necessary to examine the possibility that, although DRT does not apply to proper names, it may apply to natural kind terms.

Chapter 3 will also contain what I think is the true theory of how names, particularly general names, actually function.

CHAPTER THREE

(General Terms)

In Lecture III of *Naming and Necessity* Kripke extends his theory of direct reference to general terms; specifically, natural kind terms like 'gold' and 'tiger'. The arguments used against a sense theory for ordinary proper names are used against a sense theory for general terms. It is Kripke's contention that general terms function in a similar way to proper names in that they are rigid designators and do not refer by means of a sense. This is where he differs from Mill, who held that general terms, unlike proper names, do indeed have a sense. For example, it may seem obvious to some that the sense of 'tiger' could be given in terms of a list of properties we believe tigers to possess. Such a 'definition' could be something like "Large carnivorous quadrupedal feline, tawny yellow in colour with blackish transverse stripes and white belly" (1980, p119). Kripke employs his usual anti-sense arguments to show that 'tiger' is not short for such a definition.

First of all, we would not contradict ourselves if we were to declare that we had seen a three-legged tiger. Secondly, there is no contradiction in supposing that there are creatures having all the phenomenological properties of tigers but which are not tigers, but reptiles, for instance (Ibid. p120). A Godel-type argument is also used to show that 'tiger' cannot be synonymous with anything like a dictionary definition consisting of a list of properties because we could be mistaken about some, or even all, of these properties (Ibid. p121). According to Kripke we could discover that tigers do not in fact have a particular property we believe them to have; or indeed any of the properties we believe tigers to have; which shows that 'tiger' cannot be associated with a cluster concept. Consequently, natural kind terms like 'tiger' or 'gold' are not synonymous with any list of properties which give the meaning of these terms. A statement like "Tigers are striped" is neither necessary nor analytic in Kripke's view. It is certainly not known a priori, but is an empirical discovery. It follows from all this that natural kind terms are rigid designators which leads Kripke to conclude that there are necessary a posteriori truths involving natural kind terms - the so-called 'Scientific Identities', like 'Water is H₂O' or 'Heat is the motion of

molecules'. This is equivalent to the claim that science can discover the real essence of natural kinds (although it need not know that it has done so) - a very strong claim indeed.

As with many of the claims of *Naming and Necessity* Kripke does not offer a full-blown theory of the semantics of natural kind terms, but rather, an outline of a theory. Perhaps he does not think he needs to offer a detailed theory since he has already dealt in depth with the semantics of proper names and he considers general terms as functioning in a similar way to proper names. The question we must now ask is "Has Kripke been too hasty in equating general terms with ordinary proper names?" After all, it is far from obvious that general terms function in anything like the same way as ordinary proper names. What is certain is that the putative senses of natural kind terms involve properties of a completely different kind from those involved in the senses of proper names. For instance, it is one thing to say that Nixon might not have been a politician, but another thing entirely to say that tigers might not have been carnivorous, or four-legged. These latter are intrinsic to tigers, whereas the property of being a politician does not seem to be intrinsic in any way to Nixon. So even if Kripke's anti-sense arguments work against ordinary proper names - which is extremely doubtful as we have seen - it is by no means certain that they will work in the case of natural kind terms, where the putative sense is given in terms of different types of properties.

Actually Kripke's arguments are no more effective in the case of natural kind terms than they are with proper names. Of course we would not contradict ourselves if we declared that we had seen a three-legged tiger, but this is because the definition of a tiger applies to *normal* tigers. A three-legged tiger or an albino tiger would not be normal tigers. (They are still tigers, however, simply because their parents were tigers). As for the possibility of discovering that we were mistaken in thinking that tigers had a particular property, this merely shows that we could be mistaken about the meaning of the term 'tiger'. As with proper names we could simply adjust the sense of 'tiger' to read something like 'Creature which is *commonly believed* to normally have such and such properties'. The linguistic community could not be mistaken about this sense since it is the linguistic community that commonly believes. As for Kripke's other possibilities - the

creature which resembles a tiger in all its phenomenological properties but which turns out to be a reptile, or the possibility of discovering that tigers never have any of the properties we commonly associate with them - it is highly doubtful whether these are real possibilities at all. No animal could resemble a tiger and be a reptile because, in order to fully resemble a tiger, it would have to have a fur coat. Quite simply, no animal with a fur coat could be truly cold-blooded; and therefore could not be a reptile. Furthermore, we could not actually discover that tigers had none of the properties we believe them to have without doubting the very existence of tigers. If, for example, we had a tiger in a cage, and one morning we go to the cage and discover a creature with none of the properties we commonly associate with tigers, we would probably conclude that someone had stolen our tiger and replaced it with this strange creature. It is just unintelligible to argue that tigers could turn out not to have any of the properties we believe them to have because if we cannot be reasonably sure of at least some of the properties of a natural kind, then we simply cannot talk about natural kinds at all. (There is certainly no example of a natural kind term where we have been mistaken about all the properties of the kind)

So it seems that Kripke has been premature in making the transition from proper names to general terms. Another major difference between ordinary proper names and natural kind terms concerns what it is that they are supposed to refer to. It is fairly clear, for example, that the name 'Nixon' refers to the man Nixon (multiple bearers apart); but to what does the name 'tiger' refer? Does it refer to a particular tiger or to the whole species? What is it for a name to designate a whole species anyway - and how does a species term like 'tiger' designate the same species across possible worlds when the actual individuals who make up the species can vary from world to world? These are questions that Kripke does not adequately address in his rush to conclude that general terms function in the same way as ordinary proper names, but which will have to be dealt with at some point.

Kripke's treatment of natural kind terms can be seen as inadequate in many ways. It is left to Hilary Putnam to give us a detailed account of the semantics of natural kind terms within the framework of DRT, and it is Putnam's work in this area which will be the main concern in this chapter. Although Putnam has changed his mind about metaphysical necessity (among other things), preferring physical necessity

instead (See Putnam, 1990, Ch 4), he has steadfastly maintained his position on the meaning of natural kind terms. Despite Putnam's divergence from Kripke on the topic of necessity, a detailed examination of his theory of meaning may help Kripke's claims about metaphysical necessity, even if that is not what Putnam intends. Even if Putnam's views on meaning cannot be upheld, as I shall maintain, they will lead us to an alternative theory of how proper names and natural kind terms actually do function.

Putnam on Meaning

In works such as "Is Semantics Possible?" (1970), "The Meaning of Meaning" (1975), and *Reason, Truth and History* (1981), Putnam challenges what might be termed 'The Traditional Theory of Meaning' for general terms - specifically, terms for natural kinds like substance terms and species terms. The view he wants to criticise is characterised by him as follows:

(I) That knowing the meaning of a term is just a matter of being in a certain psychological state.

(II) That the meaning of a term (in the sense of 'intension') determines its extension (in the sense that sameness of intension entails sameness of extension).
[1975, p219]

'Intension' can be equated with sense or concept, or sometimes with 'mental representation' (See 1988, Ch.2) which determines what the term refers to. 'Intension' can also be identified with a list of properties or descriptions which give necessary and sufficient conditions for anything to belong in the extension of the term, such that, for any natural kind **K**, the statement "All **K**'s have property P" or "Anything with properties P1.....Pn is a **K**" will be analytic (where P1.....Pn are all the properties in the intension of 'K').

It seems to follow from assumptions (I) and (II) that it is something in the mind of the speaker that determines the reference of a word. Thus, psychological state determines reference, where 'psychological state' is taken to be a 'narrow' state. A narrow state is one which, according to Putnam, does not presuppose the existence of anyone or anything else (1975, p220). More importantly, an individual does not have such states in virtue of any causal connection with his environment. How he does come to have these states seems mysterious. Frege, for example, simply says that sense is grasped by anyone who is familiar with the language - although it should be pointed out that for Frege, grasping sense is not a narrow state. (Fodor and Chomsky actually hold that concepts are innate).

Putnam's challenge to the Traditional Theory takes the following form:

- (i) There is nothing in the mind of the speaker (in the narrow sense) which has the *intrinsic* ability to determine what a word refers to. This would be a 'magical theory' of reference.
- (ii) It is essential that there be a **causal, non-conceptual** connection between the world and the speaker's use of any term to refer.
- (iii) Reference is determined **partly indexically, partly socially**. (There is a division of linguistic labour)

Putnam begins by attacking the notion of a normal member of a species or kind (1975, p 140), attempting to show that the properties in the intension of a term do not give **necessary** conditions for anything to be of that kind (His attack on the *sufficiency* of the conditions comes with his 'Twin Earth' paradox). Is it not possible, he asks, that what we call normal lemons are in fact atypical? He then points out that the characteristics of a natural kind could change over time without affecting the essence of the kind. Both these considerations undermine the claim that a natural kind can be defined in terms of normal members having certain properties. If this is correct then we can use it to defend Kripke's claim that the Traditional

Theory is committed to the view that a statement like “Today I saw a three-legged tiger” would be a contradiction in terms, from the objection that it is normal tigers that have four legs and so we do not contradict ourselves in this way. According to Putnam it is simply not clear what is to count as normal.

Putnam then proceeds to a possible definition of natural kind terms which accommodates these points but leads to a *reductio ad absurdum* of the Traditional Theory. The definition goes as follows:

X is a lemon=df **X** belongs to a natural kind whose normal members have yellow peel, tart taste etc. OR **X** belongs to a natural kind whose normal members used to have these properties OR **X** belongs to a natural kind whose members were formerly believed to, or are now incorrectly believed to have these properties. (1975, p143)

Now suppose that a long time ago, before lemons were known, people found a few atypical oranges which had all the phenomenological properties of lemons. If these people took these objects to be normal members of a species, then they thought that oranges have the properties that lemons in fact have. It would follow from this and the above definition that all now existing oranges would be lemons! Hence, *reductio ad absurdum*.

The notion of a normal member of a species is a crucial one because if it makes no sense to talk of normal members of a kind it would seem to be very difficult, if not impossible, to talk about kinds at all. It seems that Putnam has not realised this. (Of course it may be that we cannot talk about natural kinds but that would be to shift the debate to a different area. In the present context we are assuming that there are such things as natural kinds and that we can make empirical discoveries about them. Kripke and Putnam certainly make this assumption). In fact there is much about natural kinds that Putnam has not realised. To begin with, how can the members of a species we regard as normal actually be atypical? The things we call normal lemons, for example, number billions and make up the vast majority of the species ‘lemon’. How can the vast majority be atypical of a species? They are surely typical *by definition* - simply in virtue of being the majority. We have nothing else to go by. The only way these lemons could be atypical would be if there was some standard of normality to which they might fail to conform. But what could this

standard be? As far as we know at the moment it could only be the DNA of lemons: but would this be the DNA of our 'normal' lemons, or would it be the DNA of our 'abnormal' lemons? The problem as posed by Putnam, still stands.

Actually there is no problem here because the only way to define a normal lemon is in terms of the majority of lemons. If the abnormal ones increased in number to the point where they became a sizeable minority, we would simply conclude that we were dealing with two species of lemon. Likewise, if we discovered that our 'normal' lemons which number billions here on Earth were actually a tiny minority of all the lemons in the universe and that the majority in the universe were like our 'abnormal' lemons, we still need not conclude that we were mistaken in thinking that our lemons are normal members of a species. There would just be more than one species of lemon. Putnam simply does not understand the concepts of 'typical' or 'normal' as applied to natural kinds. In attacking these notions he undermines his own position.

Putnam's second point is that the characteristics of a natural kind could change in some ways without the kind ceasing to be the same kind. This is true, although it is debatable to what extent these characteristics could change - or how many of them could change - without the kind ceasing to be the same kind. Putnam's example in "Is Semantics Possible" is not a good one. He considers the possibility of lemons changing their colour because of a change in atmospheric conditions which does not affect the essence of lemons. Again, though, this is easily accommodated into the Traditional Theory. The simple fact is, that if some of the characteristics of a kind did change, the meaning or concept of the kind term would change accordingly. This would be no different from the way in which a term like 'solicitor' could change its meaning (up to a point) if the function of solicitors changed. Furthermore, a normal member of a kind can be, and probably should be, defined by its apparent properties in standard conditions.

As for the alleged reductio of the Traditional Theory, it is no reductio, for the following reasons. First of all, these people surely do not have the right to conclude that a few samples are typical of a species. Secondly, if these few samples have all the phenomenological properties of lemons, even down to the taste, by virtue of what would they actually be oranges? It could only be in virtue of their DNA

structure. But then we would have a situation where two completely different types of samples, normal oranges and these lemon-resembling oranges, would be members of the same species. How is this possible if normal oranges and the atypical oranges are so completely different, not only in phenomenological properties, but even down to their DNA? If they do not differ in their DNA then we would have a situation where the same DNA can make some oranges look and taste like oranges, and make some look and taste like lemons! This either undermines the whole concept of a species or it questions the connection between DNA and phenomenological properties, neither of which we want to do in the context of the present discussion. Once again Putnam undermines the very scientific enterprise in which he is supposed to have such confidence by calling into question one of its basic tenets - namely the belief that microstructure determines macrostructure in a law-like manner.

Finally, if we return to Putnam's 'definition' at the start of his reductio we can see that we do not need to accept it. We especially do not need to accept that lemons, or anything else, would ever be defined as that kind of thing whose normal members were formerly believed to have a particular set of properties. We can be mistaken about one, or maybe two, properties, but not all the properties of a natural kind. Putnam's requirements are too strong here. He asks us to consider what is logically possible in situations where we are dealing with natural languages and how they actually function. No-one would claim that it is a logical truth that lemons have the properties they in fact have so why do we need to consider what is logically possible here? Putnam's definition does not work because there has never been a single case of a natural kind term in a language whose meaning has been based on entirely mistaken beliefs. Nor could there be. Such a situation is unintelligible.

Putnam has not succeeded in undermining the notions of being normal or typical. In fact we could question whether he even understands these notions. We could also question the motivation for undermining these concepts, because if we cannot make sense of something being a normal member of a species, then we simply cannot make sense of a species. That there are such things as species and natural kinds is not in question in the present discussion.

TWIN EARTH

That psychological state (in the narrow sense) does not determine extension Putnam seeks to show with his famous 'Twin Earth' argument. Twin Earth is just like Earth except that the liquid called 'water' on Twin Earth is not what we call water (it is not H₂O), but a completely different chemical compound, which Putnam calls 'XYZ'. XYZ has all the phenomenological properties of water but is in fact a different substance (because it has a different chemical composition). According to Putnam the people on Twin Earth even speak English, or a dialect of English, to be more precise, since their word 'water' has a different reference from our word 'water'. Now imagine two hypothetical typical speakers, Oscar₁, who lives on Earth and speaks English; and Oscar₂, who lives on Twin Earth and speaks Twin Earth English. Suppose further that Oscar₁ and Oscar₂ are qualitatively identical to each other, including their respective psychological states. If we imagine that they lived at a time before the chemical composition of water was discovered, then, according to Putnam, Oscar₁ and Oscar₂ would *understand* the term 'water' differently, even though they were in the same (narrow) psychological state. This shows that the extension of the term 'water' is not a function of psychological state by itself. (1975a, p224) [It also shows, according to Putnam, that what is included in the intension of a term is not *sufficient* to determine the extension]

The case against the traditional theory is compounded by two more 'Twin Earth' examples. First of all we are asked to imagine that aluminium is as rare on Twin Earth as it is common on Earth, and that molybdenum is as common on Twin Earth as it is rare on Earth and that typical speakers on both planets cannot tell the difference between aluminium and molybdenum. Furthermore, the words 'aluminium' and 'molybdenum' are switched on Twin Earth so that 'aluminium' is the name of molybdenum and 'molybdenum' is the name for aluminium. Once again Putnam holds the psychological state of the speakers on both planets constant while varying the extensions of the relevant terms, thus showing that psychological state does not determine extension. When Oscar₁ and Oscar₂ use the name 'aluminium', for example, it will have a different extension in the two idiolects. If this is not enough then Putnam asks us to consider the concepts 'elm' and 'beech',

which he thinks will be the same for most people; just that of a common deciduous tree. For Putnam "...‘elm` in my idiolect has a different extension from ‘beech` in your idiolect... Is it really credible that this difference in extension is brought about by some difference in our *concepts*?" We could reinforce this point by imagining another Twin Earth example where the terms ‘elm` and ‘beech` are switched on Twin Earth. Thus, when Oscar1 says ‘elm` he means *elm*, whereas Oscar2 means *beech* when he says ‘elm`. Thus Putnam famously concludes: "Cut the pie anyway you like, ‘meanings` just ain’t in the head." (Ibid. p226)

One obvious criticism at this stage is that Putnam is begging the question against the view he claims to be criticising when he says that Oscar1 and Oscar2 *understood* the word ‘water` differently in 1750. After all, if meanings *are* in the head then they understood it in exactly the same way. This would be supported by the fact that the two speakers would use the words in the same way and behave in the same way towards the stuff they both call ‘water`. However, Putnam is not actually begging the question here. What he is trying to show is that it is a mistake to conjoin the claims:

(1) Sense determines reference

and

(2) Knowledge of sense is a narrow mental state

A less obvious but potentially damaging criticism of the Twin Earth situation is that Putnam is guilty of a ‘use/mention` confusion, especially regarding the words ‘elm` and ‘beech`. If he knows hardly anything about elms and beeches, and cannot tell them apart, then perhaps he can only mention the words and not actually use them. Elsewhere in “The Meaning of ‘Meaning`” Putnam is undoubtedly guilty of the most blatant ‘use/mention` confusion involving proper names when he says the following:

One can use the proper name 'Sanders' correctly without knowing anything about the referent except that he is called 'Sanders' - and even that may not be correct. ('Once upon a time,.....Winnie-the-Pooh lived in a forest all by himself under the name of Sanders') (1975, p247)

Now, although Putnam goes on to make it quite clear that natural kind terms do not function in this way, this passage does show that he is prone to the 'use/mention' confusion, as is Kripke with his 'Feynman' example. If he is guilty of the confusion as regards the words 'elm' and 'beech' or 'aluminium' and 'molybdenum' then obviously these examples are of no use.

Actually it seems that Putnam does not have the concept of an elm or a beech. All he knows is that there are trees called 'elms' and trees called 'beeches'. The truth is that the concepts 'elm' and 'beech' are two *different* concepts. Putnam surely knows this, so why does he claim that *his* concepts are the same? If he thinks they are the same then he simply does not *have* the concepts of an elm or a beech. (He certainly does not have the concept of an elm *as opposed to* a beech).

As for switching the names 'elm' and 'beech' on Twin Earth, what does this show? Either Oscar₂ has a different word for beeches (he calls them 'elms') which is harmless because anything can be called anything, and Oscar₂ is only speaking a *dialect* of English, remember; or Oscar₂ calls beeches 'elms' because he thinks that beeches *are* elms, in which case he is mistaken and therefore not in the same psychological state as Oscar₁. (The same argument goes for the aluminium/molybdenum case, and of course the 'water' case) Either way nothing can be concluded from this example about whether or not meanings are in the head or whether intension determines extension. Putnam's attempt to show that the properties in the intension of a term are neither necessary nor sufficient for anything to fall within the extension of the term has failed. What we must do now is examine his more positive thesis about how the reference of natural kind terms are actually fixed - the socio-linguistic hypothesis.

A SOCIO-LINGUISTIC HYPOTHESIS

Putnam defends his claim that he has the concept of an elm simply by possessing the word 'elm', together with the knowledge that elms are trees, with his socio-linguistic hypothesis - the division of linguistic labour in society. For Putnam, reference is fixed partly indexically, partly socially. The idea is that reference is fixed for the community as a whole by experts who possess the relevant identifying knowledge to be able to tell if a particular tree is an elm or if something is gold, or whatever. It is a very important part of Putnam's thesis that knowing the meaning of a word is not a matter of knowing *that* at all, but of knowing *how* - i.e. how to play your part in a system of social co-operation. According to Putnam most people do not possess the specialised knowledge enabling them to distinguish, say, aluminium from molybdenum or even gold from brass, nor do they need to, just so long as someone in the community does. As Putnam says in "The Meaning of Meaning": "He (the speaker) can rely on a special sub-class of speakers. The features that are generally thought to be present in connection with a general name - necessary and sufficient conditions for membership in the extension, ways of recognising if something is in the extension (criteria) etc. - are all present in the linguistic body considered as a collective body..." (1975a, p 228). Thus the speaker does not need to know how to identify elms or gold etc., so long as someone does.

This social dimension to reference fixing is supposed to count against the Traditional View which seems to require that, since it is by virtue of something in the mind of the speaker that words refer, the act of referring must be individual, not social. If Putnam is correct about the division of linguistic labour he will have dealt a serious blow to the Traditional View that meanings are in the head - or so he believes.

The division of linguistic labour is a departure from Kripke's theory of how general terms refer, which seems to deal more with individual acts of reference. Another major difference between Putnam and Kripke is Putnam's attempt to deal with linguistic competence concerning natural kind terms - or any terms, come to that! For Putnam, being able to refer to a natural kind is one thing; acquiring

linguistic competence with the appropriate term, another thing entirely. In order to tell someone what I mean by a natural kind term, or in order to communicate using such a term, it is necessary that I acquire what Putnam calls a **stereotype** of the kind in question. A stereotype, according to Putnam, is

“...a standardized description of features of the kind that are typical, or ‘normal’, or at any rate stereotypical. The central features of the stereotype generally are *criteria* -...ways of recognising if a thing belongs to a kind, or, at least, necessary conditions (or probabilistic necessary conditions) for membership in the kind.” (Ibid. p230)

by which he means something like ‘defeasible criteria’.

As it turns out, some stereotypes can be more detailed than others. Some, for example ‘elm’, will merely contain the fact that elms are a common deciduous tree; whereas others, like ‘tiger’, will contain a lot of information about tigers. In fact, according to Putnam, a speaker would not count as knowing the meaning of ‘tiger’ unless he knows a good deal about tigers. Speakers are *required* by the community to know something about (stereotypical) tigers in order to count as having acquired the word ‘tiger’ (p248). The same goes for all natural kind words; although for some words, ‘elm’ for example, we are not required to know as much as we are required to know about tigers, for ‘cultural reasons’ (whatever that means). Putnam even goes so far as to say that it is part of the *meaning*, in some sense, of the word ‘tiger’ that tigers are striped. It is just not *analytic* that tigers are striped (p256). On Putnam’s view: “Linguistic obligatoriness is not supposed to be an index of unrevisability or even of truth; thus we can hold that ‘tigers are striped’ is part of the meaning of ‘tiger’ without being trapped in the problems of analyticity.” (Ibid.) We may not be trapped in the problems of analyticity but nevertheless “Most stereotypes do in fact capture features possessed by paradigmatic members of the class in question” (p250).

Presumably they do not always capture all of the features all of the time, and we cannot say in advance which ones will be captured. This is not a notion of meaning that determines reference, at least in the traditional sense.

Towards the end of “The Meaning of Meaning”, Putnam attempts to ‘reconstruct’ the notion of meaning by specifying what he calls a ‘normal form’, or, rather a *type* of normal form for the description of meaning (p296). According to Putnam, if we know what a normal form description of a word should be, then we know what meaning is in any scientifically interesting sense (Ibid.) Such a description would be a finite sequence or ‘vector’, the components of which would represent a hypothesis about the individual speaker’s linguistic competence - except the extension. The normal form description for water might be:

SYNTACTIC MARKER	SEMANTIC MARKERS	STEREOTYPE	EXTENSION
<i>mass noun, concrete</i>	<i>natural kind; liquid</i>	<i>colourless;</i>	<i>H2O</i>
		<i>transparent;</i>	<i>(give or take</i>
		<i>tasteless, etc.</i>	<i>impurities)</i>

According to this schema, the representations of ‘water’ on Earth and Twin Earth would be the same, except for the extension. So Oscar1 and Oscar2 would have the same linguistic competence, but a different extension (p270). Once again we must ask in what sense they *understood* the word differently? Obviously Putnam makes no connection between linguistic competence and understanding.

On the basis of this ‘meaning vector’ Putnam concludes that meaning does determine extension - “by construction, so to speak”; although he gives up assumption (I) that meanings are in the head (Ibid.)

Exactly what kind of meaning this is is not clear. (After all, if we are not ‘trapped in the problems of analyticity’ perhaps we are not trapped in the problems of meaning either). We might agree with Searle (1983) when he claims that the socio-linguistic hypothesis does not show that intension does not determine extension because Putnam merely replaces the intension of the average speaker with the intension of the experts which determines reference. As Searle puts it in *Intentionality*:

The thesis that meaning determines reference can hardly be refuted by considering cases of speakers who don’t even know the meaning or know it

only imperfectly. Or to put it another way, the notions of intension and extension are not defined relative to idiolect. (P201)

(Bach, 1987, p289,290 makes a similar point)

This is a valid point against Putnam unless he can be more precise as to how the experts fix reference for the ordinary speaker, who lacks the reference-fixing knowledge, simply by being part of the same community. Presumably Putnam would give some kind of causal explanation of this, but as we have already seen this would be of no use since the causal theory of reference really does not work.

Putnam's Twin Earth scenario has impressed many philosophers of the externalist variety. Not surprisingly, however, there has been a robust Fregean response, which could be characterised in two ways. First of all there is the internalist view of John Searle's (1983) where he steadfastly maintains that meanings are indeed in the head by questioning whether Oscar₁ and Oscar₂ are in the same psychological state. Secondly there is the challenge to the claim that the word 'water' really does have different extensions on Earth and Twin Earth, coming notably from Eddy Zemach (1976), D.H. Mellor (1977) and Nathan Salmon (1981). In his (1977, p302-4) Mellor claims that the extension of the term 'water' on the two planets in 1750 and 1950 is the same, viz. H₂O v XYZ. This is supported by Zemach (1976) and by John Dupre who argues in his (1993) that natural kinds could encompass different microstructures, in which case the word 'water' would have H₂O and XYZ as its extension on both planets. For Dupre what counts as a natural kind will depend on our interests. This is a strong point and will be discussed in more detail in chapter 6. In the meantime it counts against Sterelny's (1983) where the claim is made that it is obviously not true that XYZ and H₂O would be regarded as the same substance following the discovery of their respective chemical structures.

The main criticism of Putnam by Zemach, Mellor and Dupre is that it is the Fregean account that accords best with our actual use of natural kind terms - and not only the use by ordinary speakers, but also of the experts. For instance, even if we accept that there is a division of linguistic labour and that there are experts who fix reference paradigmatically, we are still involved in a Fregean labour. This point is

made particularly by Mellor, who argues that even where an ostensive definition of a paradigm is involved, the paradigm is designated on the basis of certain theoretical beliefs we have. This is surely correct in the sense that even if we do fix reference using a paradigm, we can only be in a position to choose the paradigm because we are aware, or believe we are aware, of a set of properties, we take to delineate a natural kind. We have nothing else to go by. This seems to be in opposition to Putnam's claim in his (1983) that paradigms are

...given existentially and not by criteria... Actual things, whatever their description, which have played a certain causal role in our acquisition and use of terms determine what the terms refer to (1983, p73)

If Putnam means by this that properties are in no way involved in choosing our paradigms then this is just a naive claim. It also contradicts his earlier view that we rely on what he calls 'operational definitions'. An operational definition for Putnam is a way of pointing out a standard so that we can investigate its properties (1975a p232). Before discovering the microstructure or other important properties of, say, water, we can, according to Putnam, identify water by the operational definition. Now this obviously involves properties, but not, in Putnam's view, an analytical specification of what it is to be water (Ibid.). These properties would be similar to Kripke's descriptions that fix reference without giving a synonym. It is Putnam's way of "loosening up" the meanings of general terms. The problem with this has already been mentioned. Properties are all we have to go by. These properties constitute the intensions of our terms. Of course we can be mistaken about the properties of a natural kind, but then we can be mistaken about the meaning of a natural kind term.

Of course we do not need to accept the claim at all that the reference of natural kind terms is fixed by paradigms. Both Mellor (1995 p74) and Zemach (1976 p123-4) make the simple point that we do not know what sample our use stems from. For Mellor this means that "our most authoritative specimens of a kind may not even be of that kind". If this is correct then Putnam's theory appears downright fanciful - although in fairness to Putnam, he does, as we have seen, grant that the properties which go to make up a stereotype are represented by defeasible criteria. Then of

course there are undoubtedly cases where no paradigms are involved - the case of theoretical particles, or elements high in the periodic table - where only some kind of Fregean theory will do. However, these latter seem to be special cases, and so perhaps Putnam still has a point to make about the general case.

Searle (1983) answers the Twin Earth paradox in terms of the Intentional contents of mental states and questions Putnam's assertion that Oscar₁ and Oscar₂ are in the same mental state as regards the word 'water'. In Searle's view two speakers can have the same, i.e. type-identical, mental states and yet the Intentional contents of the two states can be different. This is partly because Putnam replaces the definition of water in terms of phenomenological properties with an indexical definition in terms of a hidden microstructure. If the substance is different then the Intentional content of the mental state is different. Even indexical contents are in the head, according to Searle (1983, p206, 207). So the Twin Earthers may not be in the same psychological state, in some sense; although Putnam would no doubt reply at this point that we are still dealing with *intrinsic* properties of mental states to refer, and hence still involved with magical theories of reference. He would also deny that he is *defining* the term 'water' at all - at least in the traditional sense.

Putnam also takes issue with the claim that the intentional contents of the mental states of Oscar₁ and Oscar₂ are different. In the introduction to Pessin and Goldberg (1996) Putnam argues that if Searle were correct on this point then the intentional contents of speakers of different languages would be different even when using words that are translatable from one language to the other. For example, on Searle's theory, according to Putnam, what an English speaker means by the word 'elm' would be *the kind of tree that experts upon whom I rely refer to by the name 'elm'*, whereas what German speakers mean by the word 'Ulme' is *the kind of tree that experts upon whom I rely refer to by the name 'Ulme'* I.e. we seem to have two different intentional contents and yet we are dealing with the same extension and two words that are inter-translatable! However, this is not a strong point against Searle, precisely because the terms 'elm' and 'ulme' are intertranslatable. Given that it is experts who determine reference on Putnam's view, we could ask him how correctness of translation is established in such a case. At any rate, Searle is surely

correct when he claims that Putnam replaces the intension of the ordinary speaker with that of the expert, which severely weakens Putnam's case.

We can also question Putnam's division of linguistic labour and the way he wants to separate reference from linguistic competence (1975a, p246). On the one hand, according to Putnam, if gold, for example, is important to a speaker he has to acquire the word 'gold' (1975a, p227), (which is not true for a start); but he does not need to acquire a way of recognising gold, and he does not acquire a way of fixing the reference of 'gold'. (We might wonder why gold is so important to this person if he does not know how to recognise it). This is done for the community as a whole by the experts. On the other hand the speaker is obliged to acquire a *stereotype* of gold, or whatever, which does provide 'criteria', or probabilistic necessary conditions for anything to be gold. Without this stereotype the speaker cannot achieve linguistic competence with the word. In fact, according to Putnam, (p247) the speaker cannot be counted as having *acquired the word at all*, unless he acquires a stereotype. If all he has is the word and perhaps the knowledge that it has an extension, he can only be regarded as having *partially acquired* the word. (How can anyone partially acquire a word? This seems to be another case of DRT confusing a word with the concept that goes with it). So the speaker does not need a way of recognising gold but he does need to have a way of probably recognising gold, viz. the stereotype. Where stereotypes come from Putnam does not say. Are they the same things as operational definitions? Are they passed from speaker to speaker in the causal chain? Do they originate with the experts; and if not, what exactly is the role of the experts in the semantics of general terms?

Actually it seems that if there are such things as stereotypes, then they will ultimately come from these experts. They surely cannot contain information that would take precedence over the knowledge of the experts. Although we may get by for a while with what we might term 'folk stereotypes', ultimately we will rely on verbal and ostensive definitions in dictionaries and encyclopaedias - written by the experts, or with the help of experts. This is where we acquire linguistic competence. This is where the stereotypes are to be located.

As for the ‘reconstruction’ of meaning in the form of a meaning vector, this would seem to be a rather trivial analysis of the determination of reference, since the extension of the term is contained within the vector itself (or, rather, a description of the extension). On page 246 of “The Meaning of Meaning” Putnam anticipates the meaning vector in principle but not in great detail, and admits that an analysis of meaning in terms of the vector would make it trivially true that meaning determines extension, in the sense that difference in extension is *ipso facto* difference in meaning. We shall see below that a non-trivial analysis of meaning would require that **knowing** the meaning of a term, or acquiring linguistic competence with a term, would determine reference for speakers of the language. Putnam shies away from such an analysis for fear of having to locate meanings in the head. I shall argue below that there is an analysis of meaning where meanings are in the head and determine reference, but not in any ‘magical’ way.

BRAINS IN A VAT

Putnam argues for the necessity of a causal connection between language and the world in his treatment of the famous ‘Brain-in-a-vat’ problem. In chapter 1 of *Reason, Truth and History*, Putnam considers the possibility that we could be disembodied brains kept alive in a vat of nutrients and programmed by a computer to have the illusion that we are solid flesh and blood beings living in a world of physical objects, which we can see, touch, etc. and, crucially, refer to using a language. In short, the question is: Could all the experiences we regard as ‘real’ actually be computer-generated illusions? Could we be brains in a vat?

The problem is usually posed as an epistemological one; viz. “How do we *know* that we are not in this situation?” Putnam seeks to show (1981, p7) that the supposition that we are actually brains in a vat cannot possibly be true because it is self-refuting, or necessarily false (Ibid. P15). This he attempts to show by exploring the preconditions for reference, from which he concludes that reference is conditional upon causal connection to the world. It follows from this, once again,

that there is nothing intrinsic to any mental representation which can reach out to the world and refer to it.

Putnam begins by asking if there could be a definitive Turing Test for reference. Suppose, for instance, we have a machine that has been programmed to answer questions put to it by a human being. It can produce discourse about things with which we are familiar and so might seem to be referring to these things. However, this ability to refer to actual things is an illusion created by *our* ability to refer to such things. The machine may seem to be referring but is not actually doing so because it does not have the kind of relation to actual objects that we do. Specifically it lacks what Putnam calls 'language entry rules', which take us from experiences of actual things in the world to such utterances as "I see a such-and-such", and 'language exit rules', which take us from decisions expressed in linguistic form ("I desire a such-and-such", for example) to extra-linguistic acts (Ibid. p11). There is therefore no definitive Turing Test for reference because there is always the possibility that someone could pass the Turing Test for reference and still not be referring to anything. (I shall argue below that there is a definitive Turing test for sense)

This would be the situation if we imagine some brains in a vat. Although the 'vat-people', as I shall call them, could say things like "I see a tree", they would not be talking about actual trees, although someone might take them to be so doing. If they are referring at all, they are referring to computer-generated illusions or simulations of trees - and the same goes for all of their words. None of the vat-people's words ever refer to anything outside the vat because they lack the necessary causal connection to those things. It follows from this that the vat-people could not say or think that they were brains in a vat because, if they uttered the sentence "We are brains in a vat", it would be automatically false, since the words 'brain' and 'vat', as uttered by the vat-people, would not refer to an actual brain or an actual vat. The sentence "We are brains in a vat" as uttered by the vat-people, would actually mean "We are computer-generated illusions of brains in a computer-generated illusion of a vat", which is false. In short, the brain-in-a-vat hypothesis is self-refuting in the sense that, if we are brains in a vat, then "We are brains in a vat" is false. So it is necessarily false. Just to entertain the hypothesis is to imply its falsity.

Notice here that the word 'we' is translated homophonically, from "We are brains in a vat" to "We are computer-generated illusions of brains in a computer-generated illusion of a vat". This is because 'we', or 'I', are the only words in the 'vat - language' that actually refer to anything. The word 'I' in the sentence "I am a brain in a vat" refers to the actual brain in the vat, whereas all the other 'words' either only refer to illusions of things, or, to nothing. We could conclude that the words of the vat-language do not refer at all since there is nothing for them to refer to. They do not refer to anything outside the vat and there is nothing inside the vat for them to refer to. The irony here is that the brains would never realise that the word 'I' ('we') refers to a brain. They think they know what all their other 'words' refer to (or most of them) and could be embroiled in a debate about what the word 'I' refers to. Does it refer to a Cartesian self, or is it just an imposition of grammar? They'll never know.

The other point to note here is that Putnam is wrong in claiming that a sentence like "I am a brain in a vat" is necessarily false. If it is false it is not necessarily so. Actually it may not be false at all in the sense that we cannot know that it is. I.e. on Putnam's view the sentence "I am a brain in a vat" is false as uttered by anyone who can entertain the thought it expresses. But since we could never know that we are entertaining the thought, there is no refutation of scepticism here - which may mean that if causal connection to the world is necessary for reference, then we simply don't know what we are talking about!

What, then, does Putnam's attack on the Traditional View amount to, and how successful has it been? We have seen that it has been limited in its success. Putnam seems to beg the question against the Traditional View when he claims that Oscar₁ and Oscar₂ understand the word 'water' differently in 1750. There may even be a 'use/mention' confusion in his thinking concerning the words 'elm' and 'beech', as there seems to be throughout DRT in general (For example, Kripke's 'Feynman' example, and Evans' 'Louis' example). We might also agree with Searle's claim that the two Oscars are not in the same psychological state and that Putnam merely replaces the intension of the ordinary speaker with that of the expert. Certainly there seems to be a strong case put forward by Zemach, Mellor and Dupre that the

extension of the word 'water' in 1750, as now, includes XYZ and H₂O. It may even be that the Twin Earth situation is incommensurable with the situation on Earth because the Twin Earthers do not have our word 'water' but merely a word that has "the same spelling and pronunciation" (1975a, p248). After all, consider what Putnam is saying here. He wants to argue against the claim that it is by virtue of a sense in the minds of speakers that the word 'water' refers. He does this by considering a possible, though highly unlikely, situation where the people do not possess our word 'water' - neither do they have a name *for* water ; because there isn't any water in this situation! What can we conclude from this?

Putnam is quite clear (1975a, p247) that as far the word 'tiger' is concerned the Twin Earth people only count as having the word 'tiger' if it has as extension the set of tigers. For Putnam, "If Twin Earth organisms have a silicon chemistry, for example, then their 'tigers' aren't really tigers, even if they look like tigers....Thus....we have decided to say that Twin Earth speakers have not acquired our word 'tiger' (although they have acquired another word with the same spelling and pronunciation) [Ibid]. Surely the same applies to the 'water' case. The Twin Earthers simply have a different term, and so the two situations cannot be compared. We are not in the same situation whereby a word is translated from, say, English into German. In such a situation the word 'water' can be translated as 'wasser' because the two words have the same extension. In the Twin Earth scenario the earth word 'water' and the Twin Earth word 'water' have different extensions (according to Putnam) and so we are not dealing with a straightforward translation of one language into another. The charge of incommensurability stands.

The charge of incommensurability may be even stronger in the Brain-in-a-vat case. That is, we cannot conclude anything about what the brains can or cannot refer to because they do not actually refer. They only have the illusion of referring. The very word 'refer' has a different meaning inside the vat (if it has any meaning at all) and so the situation inside the vat is incommensurable with the situation outside the vat. Similarly, if we take a word like 'brain' as used inside the vat, we cannot conclude anything about what it can or cannot refer to because the brains in the vat do not have the word 'brain'. They certainly do not have *our* word (or term) 'brain' since it does not have the set of brains as its extension. In fact we could say that they

do not have words at all - only the illusion of words. Nevertheless, as we have seen, we would seem to *understand* the brains, even though they do not refer to anything. At any rate, we would not be able to tell whether we were listening to brains in a vat or real people referring to actual objects. In short, there is a Turing Test for sense - i.e. sense as understanding. This shows either that sense is independent of reference in some sense, or that reference is inscrutable.

This brings us to the crux of the matter. Just what exactly is the relation between sense and reference and between what is in the minds of speakers and what their words refer to? One thing that is certain and that is that sense is **necessary** for reference. It is necessary because reference only takes place within a language, and there can only be language if there is understanding - which means speakers grasping senses in their minds. I.e. a word like 'water' or 'gold', or whatever, only refers if it contributes to what speakers of the language understand by these words (in any sentence which uses them). *Understanding is the single most important concept in the philosophy of language because without understanding, or the possibility of understanding, there is no language. DRT fails as a semantic theory because it either deals inadequately with understanding, as Putnam does; or not at all, as in Kripke's work.*

So understanding is necessary for reference and understanding consists in speakers grasping senses in their minds. Sense, therefore, is necessary for reference. Furthermore, something in the minds of speakers is necessary for reference because understanding is a mental act. Even Putnam might accept this (at least concerning general terms) - although Kripke probably would not. What Putnam tries to do is to show that sense is not *sufficient* for reference. We may reject Putnam's arguments but nevertheless it will be more difficult to argue for the sufficiency of sense than for its necessity.

Nevertheless, the necessity of sense in determining reference is crucial to the debate over the existence of necessary a posteriori truths. It means that even if microstructure is necessary for something to be a member of a natural kind, then so is macrostructure, because it is the macrostructure, i.e. phenomenological properties of the kind, that constitute the sense of the natural kind term. For example, if it is necessary that water is H₂O, it will be just as necessary that it have the

phenomenological properties it has (or seems to have) in the actual world. Of course we must remember that we are dealing with a qualified notion of sense. For instance, the sense of the word 'water' is not

Naturally occurring substance with properties $P_1 \dots P_n$

but, rather:

Naturally occurring substance *that is commonly believed* to have properties $P_1 \dots P_n$

If this is our notion of sense then it makes a great difference to what we regard as a priori, analytic and necessary. Take, for example, the sentence

(a) Lemons have $P_1 \dots P_n$

If 'lemon' has the sense of

That which is commonly believed to have $P_1 \dots P_n$

then (a) will mean

That which is commonly believed to have $P_1 \dots P_n$ have $P_1 \dots P_n$

Clearly this may not even be true, so the question of its a priority, analyticity or necessity will simply not arise (Although some may claim that it is contingent a priori). However, these questions do seem to arise with

(b) Lemons are commonly believed to have $P_1 \dots P_n$

since, on the present view, this will mean

That which is commonly believed to have P1...Pn are commonly believed to have P1...Pn

The charge could be made here that (b) appears to be analytic but not a priori or necessary; although it is doubtful whether someone like Kripke could make such a charge. After all, if he wants to claim that there are necessary truths which are not analytic, then why can he not accept that there could be analytic truths that are not necessary? Certainly (b) is not necessary de re. It is of course necessary de dicto, and there is also a sense in which it is a priori, once we know what the terms mean.

One thing we know for sure is that sense is necessary for reference. What we must do now is analyse why it does not seem to be *sufficient* for reference. This will lead us to examine just what is meant by reference. First of all, we have seen that Putnam has not succeeded in showing that sense is not sufficient for reference with his Twin Earth experiment, especially if we accept that the extension of 'water' is, or could be, H₂O v XYZ. Even if we do not take this line it could be argued that sense determines reference in general, although there may be highly specialised cases, such as the Twin Earth case, where the relation comes under strain. But if this is so, then the Twin Earth example also presents problems for Putnam's view of reference fixing by experts. This is because the experts can be wrong, and also because what counts as an expert will change over time. For instance, in the example of water and Twin water, the sense of the earth term 'water' will be something like:

Stuff which is commonly believed to have such-and-such properties

In 1750 this will have only H₂O as its extension because speakers on Earth have no beliefs about Twin Earth. They don't even know that it exists. However, the liquid on Twin Earth that resembles water would be part of the extension of 'water' if this is determined by experts on Earth because in 1750 what counts as an expert on Earth would judge that XYZ is water. Once again it is the (qualified) Fregean theory that works best - at least before the discovery of chemistry. Once the microstructure of the two substances is discovered (if they are two substances) then the Fregean theory

will come under strain, at least as far as common usage is concerned. As far as the scientific usage is concerned, however, the sense of 'water' will be adjusted to accommodate the new information that distinguishes H₂O from XYZ. This may eventually become part of common usage or it may not. At any rate, if there is a problem with the Fregean determination of reference, there is also a problem with Putnam's view. It could be that we simply cannot talk about reference. At the very least we should be very careful about reference talk.

This last point is supported by the argument made above that sense is necessary for reference, combined with the fact that sense is very often indeterminate. In fact, strictly speaking, sense is *always* indeterminate in that we can never really know if we are in possession of the correct sense of a word. Hence the need to qualify our senses in terms of speakers' beliefs. *What this amounts to is that if sense is necessary for reference and sense is indeterminate, then reference is indeterminate. To put it bluntly, we literally don't know what we are talking about. If this is all there is to be said about reference then questions about essentialism derived from a semantic theory would hardly arise!*

There may be more to be said about the relation between sense and reference, however. So far we have spoken as if reference is something that is effected by language. This is true, but it is equally true that there is no language without speakers, or potential speakers, to understand the language. We must therefore examine the relation between speakers and the mechanism by which their words refer.

What I propose is that it is really speakers who refer when they communicate with other language users. I.e. speakers refer **each other** to what they take to be things in the world. This they do by evoking concepts in the minds of other speakers causing them to think about whatever is being referred to. For example, if I say to you:

Kripke has just published a new book

I am referring *you* to a certain person who has just done something. The sound of these words evoke certain concepts in your mind thus enabling you to understand what I am saying. If these words did not evoke these concepts in your mind you

would not know what or who I was talking about - or referring to. If no-one in the linguistic community was affected by these words in the appropriate way, or in any way, then this would not be a sentence in our language. It would be nonsense. It does not matter how we characterise this understanding for the present purpose. As far as the determination of reference is concerned understanding can be taken for granted. We can start with the assumption that language exists - which means that understanding, or the possibility of understanding, exists. So it is speakers who refer by referring other speakers by causing them to have certain thoughts. The contents of these thoughts are conceptual. Sometimes these concepts correspond to things in the world and sometimes they do not. This is a matter of empirical investigation.

This latter point answers Kripke's circularity objection mentioned at the end of chapter one where he claims that the sense of a name cannot be given in terms of what is commonly believed by the community because we must say what it is that is commonly believed without circularity. This is easily done. The content of the thoughts expressed by sentences containing the relevant name is given in terms of concepts in the minds of speakers, which may turn out to be concepts of actual things and may not be.

A few points of clarification are necessary at this point. First of all, it could be argued that this is merely Locke's largely discredited theory (I.e discredited by Wittgenstein's Private Language argument). It is true that the above theory is similar to Locke's but there are important differences, the main one being that, on Locke's theory words stand for ideas in people's minds; whereas, on the present view, words can quite legitimately be regarded as 'standing for' things in the world *by means of* evoking concepts in the minds of other speakers - or by causing them to grasp senses in their minds.

Secondly, it could be argued that surely a sentence such as

Kripke has just published a new book

will refer to Kripke and books etc. Even if it is never communicated to anyone. For example, I could simply write it down in a diary and never show it to anyone. This

is indeed the case. The word 'Kripke' will refer to Kripke even if the sentence is never actually communicated to anyone - but only in the sense that it has the *potential* to cause other speakers to have certain thoughts, or to grasp certain senses.

[The Private Language objection will be dealt with below]

The third point of clarification is an important one. It is that, although I am proposing that it is speakers who refer, I am not making the distinction between speaker's reference and semantic reference drawn by Kripke in his (1977). On the current view there is no such distinction.

Kripke begins his 1977 paper by examining Donnellan's distinction between the referential and attributive uses of definite descriptions in order to see whether the distinction really does count against Russell's treatment of descriptions in his Theory of Descriptions. Although Kripke decides that Donnellan's distinction probably does not count against Russell he nevertheless accepts that it is a genuine distinction and that it can be used to illustrate a difference between what Kripke terms 'speaker's reference' and 'semantic reference'. The background to the distinction is Grice's work on implicature - i.e saying one thing but meaning something else. The classic example is the situation where one burglar says to another, "The cops have arrived", but actually means something else, such as "Let's get out of here!". In a similar fashion Donnellan cites the now famous example of the man in the dock accused of the brutal murder of Smith when he is in fact innocent of the crime. According to Donnellan, someone present in the courtroom could refer to this man using the description "Smith's murderer", as in "Smith's murderer is insane", even though the man in the dock does not fit the description. This is the referential use of a definite description.

Kripke proposes his own example involving proper names rather than descriptions which is supposed to illustrate the distinction between speaker's reference and semantic reference. Suppose, says Kripke (1977, p257) two people see Smith in the distance raking leaves and mistake him for Jones. They have a brief colloquy: "What is Jones doing?" "Raking leaves". "Jones" in the common language of both is a name of Jones; it *never* names Smith. Yet, in some sense, on this occasion, clearly both participants in the dialogue have referred to Smith, and the second participant has said something true about the man he referred to if and only

if Smith was raking leaves (whether or not Jones was). Thus, for Kripke, the semantic referent of a designator is given by certain conventions of the speaker's idiolect (given various facts about the world). The speaker's referent on the other hand, will be that object which the speaker wishes to talk about, on a given occasion, and believes fulfils the conditions for being the semantic referent (Ibid.) In the above example, Jones is the semantic referent, while Smith is the speaker's referent.

In what follows I shall argue that what Kripke calls semantic reference is actually speaker's reference, and what he calls speaker's reference is not really reference at all. To see this, consider again the example of Smith raking leaves and being mistaken for Jones by two speakers, call them A and B. Now Kripke (and probably Donnellan) thinks it obvious that when A says to B:

(i) Jones is raking leaves

that A is referring to Smith (on this occasion) and has said something true about Smith. But there is reason to doubt this, because if A *reports* what he believes he has witnessed, he will say something false, as in:

(ii) Today I saw Jones raking leaves

or

(iii) Jones was raking leaves in the park

Even if he reports in the present tense, as in

(iv) I am watching Jones raking leaves

what he is saying is false. So if (ii), (iii) and (iv) are false, how can (i) be true? And if (i) is not true, how can A be referring to the man raking leaves? Of course A *believes* that he has referred to Jones and that he has said something true about

Jones, but he has failed to do so. I.e. he has failed to utter a truth and he has failed to refer to the person he intended to refer to. Seen in this way there is no such phenomenon as speaker's reference in Kripke's sense.

Another way of looking at the problem is to consider what is understood by the name 'Jones'. Suppose it has a sense something like:

The man commonly believed to be the Dean of the Faculty

then A is really saying to B

The man who is commonly believed to be Dean of the Faculty is raking leaves
which is, again, just false.

As far as the so-called 'Gricean Implicature' is concerned, such a phenomenon need not trouble us. When the burglar says "The cops have arrived", if he means something else by this then he is not speaking English. At least he is not speaking literally. He is speaking figuratively, or euphemistically - which is fine, so long as other people are privy to the fact that the literal use is not intended. If other speakers are not privy to this then the burglar will be talking to himself. The non-literal use of expressions in a language such as English is really a separate language because a literal speaker of English would misunderstand such a language unless someone lets him in on the secret - for example, that "The cops have arrived" really means "Let's get out of here", or "Let's give ourselves up", or whatever. Thus the expression "The cops have arrived" is really being used as a kind of code, or perhaps a euphemism. This is fine but it is not a literal use of English and so need not trouble any general theory of the semantics of English.

Reference consists in speakers referring each other to what they take to be things in the world. This requires speakers to grasp senses in their minds. In this sense meanings are in the head and sense does determine reference, since it determines what the speakers are talking about.

Another way of looking at the problem is as follows. Suppose that 'NN' is a name for something; then its correct use in a language will be necessary for it to refer.

This correct use involves speakers grasping senses in their minds. But not only is correct use necessary for reference, it is surely sufficient. If I use a name correctly then I will be referring to its referent. So not only can we say that meaning is use, we can also say that reference is use - where use involves understanding - i.e. speakers grasping senses in their minds - or the potential for this to happen. This also counts against any causal theory of reference since correct use does not seem to presuppose any causal connection to a mind-independent external world.

There are still problems, however. The main problem is where senses come from. They are surely not innate as Chomsky and Fodor claim, and we must say more than Frege when he says that senses are grasped by anyone who is familiar with the language. It is at this point that DRT seems to have the advantage over the Traditional View in its insistence on the contribution from the environment because if senses are not innate then where else could they come from? And if the meanings of our terms do depend on some kind of causal connection with the environment then meanings would definitely not be in the head. This is what Putnam tries to establish with his Brain-in-a-vat argument. We have seen that this does not work in the way Putnam wants, and also that it can be turned on its head to show the opposite of what he claims. What also counts against the claim that it is the external world that determines the reference of our terms is quite simply that there may not be a mind-independent external world. In other words, meanings could be in the head for the simple reason that *what we regard as the external world could be in the head*. DRT and the scientific realism that underlies it is guilty of making the same unjustified assumptions about reality that metaphysical realists have been accused of making since at least the time of Locke and Berkeley. What Locke and Berkeley did agree on is that the meanings of our terms come from our experience. Whether this is experience of a mind-independent reality is debatable, but it seems clear that the senses of our words are indeed derived from experience. We experience, or believe we experience, such things as water, gold and tigers and from this experience we invent names for these things and associate senses with the names, in our minds. These senses (suitably qualified in terms of what the linguistic community believes) then determine reference for the speakers of the language; where reference is construed as consisting in speakers causing other speakers to have certain thoughts,

which may be thoughts about an external mind-independent reality, or may not be. That is a metaphysical claim beyond the scope of semantics. The Private Language objection can be answered at this point because the senses of our words are derived from experience of publicly observable things, or, in the case of mental terms, publicly observable behaviour. So meaning is public because senses are derived from what we take to be a publicly observable world.

Another possible difficulty with the present theory is that it would seem to be difficult to say whether the beliefs of speakers which constitute the senses of terms are *de re* or *de dicto*. It could be that they are *de dicto*, in the first instance, and possibly *de re* on further investigation. Or it could simply be that the *de re/de dicto* distinction becomes blurred here - if not untenable. What is clear is that speakers do have beliefs and it is these beliefs that constitute the senses of their terms.

In summary, then, sense is at the very least necessary for reference. This in itself is significant for any study of the supposed necessary a posteriori statements. Putnam and Kripke have failed to show that sense is not sufficient for reference, especially if we deal with the qualified notion of sense outlined above. Sense is sufficient for reference if we take reference to be the capacity for a word to be used and understood by speakers, which involves them referring each other to what they take to be the world. Of course we can be mistaken about senses, but once the mistake has been discovered the meaning of the word changes. Not only does sense determine reference (in a sense) but meanings are in the head, in the sense that it is something in the minds of speakers which determines what they are referring to - but not in any magical way. I.e. we are not dealing with narrow states.

Even if Kripke and Putnam do not accept the sufficiency of sense in determining reference, they cannot deny its necessity. If they wish to argue that sense is indeterminate in some way then they must accept that reference is indeterminate because sense is a necessary condition for reference. Ironically, when Putnam tries to loosen up the criteria by which reference is determined he does not give us a better way of determining reference, but merely makes it more difficult, if not impossible, to talk about reference at all. But this is something we must live with. It is a consequence of the limits of our knowledge of the world.

As far as the claims about necessary a posteriori truths, or whether such statements as 'Water is H₂O' are analytic are concerned, we shall deal with these in chapter 6. In the meantime, however, we should remember the kind of language we are dealing with. We are dealing with what we could term a 'Scientific Language'; not the language of the man in the street. In the scientific language, strictly speaking, we should not make such statements as 'Water is H₂O' but only such statements as

Water appears to be H₂O

or,

The substance that is commonly believed to have properties P₁...P_n appears to have chemical composition H₂O

Seen in this way, questions of necessity, a priority and analyticity hardly seem appropriate. Ironically, such questions only arise for language as used by the man in the street and not for the scientist. Such use is, strictly speaking, incorrect. It is not the language of science.

One final point about whether names have sense which determine their reference and whether we need to possess senses in order to refer is that, strictly speaking, we do not need proper names, or even natural kind terms like 'water' and 'gold' at all. We could simply use descriptions instead. We could get rid of all proper names, and most natural kind terms and replace them with descriptions - not in the abstract and artificial way advocated by Quine, but just as we sometimes do in ordinary speech when we forget a name. We cannot do things the other way round. We cannot get rid of descriptions and simply rely on 'mere tags' because there would be no understanding - and hence no language. Of course I am not suggesting that for ordinary use we should do this; but for the scientific language we could and indeed should. There is certainly nothing to prevent it logically. We would still need general terms such as 'liquid' and 'metal', but we do not need names like 'water' and 'gold'. These 'mere tags' are just that without a descriptive sense which allows

their participation in the language. If we accept that we do not really need proper names or the type of natural kind terms dealt with by Kripke and Putnam then DRT is defeated before it starts.

CHAPTER FOUR

IDENTITY

In “Identity and Necessity” and *Naming and Necessity* Kripke makes many bold claims about the existence of necessary a posteriori truths. This is a consequence of his thesis that proper names and even names for natural kinds are rigid designators, in conjunction with his claim that identity statements involving rigid designators are necessary. This latter is derived from the metaphysical thesis that any pair of objects, x and y, if they are identical, then they are necessarily identical.

Kripke’s alleged examples of the necessary a posteriori are by now well known. They include such statements as:

- (i) Hesperus is Phosphorus
- (ii) Water is H₂O
- (iii) Nixon is the offspring of gametes G

This claim to the necessity of identity is explicated in terms of possible worlds. For example, if Hesperus is Phosphorus in the actual world and ‘Hesperus’ and ‘Phosphorus’ are rigid designators then there will be no possible worlds where the planet Venus exists and Hesperus is not Phosphorus since the names ‘Hesperus’ and ‘Phosphorus’ designate the same object in all worlds where the relevant object exists. This is what Kripke calls ‘weak necessity’ or metaphysical necessity. Such claims were originally regarded as outrageous by many philosophers who maintained that an empirical discovery must be contingent. Kripke did not falter in the face of such opposition but insisted that some empirical discoveries are indeed necessary, because identity statements are necessary.

In the present chapter we will be concerned with this claim to the necessity of identity which has become the orthodox view in philosophy. Few philosophers would now argue with Kripke’s claim that identity statements are necessary. (One who does is Ian Hacking whose views we shall consider later). Not all would agree with Kripke that statements such as (i) - (iii) are examples of non-trivial essentialism

(E.g. Salmon, 1981). Not all would even grant that they are a posteriori (Fitch, 1976 and Salmon, 1986, as we have seen, argue that statements like (i), at least, are a priori); and not all would concur with the Kripkean view that the necessity of such statements has anything to do with the rigidity of names or that their modality can be explicated in terms of possible worlds (E.g. Wiggins, 1976). On the necessity of identity, however, there is broad agreement. The initial shock has subsided and faded into an admiring acquiescence. The case is closed.

In what follows I shall present a simple, and I believe novel, argument, which although not entirely conclusive, amounts to at least *prima facie* evidence for the re-opening of the debate, and at most to a serious challenge which undermines the claim that identity statements are necessary. If this challenge is successful then serious damage will have been inflicted on the claim that ordinary proper names or names for natural kinds are rigid designators, and even on the claim that expressions of the form $a=b$ make genuine statements at all. This would vindicate Wittgenstein's position in the *Tractatus* where he describes identity statements as "expedients in presentation" (4.242, 5.535). Of course if statements of the form $a=b$ do not make genuine statements we must ask how it is that we seem to understand them. This question will also be dealt with in the present chapter. Before proceeding to the question of the modality of identity statements, however, we must be clear as to what sort of thing we are supposed to be dealing with. In other words, what does a so-called 'identity statement' actually state?

To simplify matters we shall deal initially with identity statements involving ordinary proper names rather than general terms - statements like:

'Hesperus is Phosphorus' and 'Cicero is Tully'

These are examples of what is usually termed **synchronic** identity, or *identity at the same time*; to distinguish it from **diachronic** identity, or *identity over time*, or through time, as in the case of personal identity. Later in the chapter we will question this distinction between synchronic and diachronic identity, but for the time being we shall let it stand.

Within the category of synchronic identity there are traditionally two opposing theories concerning the meaning or content of identity statements. These are the **objectual analysis**, which holds that it is the *objects* x and y, or a and b, that are identical; and the **meta-linguistic analysis**, which says that an identity statement is about the co-referentiality of linguistic items. For example, according to the meta-linguistic analysis the statement “Hesperus is Phosphorus” just means, or has as content, that the *names* ‘Hesperus’ and ‘Phosphorus’ are co-referential. In the light of Frege’s work on identity the meta-linguistic analysis could be divided into a simple analysis, like that found in *The Begriffsschrift* (1879); and a sophisticated analysis, such as the one posited in Frege’s (1892) where he introduces the notion of sense as a mediator of reference in order to solve the problem about the difference in cognitive value between statements of the form $a=a$, and those of the form $a=b$. On the other hand, Frege’s later view may actually be a form of the objectual analysis (This is how it is usually taken).

Clearly the objectual and the meta-linguistic analyses are two very different analyses of identity statements which will have different consequences for the modality of such statements. Although Frege himself was not concerned with modality it is obvious that any meta-linguistic analysis of identity statements has the consequence that they are contingent since it is a contingent matter whether an object is known by two names or referred to in two different ways (This is one of the reasons that Frege abandoned the *Begriffsschrift* account of identity). The objectual analysis, on the other hand, would seem to have the consequence that true identity statements are necessary. Although Kripke does not say what an identity statement actually means it is clear that he must opt for the objectual analysis. This is supported by what he says on page 3 of (1980):

....identity is an ‘internal’ relation: $(x) (y) (x=y) \rightarrow \Box (x=y)$If ‘a’ and ‘b’ are rigid designators, it follows that ‘ $a=b$ ’, if true, is a necessary truth. If ‘a’ and ‘b’ are *not* rigid designators, no such conclusion follows about the statement ‘ $a=b$ ’ (though the *objects* designated by ‘a’ and ‘b’ will be necessarily identical)

And again on page 4:

We must distinguish three distinct theses: (i) that identical objects are necessarily identical; (ii) that true identity statements between rigid designators are necessary; (iii) that identity statements between what we call 'names' in actual language are necessary. (i) and (ii) are (self-evident) theses of philosophical logic independent of natural language. They are related to each other though (i) is about objects and (ii) is meta-linguistic.

By 'meta-linguistic' of course Kripke does not mean what we have termed the meta-linguistic analysis of identity statements. What is clear from the above quotes is that if we accept that identity statements between ordinary names are necessary then we need the objectual analysis of identity statements to be the true analysis.

The main problem with the objectual analysis is how to account for the difference in cognitive value between trivial a priori statements of the form $a=a$, and informative, empirically discoverable truths of the form $a=b$. This is just the sort of problem that the meta-linguistic analysis deals with so well. The problem for the objectual analysis is that sentences of the form $a=a$ and $a=b$ seem to express the same *proposition*; in which case how can one be a priori and necessary while the other is a posteriori and contingent? Actually it is rather remarkable that the fundamental difference in the modal and epistemic status of these two types of statement, together with their potential to affect behaviour, or not as the case may be, has not been taken as evidence for the fact that statements of the form $a=a$ and $a=b$ have absolutely nothing to do with one another. They do not express the same proposition. They do not say the same thing; nor do they deal in the same properties or relations *because* one is a priori, trivial and necessary, while the other is a posteriori, contingent, and has the potential to profoundly affect behaviour. The objectual analysis either ignores this considerable body of evidence or it has to invent some rather tortuous and profoundly counter-intuitive ways of accounting for the modal and epistemic differences.

Attempted solutions of the problem of the informativeness of identity statements take many forms within the objectual analysis. Some we have already encountered, such as Salmon's view that statements of the form $a=b$ are actually a priori, since they express the same proposition as those of the form $a=a$; or Kripke's claim that we must distinguish between the *epistemological* claim that a statement is a priori or a posteriori, and the *metaphysical* claim that it is necessary or contingent. Since we

have already seen that Salmon's views are incoherent and since Kripke's claims involve the claim to the necessity of identity - the subject of this chapter - we shall leave them for the time being.

In his book on identity statements Thomas Morris (1984) offers two more arguments which are supposed to reconcile the informativeness of identity statements with the objectual analysis. These come from Plantinga (1974) and Stalnaker (1972, 1976) and are typical of the response put forward by advocates of the objectual analysis. Basically Plantinga and Stalnaker seek to explain the informativeness of a statement like "Hesperus is Phosphorus" by arguing that when we discover the identity we are discovering the contingent fact that this sentence expresses the necessary proposition that Hesperus is Hesperus. This is supposed to explain how the discovery that Hesperus is Phosphorus can be empirical when the core content of the sentence "Hesperus is Phosphorus" is a necessary proposition. The empirical discovery is that a certain sentence expresses a necessary proposition. Actually Stalnaker goes further, suggesting that there are two propositions involved in our understanding of a statement like "Hesperus is Phosphorus" (1976, p88-89). First of all there is the contingent proposition that the sentence "Hesperus is Phosphorus" expresses the proposition that Venus is self-identical, and secondly, there is the necessary proposition about self-identity that it does express.

The most obvious objection to this line of argument is that it makes empirical discoveries like the discovery that Hesperus is Phosphorus into linguistic discoveries - discoveries about what propositions are expressed by certain sentences - and this surely cannot be correct. When we discover that Hesperus is Phosphorus we surely make a discovery about extra-linguistic reality, not about language. Furthermore, if the likes of Plantinga are to be believed, then it seems that we can only discover that Hesperus is Phosphorus if we first have the appropriate sentence in our vocabulary. This would rule out an accidental discovery of this type of identity. We can surely discover that Hesperus is Phosphorus even if no-one has ever formulated the appropriate sentence or ever suggested or suspected that Hesperus might be Phosphorus. The theories of Plantinga and Stalnaker also lead to the conclusion that we can understand sentences without knowing what propositions

they express. If propositions do not contribute to understanding then what exactly are they supposed to contribute to?

In his (1984) Morris highlights an even more fundamental flaw in Plantinga's argument, and that is that it leads to an infinite regress. On page 10 of (1984) Morris characterises Plantinga's argument in the following way (where S stands for the sentence "Hesperus is Phosphorus", and (1) for the sentence "Hesperus is Hesperus"):

- (i) S expresses some proposition P
- (ii) (1) expresses some proposition P1
- (iii) P 'is the same as' P1
- (iv) (iii) was discovered to be true

According to Plantinga, the ancient astronomers discovered (iii), that P is P1. Now, Morris argues, if this is regarded as a *sentence* in our language, then it expresses a proposition, the same proposition as the sentence 'P1 is P1'. And so we have the same problem again of how 'P is P1' can be informative when 'P1 is P1' is not - *ad infinitum*; and so Plantinga's analysis fails.

As for Stalnaker's claim that we are dealing with two propositions, one contingent, the other necessary, this fares no better according to Morris. On page 21 of his (1984) Morris highlights an admission by Stalnaker to the effect that anyone who asserts a statement like "Hesperus is Phosphorus" is really only asserting the contingent proposition that the sentence expresses a necessary proposition. In Stalnaker's view this person does not, and need not, assert the necessary proposition itself (1976, p88-89). For Morris, this admission by Stalnaker amounts to an admission that statements of the form $a=b$ do not assert necessary propositions dealing in self-identity. As Morris himself puts it:

By Stalnaker's admission, the necessary statement of self-identity does not figure in the empirical role of any identity statement. Why then should any such necessary statement be accepted as the correct analysis of an identity statement of the form of S in the first place? (1984, p22)

The failure of the objectual analysis in itself is enough to falsify the claim that identity statements are necessary, since the necessity of identity statements of the form $a=b$ seem to derive their claim to necessity from the fact that at the core of such statements lies the necessary truth that one object is self-identical. I am not denying that the self-identity of an object is necessary, provided we can make sense of the concept of self-identity. What I do deny is that statements of the form $a=b$ have anything to do with self-identity. The above arguments bear this out.

What also bears this out is the meta-linguistic analysis of identity statements. According to this analysis, remember, a statement of the form $a=b$ is not about any objects which are designated by 'a' and 'b', but about the linguistic items 'a' and 'b' themselves. In other words the meta-linguistic analysis says that 'a' and 'b' are co-referential. Now as we have already seen, it may be that there are two versions of the meta-linguistic thesis. There is that dealt with by Frege in his early work, where he deals with the simple co-referentiality thesis; and then there is Frege's later work on sense and reference which advances an analysis of identity statements according to which 'a' and 'b' are not simply co-referential but have senses which contain different modes of presentation of the same object.

Of course Frege's later work may not be any kind of meta-linguistic thesis; and even if it is, it is not straightforwardly so because of the introduction of these modes of presentation which do not seem straightforwardly linguistic. Dummett (1973) certainly regards Frege's later theory as an objectual one. As he says on page 544:

In *Begriffsschrift* Frege held that identity was a relation between names not between things. His motive for this view was to give an explanation of the informativeness of a true identity statement: but it makes nonsense of the use of bound variables on either side of the sign of identity.

On the other hand, Linsky (in his 1977) takes Frege's (1892) to be an analysis of identity which is neither straightforwardly objectual nor straightforwardly meta-linguistic (1977, p22). This is certainly preferable to Dummett's view since Dummett is surely mistaken in claiming that Frege rejected the meta-linguistic thesis because of its inability to cope with bound variables on either sign of the identity sign. As Morris rightly points out: "The occurrence of the sign '=' between bound variables does not *make* an identity statement, it constitutes a statement *form*

for an identity statement.” (1984, p41). Once we assign values to the variables we have an identity statement, but not before then.

Whether Frege’s later thesis is meta-linguistic or not, it is clear that he rejects his early analysis of identity - what we have been calling the simple meta-linguistic thesis. We shall see below that Frege may have been too hasty in his rejection of his early work on identity.

Frege apart, there are two classic objections to the simple meta-linguistic thesis. The first of these is that such an analysis leads to an infinite regress. David Wiggins explains perfectly how the regress is supposed arise in his (1965) thus:

The regress which the two signs theory causes, at least in Frege’s formulation, arises in the following way. Asking for the sense of ‘a=b’ I am told ‘a’ and ‘b’ have the same content, or designate only one thing. Unless something is said to justify calling a halt here, the explanation generates a new statement of the same form as the original explicandum - ‘The content or designatum of “a” = the content or designatum of “b”.’ Applying the same explanation to this we get ‘The content or designatum of “the content or designatum of ‘a’” = the content or designatum of “the content or designatum of ‘b’”.’ But evidently we never can reach in this way what seems to be needed to carry the explanation through a statement only about signs. (1965, In Butler (Ed.), p51)

It seems to me, however, that the simple meta-linguistic thesis does not, or need not, lead to an infinite regress if it is formulated in the correct way. If we take the general statement form ‘a=b’, then instead of explicating it by simply saying that ‘a’ and ‘b’ are co-referential, or have the same content, as Wiggins does, we formulate it in the following way:

$(\exists x) ('a' \text{ designates } x \ \& \ 'b' \text{ designates } x)$

This does not lead to an infinite regress because the sign for identity does not even appear. This is a perfectly reasonable formulation of the meta-linguistic thesis (Although Frege would probably object to it. [See (1892,p1)])

The other classic objection to the meta-linguistic thesis is simpler than the above but every bit as devastating if upheld. The objection is quite simply that the meta-linguistic analysis reduces what is supposed to be a discovery about the world to one

about how two names have been used, which is surely wrong. Furthermore, it seems that on this analysis, someone who discovers that Hesperus is Phosphorus, for example, makes a different discovery from the person who discovers that the evening star is the morning star, or the person who uses names other than 'Hesperus' and 'Phosphorus' for the planet Venus; whereas, intuitively, these people make the same discovery that we do.

These objections do seem to tell heavily against the meta-linguistic thesis, but we must not be too hasty here. Although it is true that the person who discovers that Hesperus is Phosphorus makes the same discovery as the person who discovers that the evening star is the morning star, this is because the name 'Hesperus' has the same *sense* as the expression 'evening star', while the name 'Phosphorus' has the same sense as the expression 'morning star'.

As for the objection that the meta-linguistic analysis fails to account for the fact that we are dealing with an empirical discovery about the world and not a discovery about language use, this takes us to the very heart of the problem about what it is that an identity statement actually states. Of course it is true that in discovering that Hesperus is Phosphorus we are discovering something about the world; although we do make a linguistic discovery at the same time. We could say that we are discovering that two names we previously believed to refer to two distinct objects actually refer to the same object. One of the features of the meta-linguistic analysis which no doubt commends it to many people is that it is in fact true. The items 'a' and 'b' really are co-referential, and this is at least part of what we are discovering. It can certainly be informative to tell someone that 'a' and 'b' refer to the same thing if that person has hitherto believed that they refer to different things.

The problem for Frege, of course, is that, as far as the meta-linguistic analysis is concerned, $a=b$ would only be informative about the language in question, not the ostensible subject matter. As he says in his (1892):

..The sentence $a=b$ would no longer refer to the subject matter, but only to its mode of designation; we would express no proper knowledge by its means...If the sign 'a' is distinguished from the sign 'b' only as object (here, by means of its shape) and not as sign (i.e. not by the manner in which it designates something), the cognitive value of $a=a$ becomes essentially equal to that of $a=b$, provided $a=b$ is true. A difference can arise only if the difference between

the signs corresponds to a difference in the mode of presentation of that which is designated. (In Moore, Ed. 199 , p23,24)

The mode of presentation is of course equivalent to the sense of the name in question - or is contained within the sense of the name.

The meta-linguistic theory does eventually fail on the point that it does not take into account that the discovery being expressed here is a discovery about the world and not purely a discovery about language use. We may be discovering *something* about language when we discover that Hesperus is Phosphorus but basically we are making a discovery about the heavens. Although I think that the meta-linguistic analysis is preferable to the objectual analysis because it is not nearly so deficient as most commentators claim, it is ultimately a failure as an analysis of the content of identity statements of the form $a=b$.

So if the objectual and the meta-linguistic analyses both fail where does it leave us with respect to the content of identity statements? According to Morris we are left with a third analysis - a **functional analysis** - in which we make a fundamental switch of emphasis from asking *what* an identity statement informs us *of*, to asking *how* it informs us. The theory is not well known, and it is Morris' intention to remedy this. It does have a strong pedigree, however, in the sense that it originates with P.F. Strawson (1974) and A.J.Ayer (1975), as well as being advocated by Michael Lockwood (1971) and possibly Butchvarov (1979). The central tenet of the functional analysis of identity statements is that such statements are to be characterised in terms of their cognitive role or function. In Strawson's words, we cease to treat identity statements as predications and concentrate on *how* they inform us. The cognitive role of identity statements is neatly characterised by Morris as follows:

According to a functional analysis an identity licenses or warrants the collating of bodies of information, both previously acquired and yet to be obtained. It does not predicate any particular property of any object or objects. Specifically, it does not attribute a logically simple, essential metaphysical property of synchronic self-identity to any object, nor does it assert any relation to hold between any two linguistic items. (1984, p68)

This is the epistemic function of identity statements. There is also an ontological function, which is to simplify our view of the world by reducing the number of entities in our ontology. For example, prior to the discovery that Hesperus is Phosphorus, we have two bodies of information which we associate with the two names - one body of information per name. On discovering that Hesperus is identical with Phosphorus we are licensed to merge the two bodies into one and thus reduce the number of entities in our ontology by one. This is the cognitive function of identity statements.

The functional analysis of identity statements is fundamentally flawed, however. It does not work either in detail or in principle. In fact in its abandonment of the search for the content of identity statements it simply gives up on the concept of identity altogether. The main reason for its failure lies in its inability to say *how* an identity statement licenses the collating of bodies of information. For example, if an identity statement is to license anything at all, it must be *true* - and presumably the people who are being informed must *know* that it is true. This of course leads to all sorts of problems about what is to count as knowledge. We might, for instance, have good reason to believe something to be true and thereby believe that we are being licensed to collate two bodies of information when actually the identity in question is false. Even if the identity statement in question is true and we do collate two bodies of information into one, this might not be informative. Take, for example, the following statement:

Clark Kent is Superman

If the functional analysis is correct then this should license the collation of the following two bodies of information:

Clark Kent
mild-mannered
meek
timid
newspaper
reporter

Superman
strong
brave
superhero
able to leap
tall buildings
with a single bound

This would allow us to conclude that one person is called 'Clark Kent' and also called 'Superman', and is a mild-mannered, meek, newspaper reporter, and also a brave daring, superhero etc. This is not informative, but downright confusing.

These last two points of criticism are only problems with the detail of the functional analysis, however. The main objection concerns a problem with the analysis, *in principle* - and this is the problem of *how* an identity statement licenses the collation of bodies of information. As already mentioned, the statement in question would have to be true, and we would have to know that it was true. *But how can we know that a statement is true unless we know what its content is* - i.e. unless we know what it is a statement of? The functional analysis fails because it is designed to avoid the problem of content, but it cannot do so. No sooner does it try to side-step the question of content than it is pulled straight back to it. It is therefore of no use as an analysis of identity statements. Even if we weaken the knowledge requirement and substitute justified belief or something similar for certain knowledge, we still have the same fundamental flaw. For instance, we might want to say that an identity statement licenses the collating of two bodies of information if people have good reason to accept the statement, or are justified in believing the statement. But this still involves us in the problem of content - i.e. what is it that these people are justified in believing? That $a=b$? Fine, but what exactly is that? We are back to square one. The functional analysis is a complete failure.

So far, then, we have seen that the traditional analyses, and not so traditional analyses of identity statements have failed to provide us with an adequate characterisation of what it is that an identity statement actually *says*. We have seen that the objectual analysis fails because of its inability to make sense of the difference in informativeness between statements of the form ' $a=a$ ' and those of the form ' $a=b$ '. The meta-linguistic thesis, although not nearly as inadequate as some would claim, cannot account for the fact that an identity statement is supposed to deal with an empirical discovery about the world and not just a discovery about the use of two linguistic items. Frege's later work on the sense/reference distinction provides no real improvement on his earlier meta-linguistic thesis for various

reasons. First of all, it may be that it is really a version of the objectual analysis where, instead of simply talking about the objects a and b, he talks about those objects in a certain way; or as presented to the mind in a certain way. According to Frege, this analysis resolves the difficulty of explaining the difference in cognitive value between 'a=a' and 'a=b' because of the difference in the senses of 'a' and 'b', which contain different modes of presentation of a and b. But this is no improvement on the early theory because the names involved in this early theory would have to have senses anyway, in order to be names. Secondly, if Frege's later theory is an objectual theory then it fails because he still has to talk about two objects being identical. How exactly can two objects be identical? In the case of the identity 'Hesperus is Phosphorus' why do we talk about the *objects* Hesperus and Phosphorus being identical? Are we dealing with two objects, one called 'Hesperus' and one called 'Phosphorus'? Surely not; although this may depend on how we are to define an object. [See e.g. Butchvarov (1979) who distinguishes between objects and entities in order to deal with the identity puzzle] This alone should tell us that there is something wrong with the objectual analysis.

The Strawson/Ayer/Morris functional analysis, as we have seen, is hopefully flawed and therefore no alternative to the traditional analyses. This seems to lead to the conclusion that there is no adequate way in which we can make sense of so-called identity statements of the form 'a=b'. If the above-mentioned analyses exhaust the possibilities then we must conclude with Wittgenstein that expressions of the form 'a=b' are only expedients in presentation (*Tractatus* 4.242) and even that the identity sign is not an essential constituent of logical notation (*Ibid.* 5.533); although we should also remember that in the *Investigations* Wittgenstein seems to retreat from the certainty of the *Tractatus* when he says at 215:

But isn't *the same* at least the same?

We seem to have an infallible paradigm of identity of a thing with itself....If you are seeing a thing you are seeing identity.

This is clearly a departure from the *Tractatus* view where Wittgenstein says that to say of two things that they are identical is nonsense whereas to say that a thing is identical with itself is to say nothing (5.5303)

Obviously Wittgenstein like everyone else who has written on the subject of identity, sees little or no difference between statements of the form 'a=a' and those of the form 'a=b'. Here we must be clear about the purpose of the present chapter. Our intention is not to challenge the validity or legitimacy of statements of the form 'a=a', or self-identity, but to question the necessity, and even the legitimacy of statements of the form 'a=b'; what Butchvarov calls 'material identity', as opposed to the formal identity statements 'a=a'. It is a contention of the present chapter that statements of the form 'a=b' have nothing whatsoever to do with self-identity because they are, or can be, empirical discoveries, whereas self-identity *cannot* be an empirical discovery. So far we have seen that all the traditional analyses of material synchronic identities have failed. If we are to conclude from this that such expressions do not make genuine statements then obviously they cannot be necessary truths, and so we have succeeded in showing that there are no necessary a posteriori truths involving identity statements.

If the foregoing is not conclusive then there are other ways in which we can question the necessity of identity. First of all, consider the alleged proof of the necessity of identity from Marcus (1947) which goes something like the following:

1. $(\forall x) (\forall y) (x=y \rightarrow (Fx \leftrightarrow Fy))$	Assumption
	Leibniz' Law
2. $x=y$	Ass.
3. $x=x$	=I
4. $\Box (x=x)$	3, \Box I
5. $\Box (x=y)$	2,4,=E
6. $x=y \rightarrow \Box (x=y)$	2,5 \rightarrow I

Quantifying over x and y we get

$$7. (\forall x) (\forall y) (x=y \rightarrow \Box (x=y))$$

In other words, if necessary identity with x is a property of x , and $x=y$, then by Leibniz' Law, it must be a property of y . When we substitute proper names for x and y we get the necessity of statements like 'Hesperus is Phosphorus'. That is, we get:

8. $(a=b) \rightarrow \Box (a=b)$ where 'a' and 'b' are proper names

Responses to this proof are varied, ranging from the Russellian response which claims that a statement like, for example, 'Cicero is Tully' is not a genuine identity statement (because it does not involve what Russell regards as logically proper names) but a disguised existential statement, to Quine's retort that substitution of co-referential terms fails to preserve truth in modal contexts. More recently there has been the response of Hacking (1976) who argues that the Marcus proof fails because identity is not a relation between individuals. Taking his definition of identity from *Principia Mathematica*, where identity for individuals must be defined in terms of predicates of Type I and Level I (in Church's terminology), Hacking claims that identity of an individual, and hence necessary identity of an individual, is of Type I, Level 2 (1976, p153). I shall not deal with these views here but merely mention them as an indication of the controversy that has arisen in the past over the question of the necessity of identity.

In recent times this controversy seems to have been resolved by Kripke in "Identity and Necessity" and *Naming and Necessity*. Kripke's response to (8) is to accept it so long as 'a' and 'b' are rigid designators - i.e. they designate the same object in all possible worlds where the relevant object exists. In "Identity and Necessity" Kripke argues as follows:

If names are rigid designators then there can be no question about identities being necessary, because 'a' and 'b' will both refer to this same object x , and to no other, and so there will be no situation where a might not have been b . That would have been a situation in which the object we are now calling 'x' would not have been identical with itself... One could not possibly have a situation where Cicero would not have been Tully or Hesperus would not have been Phosphorus (p154)

He goes on to claim that anyone who thinks he can imagine a situation where the identity would not hold is deluding himself. According to Kripke any supposed counter-example to the necessity of identity involving proper names trades on the confusion between descriptions which merely fix the reference for us in the actual world and descriptions which are synonymous with the names. Indeed Kripke has become so adept at rebuffing alleged counter-examples to his thesis that most philosophers have given up trying to think of counter-examples. Kripke's view has become the orthodoxy.

In "Identity and Necessity" Kripke imagines a possible counter-example to the claim that Hesperus is necessarily identical to Phosphorus and that is the possible situation (possible world) where a planet other than Venus is visible in the morning and is called 'Phosphorus'. Venus is still visible in the evening and is called 'Hesperus'; and, so the argument goes, this would be a world where Hesperus is not Phosphorus. But this is just a mistake, in Kripke's view. Although the people in this other world would be able to *say*, truly, "Hesperus is not Phosphorus" that is irrelevant. It does not matter how names are used in other possible worlds, but only how *we* use them in the actual world (even though we could be the people in these other worlds). What is important is that we use the names 'Hesperus' and 'Phosphorus' *in the actual world* as rigid designators of the planet Venus in all possible worlds where it exists; and so it is necessary that Hesperus is Phosphorus, in Kripke's special sense of 'weak necessity'. The alleged counter-example fails because it mistakes a reference-fixing description such as 'The heavenly body in such-and-such a position in the evening/morning sky' for a description which gives the meaning of the name. If the names in question were synonymous with descriptions then, according to Kripke, the identity would be contingent since there are possible worlds where the heavenly body in such-and-such position in the evening sky is not the same as the heavenly body in such-and-such position in the morning sky.

Before proceeding to the question of whether or not there are counter-examples to Kripke's claims I wish to examine more closely the formal proof, particularly the claim that there is such a property as being necessarily identical with x, or even

being identical with x. I shall also question the substitution of identicals in modal contexts, but for reasons other than those given by Quine.

As is well known, Quine (1961) blocks the proof of the necessity of identity by rejecting the claim that co-referential terms are always and everywhere intersubstitutable *salva veritate*. In intensional contexts such as epistemic, doxastic and modal contexts, substitution fails to preserve truth. For example, John may believe that Cicero denounced Catiline but not that Tully did so; even though the names 'Cicero' and 'Tully' are co-referential. These are Quine's famous referentially opaque contexts, according to which he would not allow the substitution of 'y' for 'x' in line (4) of the Marcus proof, thus blocking it.

Whatever the merits or deficiencies of Quine's argument I shall not follow his example, but nevertheless I do wish to question the substitution of 'y' for 'x' in line (4). My criticism is simpler than Quine's. I do not contend that Marcus goes too far in her substitution, but rather that she does not go far enough. What I mean by this is that, if we are to substitute 'y' for 'x' in the formula

$$\Box (x=x)$$

then, quite simply, we must substitute for **both** occurrences of 'x'. If we do this, all we can conclude is the formula

$$\Box (y=y)$$

which is uncontroversial because it is trivial.

It seems to be generally accepted that substitution of identicals in formulae need not take place for every occurrence of the proposed substituent. If this is the case then it is wholly unjustified. If we are to substitute 'y' for 'x' in any formula, we must surely substitute for *all* occurrences of 'x' because there is no difference between them. They are all the same object, or term, or whatever.

The other way in which we can look at the Marcus proof is in terms of properties rather than substitution. Seen in terms of properties the Marcus proof would go as follows:

x has the property of necessary identity with x

$x=y$

Ergo y has the property of necessary identity with x (by Leibniz' Law)

This is argued for by Wiggins (1976) using the Lambda Calculus, thus:

- (1) $(x)(y)((x=y) \supset (Fx \equiv Fy))$
- (2 λ) $(x)(y)((x=y) \supset ((x \text{ has } (\lambda z)[[\text{Nec}[\lambda r\lambda s[s=r]]], [x,z]] \equiv (y \text{ has } (\lambda z)[[\text{Nec} [\lambda r\lambda s[s=r]]], [x,z]])))$
- (3 λ) $(x) (x \text{ has } (\lambda z)[[\text{Nec}[\lambda r\lambda s [s=r]]], [x,z]])$
- (4 λ) $(x)(y)(x=y) \supset (y \text{ has } (\lambda z)[[\text{Nec}[\lambda r\lambda s[s=r]]], [x,z]])$ (1976, p112)

Of course we could also express the proof thus:

x has the property of necessary self-identity

$x=y$

Ergo y has the property of necessary self-identity

But this is trivial and does not require any kind of proof. If we want the substantial result we must adopt the first version where we talk of the property of identity with x , rather than self-identity. But this requires that we *distinguish between* these two properties. But what exactly is the difference between x 's being self-identical and x 's being identical with x ?

One obvious line of reply is that whereas all objects have self-identity, only x has the property of identity with x ; but this would be to miss the point of the criticism. What we really want to know is what is the difference between x 's self-identity and x 's being identical with x ? Unless this question can be answered the proof of the necessity of identity fails in this form.

Again we can cite Wiggins' (1976) on the difference between the property 'a is identical with a' and the property 'a is self-identical'. As Wiggins explains (p165,166):

That $\lambda x(x=a)$ and $\lambda x(x=x)$ are different properties can be explained by tracing their actual relationship, which may be displayed as follows. Start with the primitive predicable 'x=y' and abstract the property $(\lambda x)(\lambda y)(x=y)$. That this is satisfied by a pair is recorded in our notation by $(\lambda x)(\lambda y)(x=y),[w,z]$. To get a's own peculiar predicate $(\lambda x)(x=a)$ or $(\lambda w)[(\lambda x,\lambda y(x=y)),[w,a]]$ substitute a designation of a for z in the second free argument place in $(\lambda x)(\lambda y)(x=y),[w,z]$ and bind the only free variable with λ . Or to get the simpler designation $(\lambda x)(x=a)$, take the constant a as a value of z in $(\lambda x)(\lambda y)(x=y),[w,z]$ and apply λ -conversion to get 'w=a'. Replace the variable w by the variable x and abstract on x to get $(\lambda x)(x=a)$. To get the universal predicate $(\lambda x)(x=x)$ or $(\lambda w)[(\lambda x)(\lambda y)(x=y),[w,w]]$ substitute the first argument for the second argument place in $(\lambda x)(\lambda y)(x=y),[w,z]$ to get $(\lambda x)(\lambda y)(x=y),[w,w]$ and bind the free variable with λ . To get the simpler formula put t for both w and z in $(\lambda x)(\lambda y)(x=y),[w,z]$, apply λ -conversion to get 't=t', replace t by x to get 'x=x' and abstract on x to get $(\lambda x)(x=x)$.

The problem with the Marcus proof (and also the Wiggins version) is that it assumes a reading of $\Box(x=x)$ which it does not justify. That is, $\Box(x=x)$ must be read as attributing the property 'necessary identity-with x' to x, rather than the property 'necessary self-identity' to x. In λ -abstraction terms the logical truth $\Box(x=x)$ is most easily and naturally seen as instantiation by x of the property of necessary self-identity: $\lambda y. \Box(y=y)[x]$. But then Leibniz's Law yields only the trivial result

$$x=y \rightarrow (\Box(x=x) \rightarrow \Box(y=y))$$

If $\Box(x=x)$ is read as either attribution to x of the property of necessary identity with x, viz. $\lambda z. \Box(z=x)[x]$, or as attribution to the pair x,y of the property of necessary self-identity, $\lambda z\lambda w. \Box(z=w)[x,y]$, then in both cases it is being *presupposed* that x (or x and y) has this property. But this is just begging the question. There is no reason to read $\Box(x=x)$ in this way, rather than simply as an attribution of self-identity. The Marcus/Wiggins 'proof' fails without further justification of their particular reading of $\Box(x=x)$.

Marcus regards the formula

$$(7) \quad (\forall x) (\forall y) (x=y \rightarrow \Box (x=y))$$

as a theorem which falls out naturally from her treatment of quantified modal logic; but this is just wrong. (7) is not a natural result but a wholly artificial one arising from an illegitimate and unjustified move. I am not of course accusing Marcus of any kind of gerrymandering or logical sleight of hand because the move in question is quite open; but I do claim that she must justify what seems an arbitrary move if the proof is to be accepted. This, I maintain, she cannot do. There is no way in which anyone can justify this picking and choosing among the occurrences of 'x' in the formula $\Box (x=x)$ when we are supposed to be substituting 'y' for 'x'. To substitute only one occurrence rather than both must be justified because of the considerable difference in the result. Marcus must explain why her move is not arbitrary if we are to accept her substantial result. I see no way in which this can be done, and there is no use in appealing to precedent here. This is supposed to be a logical proof, not a legal case. I maintain, then, that if we are to substitute 'y' for 'x' in any formula we must substitute for all occurrences of 'x', and so in the present case, all we can conclude is that it is necessary that $y=y$, which is trivial.

So far, then, we have seen that the claim to the necessity of identity fails for the following reasons:

1. Names may not be rigid designators, or indeed any kind of designators
2. The failure of the objectual, meta-linguistic and functional analyses to give an acceptable account of what identity really is
3. The failure of the Marcus 'proof' on two counts
 - (i) The illegitimate substitution of only one occurrence of 'x' in $\Box (x=x)$
 - (ii) The failure to explicate the difference between the properties 'necessary

self -identity` and `necessary identity with x`

What remains is to challenge Kripke's claim that there are no counter-examples to his alleged examples of the necessary a posteriori - worlds where Hesperus is not Phosphorus and Cicero is not Tully etc. - and also to give an account of why it is that we use such statements and seem to understand them if they are not genuine statements.

I shall maintain that there are indeed counter-examples to Kripke's alleged necessary a posteriori truths - at least where ordinary proper names are concerned. If we take the stock example:

Hesperus is Phosphorus

I contend that there is a possible world where the planet Venus exists and yet it is not the case in that world that Hesperus is Phosphorus - or, at least it does not make sense to say that Hesperus is Phosphorus. Such a world could be the following (call it **W**):

The possible world where Venus exists and is visible from Earth, but only in the evening - never in the morning - and is called simply 'Hesperus'

Now *prima facie* this would seem no different from the worlds already dealt with by Kripke where he points out that it is irrelevant what the object in question is called in another possible world or how the people in that world use names. As Kripke is at great pains to emphasize, it is how *we* use the names in question in the actual world that is important. In the Hesperus/Phosphorus case we use the names 'Hesperus' and 'Phosphorus' as rigid designators of the planet Venus, which means that there is no possible world where Venus exists and Hesperus is not Phosphorus since the planet will be designated by the names 'Hesperus' and 'Phosphorus' in all worlds where it exists. It will be the case that Hesperus is Phosphorus in any arbitrarily chosen world simply if the planet Venus exists in that world. What the inhabitants of the

world say, or how they use certain names, is totally irrelevant to the question of whether it is true in that world that Hesperus is Phosphorus.

This has become a fairly standard view in philosophy, but nevertheless I maintain that it is mistaken. It *is* important what people say in other worlds; or rather, it is important what they are in a position to say, in principle. In the above example, for instance, it seems that the inhabitants of **W** could not make the *empirical discovery* that Hesperus is Phosphorus. The same people who discover that Hesperus is Phosphorus in the actual world cannot make this discovery *by any means whatsoever* in **W**; so of course they will never be in the position to assert that Hesperus is Phosphorus in **W**. This failure to be able to discover what is supposed to be uncontroversially true in certain worlds is the essence of my criticism of Kripke here. The question we must answer is ‘Why is it that people cannot discover that, say, Hesperus is Phosphorus in a certain world, when they can discover it in other possible worlds?’ This is a type of counter-example which Kripke has not considered, but which seriously undermines his claim to the necessity of identity because there does not seem to be any explanation of why it is that there are some worlds where certain things are supposed to be the case and yet it is totally beyond the powers of the inhabitants of that world to discover these things.

To make the point clearer, suppose we imagine a possible world, again call it **W**, where Dr Jekyll exists but does not drink a potion which turns him into a hideous creature called ‘Mr. Hyde’, or anything else. He simply lives a quiet, respectable and unremarkable life. Moreover, there is no-one in **W** called Mr. Hyde. Now if we want to say that this is a world where Dr. Jekyll is Mr. Hyde (which Kripke undoubtedly would) then what I want to know is this. Could the people in **W** ever make the *empirical discovery* that Dr. Jekyll is Mr. Hyde? It seems to me that they could not. They certainly could not discover it in the way in which it is discovered in the actual world (or the *fictional* actual world) - i.e. by discovering that Dr. Jekyll has drunk a potion which transforms him into a hideous creature - so how exactly would they make the discovery? Remember, these could be the same people who make the discovery in the actual world - the same people with the same means at their disposal for making discoveries about the world - so they should be able to discover

that Dr. Jekyll is Mr. Hyde, if it is true in **W**. They need not *actually* make the discovery of course; but they should be able to - *in principle*. Furthermore, they should be able to express this discovery, *in a non-trivial way*, in whatever language they speak. And yet there does not seem to be any way in which they could do this. It is not just very difficult for them to make the discovery. Neither is it a case of them having to make the discovery in a different way. It is quite simply beyond these people to discover empirically something they can discover in other worlds. This is very strange, considering that the only difference between **W** and the actual world is a difference that Kripke, and just about everyone else, would regard as irrelevant to the question of whether **W** is a world where Dr. Jekyll is Mr. Hyde.

Now this inability to make the discovery that Dr. Jekyll is Mr. Hyde in **W** does not in itself show that this is not true in **W**, but it does require an explanation. If this fact can be discovered in the actual world then why can it not be discovered by the same people in **W**? One possible answer to this question is that it is simply not true in **W** that Dr. Jekyll is Mr. Hyde. That is why these people cannot make the discovery. **W** is a world where Dr. Jekyll exists but is not identical with Mr. Hyde; and so (7) is not true. Identities involving ordinary proper names are not necessary.

Before dealing with possible objections to this counter-example there is another way in which we could argue the point, and that is in terms of behaviour. In the actual world, when people discover that Dr. Jekyll is Mr. Hyde, the discovery is shocking and outrageous, and profoundly affects people's behaviour. In **W**, on the other hand, there is nothing that these people can discover about Dr. Jekyll that is remotely shocking or outrageous; and so we can conclude that they cannot discover the same piece of knowledge.

To the objections, then. I shall consider three.

- (i) The people in **W** cannot discover that Dr. Jekyll is Mr. Hyde because they lack the *concept* of being Mr. Hyde
- (ii) The behavioural analysis fails because the reason that the discovery is so shocking in the actual world is that it is made against a certain background,

where a hideous creature has been terrorising the streets of London - a background that does not exist in **W**

(iii) There is nothing to prevent names from being rigid designators and having sense given by a definite description, so long as the description is made rigid.

Given this, the reason the inhabitants of **W** cannot discover that Dr. Jekyll is Mr. Hyde is simply that they lack the necessary name with the necessary sense.

To deal with (i) first of all, it really just shifts the problem. If we cannot explain why people in **W** cannot discover that Dr. Jekyll is Mr. Hyde it is just as difficult to explain why they do not have the concept of being identical to Mr. Hyde, because it seems that they never *could* have the concept of being Mr. Hyde. This means that they can never conceive of the possibility of the identity, or conjecture about the identity, the way they can do in the actual world. Again this must be explained.

(ii) and (iii) can be answered in the following way. Suppose that there is someone in **W** called 'Mr. Hyde', who looks like the Mr. Hyde of the actual world, and behaves in the same way, but he is not Dr. Jekyll. Now it seems that qualitatively, or epistemically, these people would have the necessary name with the necessary sense, plus the necessary background against which the discovery could be made. They are in the same epistemic state that they would be in the actual world prior to the discovery that Dr. Jekyll is Mr Hyde, and yet nothing they can discover empirically about Dr. Jekyll in **W** would count as the discovery that he is Hyde - so the problem remains.

If I am right about the Dr. Jekyll and Hesperus/Phosphorus cases then they represent a whole class of counter-examples to Kripke's claims. All we have to do is imagine a world where $a=b$ is supposed to be true and yet the inhabitants of this world cannot make the discovery by any means whatsoever. This requires an explanation, and one possible explanation is that $a=b$ is not true in that world after all. In the absence of any other explanation I feel entitled to propose this.

The foregoing argument trades on a phenomenon that Kripke regards as irrelevant; namely, what the inhabitants of other worlds say, or are able to say. More precisely, it is concerned with what these people could discover, in principle. This is the novelty of the argument. It is a way of opposing the claim that identities are necessary that Kripke has not had to deal with before, and it would be interesting to see how he would deal with it.

It should also be pointed out here that the meta-linguistic analysis of identity explains perfectly why the inhabitants of **W** cannot make the same discovery they make in the actual world. It also gives a straightforward reason to reject the necessity of identity.

If identities of the form $a=b$ are not necessary, then, what are the consequences for Kripke or anyone who holds similar views? The main consequence is that he cannot cite statements like 'Hesperus is Phosphorus' or 'Cicero is Tully' as examples of the necessary a posteriori - although the so-called scientific identities may be different (More of this in the next chapters). Secondly, we could have here a *reductio ad absurdum* of the claim that ordinary proper names are rigid designators as follows. Kripke argues that if proper names are rigid designators then identities of the form $a=b$ involving proper names will be necessary. Well if these statements are not necessary, then proper names are not rigid designators, by *Modus Tollens*. And if names are not rigid designators then Kripke's whole philosophy is in serious trouble. The only way to rescue the thesis that names are rigid is to blame the *reductio* on identity statements themselves. That is, to agree with the early Wittgenstein that identity statements make no sense - or at least we do not really know what they mean. Again, though, Kripke is in serious trouble because most, if not all, of his examples of the necessary a posteriori seem to be in the form of identity statements.

Before concluding, there are still a couple of objections to be dealt with. First of all, it might be objected that the claim that statements of the form $a=b$ are not necessary does not count against the *metaphysical* view:

$$(7) (\forall x) (\forall y) (x=y \rightarrow \Box (x=y))$$

Actually I think it does, indirectly, in that (7) could never be instantiated. There could never be a particular instance of (7) substituting singular terms for x and y.

Another possible objection to the necessity of identity concerns statements of arithmetic. Surely we have here the very paradigm of necessary identity statements. This objection can be answered in the following way. Although statements of arithmetic are undoubtedly necessary truths, they are not identities. Take, for example, the statement:

$$2+2=4$$

This is not an identity because the left hand side of the statement, namely '2+2' is an operation, whereas the right hand side is not an operation. By Leibniz' Law, therefore, the statement '2+2=4' cannot be an identity. It is still a genuine statement of course, but it is not an identity. Instead of saying that 2+2 *equals* 4 we must say something like "2+2 *results in* 4 ". Or, the operation of adding 2 to 2 results in 4. This would seem to be necessary, but it is not an identity. At the very least, statements of arithmetic need not be regarded as identity statements.

In conclusion, then, there is a strong case to be made against the necessity of identity - even against the very legitimacy of expressions of the form a=b or x=y. It is very difficult to find a way of analysing what an identity statement is supposed to mean, and Kripke certainly does not say what they are supposed to mean. There are also basic problems with the Marcus 'proof' of the necessity of identity, and also, as we have seen, the real possibility of counter-examples in terms of possible worlds where a certain identity is supposed to hold but no-one in that world can make the discovery. This makes it very difficult to claim that the identity actually holds in that world. At the very least it becomes difficult to see what it means to say that the identity holds in that world.

There is one final point to be made, and it concerns the question of why we use statements of the form a=b and seem to understand them, if they are not genuine statements. The first point to make about this is that I do not believe that we really do understand such statements at face value. I contend that it is impossible to understand a statement like 'Hesperus is Phosphorus' without also understanding the situation it stands for; namely the situation whereby one planet is visible at two

different times of the day. In other words, so-called identity statements may state *that* what could be taken to be two objects are really one, but they do not state *how* they could appear to be two. In order to be understood an expression of the form $a=b$ must not only say *that* 'a' and 'b' are identical, but also *how* they are identical. The bare expression ' $a=b$ ' does not do this, and so it is not a genuine statement. Once again understanding is the key concept here. If statements of the form $a=b$ make genuine statements they must be understood in a certain way - otherwise they would not even be part of a language. Now the key question becomes: "What must be *understood* by sentences of the form $a=b$?" Are we to understand them as making statements of self-identity? I.e. trivial uninformative statements whose content can only be known a priori? Surely not, since such statements can be the subject of serious speculation, conjecture and debate - debate which can be resolved empirically. They must, therefore be understood in a different way. The reason we use such expressions I suggest comes from a confusion between synchronic and diachronic identity. When we say, for example, that Hesperus is identical with Phosphorus, what we are really saying is that the same object which we identify at one time (and in a certain way) as Hesperus can be *re-identified*, *at a later time* (and in a different way) as Phosphorus - which is contingent because we might not have identified the planet Venus in these ways. This also applies to other examples, such as the Cicero/Tully example and the Jekyll/Hyde example. So-called 'synchronic identity' of the form $a=b$ is really an illegitimate form of diachronic identity, and so not really a distinct form of identity at all. However we regard identity statements involving ordinary proper names, they are not necessary truths.

I

The Contingent A Priori

We come now to the crux of the matter. Are there, as Kripke and others claim, *contingent a priori* truths and *necessary a posteriori* truths, and can we establish their existence from a theory of how names function in ordinary languages? The necessary a posteriori will be the subject of Chapter 6. In the meantime we turn our attention to the claim that there are contingent a priori truths, a claim that, not surprisingly, has led to considerable debate among analytic philosophers.

Forbes (1989) distinguishes two categories of the alleged contingent a priori. The first involves indexicals such as:

1. I am here now

and are due to Kaplan (1979). The second concerns the more familiar examples of what Evans (1979) calls 'descriptive names', and are mostly associated with Kripke's work in *Naming and Necessity*. These would include such examples as Kripke's standard metre case and his claim that the astronomer Leverrier could have known a priori that Neptune was the cause of certain perturbations in the orbit of Uranus if he introduced the name 'Neptune' into the language via the description 'The planet that is causing the perturbations in the orbit of Uranus', or something similar. We may also add Evans' example of the name 'Julius' which is introduced into the language as a rigid designator of whoever invented the zip, such that the statement

2. If anyone invented the zip, Julius did

would have the status of a contingent a priori truth (Evans, 1979). Crucial to the success of the examples involving names is Kripke's distinction between descriptions that give a synonym and those that merely fix a reference. Kripke introduces the distinction on page 14 of (1980) as follows:

Let "a" (rigidly) denote the unique object that actually has property F, when talking about any situation, actual or counterfactual. It seemed clear that if a speaker did introduce a designator into a language that way, then in virtue of this very linguistic act, he would be in a position to say 'I know Fa' but nevertheless 'Fa' would express a contingent truth.

Donnellan provides us with a neat summary of Kripke's position in his (1979) where he says:

....suppose we propose to introduce the term "t" as the name of the denotation, if there is one, of the definite description "the ϕ ." And we do not intend that "t" shall *mean* "the ϕ " or be a mere abbreviation for it. Rather "t" is to designate whatever happens to be the ϕ ; the definite description serves solely to fix the reference. If our procedure serves to introduce "t" as a rigid designator, "t" will designate the same entity in all possible worlds (in which it designates anything at all). It will designate the same thing that it does in this, the actual world. But there will be possible worlds in which what is the ϕ in the actual world, that which "t" designates, is not the ϕ in that world. Thus the statement that if the ϕ exists, t is the ϕ is merely contingently true, because there are possible worlds where it is false. Yet, if the reference of "t" has been fixed solely by being the denotation of the description "the ϕ " it looks like it can be known *a priori* that if the ϕ exists, "t" is the ϕ (1979, p45)

Kripke's most famous example of the contingent a priori concerns the standard metre stick near Paris. In (1980, p54) he imagines a situation where a certain person fixes the reference of the term 'one metre' by stipulating that it is to be the length of a certain stick, S, at a definite time, t_0 . The description 'The length of S at t_0 ' is not synonymous with the term 'one metre', but rather fixes its reference as a rigid designator of the length of S at that moment in time. According to Kripke, anyone who fixes the reference of the term 'one metre' in this way would know a priori, or

‘without further investigation’, that S is one metre long at t_0 , even though it will be contingent that S has this length at that time.

In what follows I shall differ slightly from Forbes in identifying three categories of purported contingent a priori truths rather than two. That is, the examples involving indexicals (Type 1); the so-called ‘descriptive names’ like ‘Julius’ and ‘Neptune’ (Type 2); and a third category consisting of definitions of units of measurement such as Kripke’s Standard Metre example (Type 3). These latter form a category distinct from Type 2 because whoever fixes the reference of the measurement term is in a position to actually baptise the standard and have causal contact with it, unlike the ‘Julius’ case, for example, where we could question whether we really are able to name anything. Consequently we must evaluate the various categories differently - at least in some respects. We will not spend too much time on the Type 1s since they do not involve names.

How we evaluate the claims to the existence of the contingent a priori will involve us in answering the following question. Can we really come to know something contingent about the world purely as a consequence of a linguistic act - namely, the act of introducing a name into the language by means of a non-synonymous reference-fixing definite description? Another way of posing the question is this. When we introduce a name like ‘Julius’ or ‘Neptune’ or ‘One Metre’ into the language in the way Kripke suggests, have we really come to know anything that we did not already know? For example, before we introduce the name ‘Julius’ as a name for whoever invented the zip, we already know

3. The inventor of the zip invented the zip (if anyone did)

Furthermore, and assuming that there are any contingent truths at all, we know

4. The inventor of the zip might not have invented the zip

So what new piece of knowledge do we acquire when we are in a position to say

5. *Julius* invented the zip, although *Julius* might not have invented the zip

What Kripke must do is show that (5) is an improvement or advance on (3) and (4) in terms of furnishing us with new knowledge, and it is very difficult to see how this can be done, especially with the Type 2 examples. If this cannot be done then the contingent a priori would seem to be no more than an illusion, the result of some logical sleight of hand.

That the claims to contingent a priori knowledge are indeed based on some logico-semantic delusion which does not give us any real knowledge about the world is supported by the fact that we seem to be able to introduce names into the language which give us contingent a priori knowledge of almost anything. For example, if there is a car parked outside my house I can reason that someone must have put it there. It did not get there by itself. I do not know who actually did put the car there, however, although I could follow Kripke's procedure and introduce a name into the language as a name for whoever put the car outside my house. Do I thereby come to know a priori some contingent truth such as :

If anyone parked this car outside my house, N did

where "N" is the name I have introduced into my language via the appropriate reference-fixing description? If I do then it is surely a cheap way of acquiring knowledge and the contingent a priori would not amount to much. Indeed it does not amount to much because I still don't know who parked the car outside my house. That is something I can only come to know a posteriori.

Another possible example will indicate more dramatically the difficulty Kripke has in establishing his position. Suppose, for instance, that I don't know who invented bi-focals but I want to introduce a name into the language in the reference-fixing Kripkean manner as a name for whoever has this distinction. Suppose further, that by pure chance I decide to use the name 'Benjamin Franklin' as a name for whoever invented bi-focals. According to Kripke I would be in a position to know a priori:

6. If anyone invented bi-focals, Benjamin Franklin did

Now intuitively this cannot be the same piece of knowledge as the *a posteriori* knowledge that Benjamin Franklin invented bi-focals and the difference can only be traced to the way in which the name 'Benjamin Franklin' is being used in the two cases. In (6) the name does not give us access to any particular person but merely abbreviates the description "The inventor of bi-focals, whoever that may be". (6) does not allow us to conclude anything beyond the knowledge that whoever invented bi-focals, the inventor of bi-focals did.

One obvious difference between the *a priori* (6) and the *a posteriori* knowledge that Benjamin Franklin invented bi-focals is that the latter does not need to be existentially quantified in the manner of (6). This quantification is thought necessary as an answer to what has been termed 'The Existential Complaint', one of the earliest criticisms of Kripke's claims. (See especially Carter, 1976; also Ray, 1994; and Cowles, 1994).

The Existential Complaint is quite simply the claim that we cannot know *a priori* a truth like

Julius invented the zip

unless we first know that there is such an inventor, which is something we can only know *a posteriori*. As Carter (1976) says concerning the standard metre stick:

One of the things that must be true if it is to be true that S has the property of being one metre long is that there exists such an object as S.... But our knowledge that there is such an object is clearly not *a priori*. (1976, p105)

Not only does this make Kripke's claim to the contingent *a priori* depend on an *a posteriori* premise, it makes such examples trivial according to Carter. To say that S is one metre long is as trivial as saying that I know *a priori* that my fountain pen is on my desk once I have observed it to be there. (Ibid. p106)

So conditionalising the target sentences in the manner of (2) is supposed to avoid the existential complaint and has become the standard way of formulating the puzzle about the contingent *a priori*. It is certainly the form adopted by Donnellan

(1979), Forbes (1989) and even Kripke himself, at least with respect to the 'Neptune' example where he makes the following qualified claim:

If the perturbations of Uranus are caused by a planet, they are caused by Neptune (1980, p79n)

Actually this needs to be further conditionalised since there may not be any perturbations of Uranus, and even if there are, we certainly cannot know this a priori. This can easily be done, however.

Be that as it may, the existential complaint has been subjected to a fair amount of criticism by at least some commentators. David Cowles in his (1994) argues that the 'Existential Worry', as he calls it, is not a telling objection to Kripke's examples because there could be examples free from existential worry. Such an example, according to Cowles, must not involve the name in question in existential commitment to any contingently existing thing. The example used by Cowles is that of a name introduced as a name of a number - for example, the number of planets. If we introduce the name 'alpha' as a rigid designator of the actual number of planets, whatever it is, then we would know a priori

(7) Alpha numbers the planets

or

(7*) There are alpha planets

There is no existential worry here because 'alpha' gets a reference even if there are no planets. We can know a priori that numbers exist and also that the definite description 'The number of planets' will be satisfied, so we know (7*) a priori. (1994, p140). Curiously Cowles does not take this to show that it is possible to have a priori knowledge of contingent truths, but he does claim that it shows that the existential worry is not a telling objection to Kripke's examples (Ibid).

Cowles is surely correct in arguing that his example is free from existential worry, and also that it does not show that there are contingent a priori truths, since

we do not know what number alpha is. But if this is so, then what is the point of the example? And how does it show that the existential worry does not arise for the examples we have already encountered? Cowles' examples involving numbers are surely a special case where it is built in to the example that there will be no existential worry. The Existential Complaint still stands for the other examples.

Evans (1979) also argues that the Existential Complaint does not count against the puzzle but for reasons very different from Cowles. Quite simply, Evans regards it as irrelevant to the puzzle since the puzzle can only arise if the use of free logic is accepted. As he says in his (1979)

Unless a sentence containing 'Julius' can be formulated which is free of existential commitment there is not even a candidate for the status of the contingent a priori.... No matter how a name may be embedded in a sentence, in a classical language that name is accessible to existential quantification and the truth of the whole sentence requires that the name refer. (p172)

Consequently, in Evans' view, conditionalising the sentence in question is "quite pointless" (Ibid). Evans' main target here is Donnellan's (1979) treatment of the puzzle where it is argued that the person who knows only the reference-fixing description does not *understand* the name (in this case 'Julius') nor sentences containing it. In Donnellan's view, we can only understand a name like 'Julius' if we know *of* some object that it is the referent of the name. Consequently Donnellan seeks to show that the contingent a priori is uncontroversial because it does not give us any knowledge of the world. Donnellan's main contention is that we must distinguish between knowing that a certain sentence expresses a truth and knowing the truth of what is expressed by the sentence (1979, p51). With a sentence such as (2) we know a priori that it is true, but we do not know a priori the truth that it expresses.

Intuitively this is a very promising line of argument and has echoes in the work of other writers. Salmon (1986), for example, argues that a priority applies primarily to propositions or the contents of sentences (1986, p133). Thus, an agent could know a priori that the sentence 'Stick S, if it exists, is exactly one metre long at t' is true, but not the content of the sentence - i.e. the fact that this stick is a certain length at a certain time. This could only be a posteriori, in Salmon's view (More of this later).

Along similar lines, Kaplan, in “Quantifying In” (1968), contends that in order to have a *de re* propositional attitude toward an entity an agent must be *en rapport* with it (In Linsky, Ed., 1971, p131) which would not seem to be the case with the ‘Julius’ example, or the Type 2 examples in general. Consequently, Kaplan goes on to state:

...I am unwilling to adopt any theory of proper names which permits me to perform a dubbing, in absentia, as by solemnly declaring ‘I hereby dub the first child to be born in the twenty-second century “Newman 1”’, and thus grant myself standing to have beliefs about that as yet unborn child (Ibid. p135)

Clearly, Kaplan, Donnellan, and Salmon are on a similar track here, as we might expect from such adherents to the direct reference theory. Despite the plausibility of these views, however, there are objections.

Returning to Evans, first of all, he agrees with Donnellan’s claim that the sort of statements we have been dealing with do not say anything informative about the world; but this is not the point for Evans, according to whom a knowledge of the reference-fixing description in these cases is sufficient for understanding the name in question. Consequently Donnellan misses the point of the puzzle, which is not how we can know a priori something informative or interesting about the world, but how we can know a priori something contingent, and hence, how something contingent can be uninformative (Op.Cit. p174).

Actually I don’t think Donnellan does misconstrue the puzzle - at least not as it is characterised by Kripke in *Naming and Necessity*. Although Kripke does seem to have second thoughts about the standard metre case on the question of whether it furnishes us with any information about the world (See 1980. p63n), when it comes to the ‘Neptune’ example (1980,p79n), which is more like the kind of example used by Donnellan and Evans, he definitely seems to hold the view that Leverrier could have known a priori that *Neptune* was the cause of certain perturbations, and it is this sort of claim against which Donnellan is arguing.

Forbes (1989) also takes issue with the Kaplan/Donnellan view that there is not one thing that is both contingent and known a priori. In the final chapter of his book Forbes objects to Donnellan’s treatment of the Type 2 examples as well as Kaplan’s

early view on the Type 1 example (“I am here now”). According to Forbes if we cannot know a priori the proposition expressed by the sentence “I am here now” because we cannot know a priori where ‘here’ is and what time it is, then an agent who is, say, lost in the jungle and does not know what time it is, would not express any thought if, on seeing an elephant charging toward him, were to think ‘I’m in danger here, right now’. Furthermore, if the agent were a radical amnesiac, then, according to Forbes, he would not even be able to grasp the proposition expressed by the sentence ‘I am the subject of this context’, which would be clearly absurd. (Op.Cit. p154). In Forbes’ view the agent would grasp these propositions, and in a *de re* sense, because the indexicals involved make these propositions or thoughts *about* the agent himself, his location, and the present time (Ibid,p 155). As for the Type 2 examples, Forbes concedes that the propositions in question are not *de re* with respect to whoever invented the zip. Nevertheless, knowledge of the reference-fixing definite description does enable one to express a non-metalinguistic thought with a sentence like (2), albeit a *de dicto* thought (Ibid.) Consequently, (2) will be contingent a priori according to Forbes.

Plausible though Forbes’ position may seem, I shall argue that he is mistaken on these points. It is just not true that on the so-called Kaplan/Donnellan account of the contingent a priori the agent in the jungle would not express any thought by ‘I’m in danger here right now’ or ‘I am the subject of this context’, for the following reasons. First of all, in the sentence ‘I’m in danger here right now’, the words ‘here’ and ‘now’ are strictly speaking redundant. All the agent need be aware of is the fact that he is in danger. This thought can easily be expressed, even by a radical amnesiac. Likewise with the thought expressed by ‘I am the subject of this context’. As far as the Type 1 examples are concerned Forbes has completely misunderstood the situation. Ironically, so has Kaplan, if he thinks that a statement like

(1) I am here now

could present us with a genuine example of the contingent a priori. This is not a genuine example because the words ‘here’ and ‘now’ are redundant. They are redundant because (1) uses the present tense. When using the present tense there is

no need to use the word 'now'. The present tense does not need to be augmented in this way. In fact the word 'here' is also redundant because (1) uses the present tense. Strictly speaking, (1) is no different from the statement "I am" or "I exist", because to say that I am is to say that I am or I exist at a place and at a time - i.e. we exist in space and time. We add nothing to the statement "I am" by adding the words 'here' and 'now', just as the person in the jungle adds nothing to the statement "I'm in danger" by adding the words 'here' and 'now'. Thus the Type 1 example reduces to the statement "I am" (or 'I exist'). But now we have to decide if this is a genuine example of a contingent a priori truth.

There are several ways in which we might proceed here. First of all, we might consider the claim that statements such as "I am here now" and "I exist" may not even be true - for example, when recorded on a tape recorder or when written on a piece of paper. Stephano Predelli (1998) makes a convincing case for just such a claim. Obviously if this is correct then we cannot know these truths a priori since they may not be true at all. One may wonder if some trick is being used in these cases, but it is difficult to see what this trick could be. We could question whether a recording of an utterance is a genuine utterance; or, even if it is, whether the word "I" is serving as a genuine indexical in these cases, and doesn't take on a sense something like "The person who made this recording" or "The person who wrote this note", but apart from this, there would seem to be examples of the Type 1 statements which we do not know a priori because they may not be true at all.

Another tack would be to look again at the claim made by Donnellan and Kaplan that there is not one thing that is both contingent and known a priori. This would seem to be more promising with the Type 2 and Type 3 cases than with the Type 1s but nevertheless it may have some application to the Type 1s. If we consider the statement "I exist", for example, we may well ask what it is that we know a priori (if anything). What we seem to know a priori is that, *pace* Predelli, the statement "I exist" is true, or will always be true on any occasion of utterance. But this could be seen as necessary. I.e. it will always be true - so long as the meanings of the words remain constant; and this is a qualification that must be made with any necessary truth. What is contingent here is that the particular person who makes the utterance exists. But we do not know a priori that any particular person exists, and so we do

not have a genuine example of the contingent a priori with the Type 1 cases. Against this it could be argued that surely *I* know a priori that *I* exist; but this is not so obvious. Perhaps I only know that I exist through experience; although, again, this does not seem to be straightforwardly the case either. Perhaps a statement like “I exist” is neither a priori nor a posteriori, but in a class of its own.

Forbes’ treatment of the Type 2 examples is more interesting and deserves detailed examination. First of all, it is not propositions or thoughts that are the primary bearers of modality in Forbes’ view, but states of affairs, where a state of affairs is defined as consisting in certain objects having particular properties and standing in particular relations (1989, p131). Propositions, which are ways of thinking of states of affairs, are primarily a priori or a posteriori and only derivatively necessary or contingent (Ibid.). Another way of putting this is to say that states of affairs are the referents of sentences, with propositions containing the mode of presentation of the states of affairs (p137). Thus with the sentences

Hesperus is a planet

Phosphorus is a planet

Forbes agrees with the Fregean that two propositions are involved, one of which could be believed while the other is disbelieved, but only one state of affairs, that of Venus being a planet. This furnishes Forbes with a neat way of avoiding Kripke’s modal objections to the description theory of proper names as follows. If we take for example, the sentences

(A) Aristotle was a teacher

(B) The teacher of Alexander was a teacher

we can see that (A) is not a necessary truth because it involves a contingent state of affairs, viz, Aristotle being the teacher of Alexander. This is the case even if

‘Aristotle’ has the sense of ‘The teacher of Alexander’ so long as the sense of the description is not identical with the sense of the name but merely “articulates its cognitive significance” (p133). (B) on the other hand, is necessary so long as the necessity operator is given the widest possible scope. This is because, according to Forbes, (B) involves a necessary state of affairs - that of someone’s being the teacher of Alexander. This is necessary in the weak sense that it obtains in every world where Alexander has a teacher - regardless of who that teacher is. The same state of affairs obtains.

So far Forbes’ position seems quite reasonable, but things are not so straightforward when he comes to deal with the problem of the contingent a priori, especially the Type 2 examples. Once again let us consider the ‘Julius’ sentences [using Forbes’ numbering]

(3) If anyone invented the zip, Julius did

(7) If anyone invented the zip, the zip-inventor did

For Forbes, the necessitation of (3) fails while the necessitation of (7) succeeds, though both are a priori, and it is the purpose of his semantics to account for this. What makes Forbes’ position especially difficult is the fact that he agrees with Evans that (3) and (7) have the same sense since (3) cannot be understood without ‘Julius’ having the sense of ‘The inventor of the zip’. The problem is to explain how two expressions which have the same sense can differ in their modal properties. Forbes does not regard this as a problem because, in his system, modality is a function of reference not sense. But now he does have a problem because, if (3) and (7) have the same sense, and sense determines reference (as a Fregean should accept) then the sense of (3) and (7) should determine the same reference - in this case the same state of affairs. But if states of affairs are the bearers of modality then (3) and (7) should have the same modality. Forbes, however, wants to argue that (3) is contingent and (7) necessary (on a wide scope reading).

Clearly (3) and (7) must have a different reference, and this Forbes accomplishes by abandoning the traditional Fregean mechanism whereby sense determines reference, in the case of descriptive names like 'Julius'. According to Forbes:

When we stipulate that a name is to have such and such a reference descriptively specified, we are associating the name with a reference of the kind ordinary names have by bestowing a descriptive (that is, non-de re) sense on it.....[I]f descriptive stipulation is in fact one of the ways in which the reference relation can be extended there is no need to have a de re sense associated with the name to provide its reference. So we can allow for a name whose reference is not the reference of its sense (Op.Cit. p161-162)

The upshot of all this is that the *sentence* (3) does not have the same reference as the *sense* of that sentence since the name 'Julius' has a different reference from that of the sense of 'Julius' (i.e. 'The individual, whoever it is, who invented the zip'). Consequently, while the proposition expressed by (3) is necessary and a priori, by virtue of the fact that it presents a necessary state of affairs (the same state of affairs as that presented by the proposition expressed by (7)), the sentence (3) is contingent and a priori because it refers to a contingent state of affairs, namely the state of affairs where a particular person invents the zip.

Clearly this is not a mechanism with which Frege would be familiar. In fact he might even regard it as perverse. It is certainly very odd to talk about the reference of a sentence or a name as distinct from the reference of the sense of the sentence or name. Sense is supposed to determine reference not *have* a reference. Indeed, how is reference determined here? Is the reference of the sense of (3) determined by another sense? And, if so, what would that sense be? It would have to be different from the sense that determines the reference of the sentence (3) if it is to determine a different state of affairs - but how could it be? It seems that if the senses of (3) and (7) are the same then we cannot escape the Fregean conclusion that they determine the same reference; in which case either both or neither is contingent.

Of course Forbes would reply that the strange mechanism he employs here only arises for the special case of descriptive names where a name is introduced into a language using a definite description to fix its reference. The fact that we are forced into this bizarre manoeuvre of having to distinguish between the reference of a

sentence and the reference of the sense of the sentence should tell us that we cannot really introduce a name into our language in this way. (More of this below)

Forbes seeks to dissolve the puzzle of the contingent a priori and demonstrate that it constitutes no threat to an empiricist concept of how knowledge is acquired by maintaining that while the proposition expressed by a sentence like (3) is a priori and necessary, the sentence itself is contingent and a priori - the contingency being due to the special way in which the name is introduced into the language.

This Forbes contrasts with Donnellan's claim that there is not one thing that is both contingent and known a priori, and also with Evans' contention that a statement like (3) is contingent (and a priori) but only 'superficially' contingent as opposed to 'deeply' contingent. According to Evans the puzzle of the contingent a priori arises due to our failure to distinguish these two types of contingency which normally coincide but which can diverge. A statement is deeply contingent if its truth depends on some feature of the actual world, and merely superficially contingent if it is true at some worlds and false at others - or, depending on how it embeds inside the scope of modal operators (Evans, 1979, p161). According to Evans, although (3) and (7) have the same sense, the necessitation of (3) would fail, making it contingent but only superficially so. As Evans explains on p185 of his (1979):

A deeply contingent statement is one for which there is no guarantee that there exists a verifying state of affairs. If a deeply contingent statement is true there will exist some state of affairs of which we can say both that had it not existed the sentence would not have been true, and that it might not have existed. The truth of the sentence will thus depend upon some contingent feature of reality.

(3) will be merely superficially contingent according to Evans, because its truth depends on the set of states of affairs

{ x's being the inventor of the zip; y's being the inventor of the zip; z's being the inventor of the zip... }

(3) is made true by the absence of every member of this set (which falsifies the antecedent), or by the presence of any member of the set (which verifies the

consequent). But this demands nothing of the actual world which is why knowing it to be true constitutes no paradox (Ibid.)

Forbes rejects Evans' distinction between these two types of contingency as a solution to the present puzzle. For Forbes there is no difference between the type of contingency exhibited by

(3) If anyone invented the zip, Julius did

and

(14) If anyone invented the zip, α did

where ' α ' is the zip-inventor's real name. In both cases it seems clear that contingency traces to the fact that who invents a certain thing is a contingent property (Op. Cit. p166).

Some might object here that Forbes is missing the point about how the name 'Julius' is introduced into the language but actually I think he does not; especially when we consider that in order to explain how (3) and (7) embed differently in modal contexts (while having the same sense) Evans supposes that 'Julius' is really behaving like 'The actual inventor of the zip' in (3), which gives him

(13) If some one person invented the zip, the actual zip-inventor did

It is indeed difficult, as Forbes says, to see how (13) involves a different kind of contingency from that of (14). So it seems that Evans' distinction is of no use in the present puzzle.

Is Forbes' position any better? We could certainly question his distinction between a proposition and a state of affairs, as well as his claim that states of affairs are the bearers of modalities. We must be careful when trying to decide what is to be regarded as necessary or contingent. For Salmon, for example, it is propositions that are contingent. For Kripke on the other hand (who does not want to deal with propositions), it is simply statements that are necessary or contingent - i.e.

necessarily *true* or contingently *true*. For Forbes it is states of affairs that are contingent in the sense that some states of affairs will obtain at some worlds and fail to obtain at other worlds. This is all very well, but we must surely avoid saying that states of affairs might have been different. All we can say is that a state of affairs might not have obtained. Another state of affairs might have obtained in its place. This suggests that states of affairs cannot be said to be contingent, just as they cannot be said to be true or false. They simply obtain or fail to obtain at any world.

As for the claim that propositions derive their modal status from states of affairs, let us once again consider the sentences

(a) Hesperus is a planet

(b) Phosphorus is a planet

According to Forbes we have two propositions here, since a speaker could believe (a) and fail to believe (b) (Op.Cit. p 130). There is, however, only one state of affairs involved, namely, the state of affairs of Venus being a planet. But consider Forbes' definition of a state of affairs on p131 of (1989) as consisting in "certain objects having particular properties and standing in particular relations". There may be only one object involved but there is more than one property since the property of being visible in the evening is different from the property of being visible in the morning. So we could actually have two states of affairs here; namely, that of Venus appearing in the evening and that of Venus appearing in the morning. On the other hand, if we accept Forbes' claim on p137 that states of affairs need not be static but can include processes, then perhaps there is just one state of affairs here; viz, the state of affairs whereby the planet Venus appears sometimes in the evening and sometimes in the morning. Certainly Forbes wants to claim that the sentences (a) and (b) involve one state of affairs - that of Venus being a planet; but this is not good enough, because, while it explains why (a) and (b) are contingent (in Forbes' terms) it does not explain why (a) and (b) can express different thoughts. Only the involvement of different properties can do this. If we accept this then we may as well simply talk about general propositions rather than states of affairs. In fact to

talk of only one state of affairs here is too redolent of Kaplan's notion of a singular proposition, or even of Salmon's idea of one proposition and two guises - views a Fregean like Forbes must surely avoid.

The difficulty in talking about states of affairs in connection with the present puzzle is further compounded by Forbes' distinction between static states (such as Venus being a planet) and processes (such as my crossing the street). Is a process, for instance, not merely a sequence of static states? If not, can we not simply claim that there is only one state of affairs - namely the process of the entire history of the universe? The individuation of states of affairs remains a difficult problem. As far as Forbes is concerned, we may as well abandon talk of states of affairs and return to the Fregean idea of a general proposition.

The consensus among philosophers is that the contingent a priori does not stand up to scrutiny. Even if there are examples of statements which are contingent yet known a priori they are harmless. Either they don't give us de re knowledge (Donnellan) or they trade on some innocuous notion of contingency like Evans' superficial contingency.

A slightly different approach to the problem is that taken by Bostock in his (1988). Although Bostock argues that Kripke's examples do not work, he wants to claim that there are, or could be, examples of the contingent a priori - albeit of a special kind. Bostock begins by pointing out that if Kripkean rigid names are admissible then there must be rigid predicates, even rigid sentences (1988, p351-354). 'Cordate', for example, could be a rigid predicate if it is introduced into the language as designating in all possible worlds only those objects which have hearts in the actual world. The description 'creature with a heart' would merely fix the extension of the term 'cordate', and not be synonymous with it. As Bostock says, "There is nothing which a thing must be if it is to be a cordate, other than simply to be one of the objects in question" (Op. Cit. p353). Armed with such rigid predicates we can apparently know a priori a conditional something like

If there is exactly one object that falls under the predicate, then it is necessary that there is no more than one

And yet, according to Bostock, this is not a necessary truth for it is guaranteed to be true only when evaluated with respect to the actual world and with respect to other possible worlds it will be false. It will be contingent when its antecedent is false, and yet anyone who understands the predicate in question surely must assent to the conditional statement (p378,379). It would seem, therefore, to be a priori.

This is clearly a different sort of example from Kripke's in that it involves predicates rather than names. It is also more 'artificial' than Kripke's examples. That is, Kripke's position seems to be that it would be quite natural for a name to be introduced into a language via a reference-fixing description. Bostock's move is not a natural one, but it would seem to be a theoretical possibility, and that is what concerns Bostock.

Prima facie, at least, Bostock seems to have a point here. It does seem possible that we could introduce a rigid predicate into our language using a description which fixes its extension but is not synonymous with it. Certainly we may want to talk about counter-factual situations where the creatures with hearts in the actual world lack this property or have some other property. I.e. we want to designate only those objects in another world who are cordates in the actual world. Problems arise, however, if we consider a world where there are creatures with hearts who do not exist in the actual world. According to our definition of the rigid predicate 'cordate' these creatures would not belong in the extension of 'cordate'. How, then, are we to designate these creatures? We cannot call them 'cordates' on the present view because 'cordate' does not *mean* 'creature with a heart'. Yet what else are we to call them? We could of course invent some other word to mean 'creature with a heart', but what is the point of that? Surely the thing to do would be to call creatures with hearts 'cordates' in any world, and use some other word for our rigid predicate. Even if we did this, though, things would not work in the way envisaged by Bostock. Whatever word we use as our predicate it would not simply have its extension fixed by the description 'creatures with hearts in the actual world'. As we have already seen in earlier chapters, any word that is part of our language must contribute to what speakers understand by any sentence containing that word. What else could our predicate contribute to understanding but the description which is supposed to merely fix its extension? There is no doubt that the predicate would be

synonymous with this description. What this means for Bostock's example of the word 'cordate' is that there *is* something that an object has to be in any world in order to belong to the extension of 'cordate'. It has to be one of the creatures who have hearts in the actual world. This simply re-introduces the 'actually' operator into the puzzle, which Bostock is supposed to have eliminated (p360-364) by treating it, not as a new modal operator, but as a device for indicating relative scope (p360). This re-emergence of the 'actually' operator means that we must return to either Davies and Humberstone's two notions of necessity or Donnellan's two types of contingency - both of which Bostock rejects. It also means that unless we get rid of the 'actually' operator we once again have the problem of having to say that, although 'p' is contingent, 'Actually p' is necessary - a view we have already rejected.

Bostock's argument fails, then, but it does bring us to another much simpler criticism of the attempt to generate contingent a priori truths. The criticism is simply that the distinction between descriptions which give a synonym and those that merely fix a reference cannot be made. As with a proper name, any word in our language must have a sense if it is to contribute to understanding. In the Type 2 examples the only sense available is that of the description which is supposed to be fixing reference. The putative name will be synonymous with the description. But we must be careful here. Remember our definition of a genuine name as that which actually *names* something. Remember also how Kripke has forced us to qualify the sense of a proper name - not 'The F', but 'That which is commonly believed to be The F'. We must conclude from this that the so-called 'names' in the Type 2 examples are not functioning as genuine names at all. This is because, first of all, the descriptions used by Kripke, Evans etc. must be conditionalized to avoid the Existential Complaint. By our definition of a genuine name the description which gives the sense of the name simply cannot be existentially conditionalized in this way. If we don't know if there is anything to be named here, then we do not know if our 'name' really is a genuine name. Secondly, it is at least highly suspect to give a sense to a word like 'Julius' of 'The person who is *commonly believed* to be The F'. The so-called 'names' in the Type 2 need not be regarded as genuine names at all, since we don't know if they really name anything; although we could regard them

as conditional names, or provisional names. Donnellan is wrong, therefore, when he argues at p57 of (1979) that “we are in the somewhat odd position of possessing a mechanism for introducing a name that rigidly designates something, but a mechanism that is not powerful enough to allow us to use the name!”

Actually Donnellan has things the wrong way round here. We can use these ‘names’ but only as abbreviations of descriptions, the way in which the police, for instance, might use a name like ‘The Jackal’ as a name for a criminal whose identity is unknown to them. This would simply make it easier to talk about whoever is responsible for the crimes in question. Such a ‘name’ would undoubtedly be an abbreviation of the definite description and would therefore have to be regarded as synonymous with it. It would certainly not be a name in the same way that ‘Kripke’ or ‘Nixon’ would be, because we know that we are designating something with these words, but we are not sure what, or even if, we are designating anything in the Type 2 cases. We are certainly not in a position to claim to have *baptised* anything in such cases, and so Kripke seems to be stretching the definition of a name too far here.

Although the Type 1 and Type 2 examples have failed to provide examples of the contingent a priori we are not yet finished with the puzzle. We have still to deal with the Type 3 examples involving measurement terms like ‘metre’. These form a separate class from the Type 2 cases for two reasons. First of all, when we introduce a term like ‘one metre’ into our language, we do not do it ‘blind’, so-to-speak, the way we do with ‘Julius’. That is, we can see that the object which is to serve as the standard for our system exists. We can therefore perform a straightforward dubbing ceremony which we cannot do with ‘Julius’. There is no doubt that we are causally connected to some object, although there may be problems with naming a *length* in all possible worlds rather than simply an object. Secondly, words like ‘metre’ or ‘kilogram’ are not obviously similar to proper names in the way that ‘Julius’ is. They are more like predicates and so the argument against the Type 1 and Type 2 examples may not apply to the Type 3 cases.

The passage in *Naming and Necessity* where Kripke introduces his standard metre example is by now well known. He begins on page 54 with a passage from

Wittgenstein where Wittgenstein claims that we cannot say of the standard metre stick either that it is a metre in length nor that it is not a metre in length because of its “peculiar role in the language game of measuring with a metre rule” (p54). Kripke finds this claim strange. He thinks the standard metre stick is obviously a metre in length, although it might not have been. It is surely a contingent truth that this object (which Kripke calls ‘S’) is one metre in length. Specifically, it is contingent that S is one metre in length at a particular moment in time - the moment when the term ‘one metre’ was first introduced into the language via the description ‘the length of S at t_0 ’. However, according to Kripke, the person who performs this reference-fixing ceremony is in a position to know a priori the contingent truth that S is a metre long at t_0 - simply by his reference-fixing act; and so we have an example of a contingent a priori truth.

Once again we have the crucial distinction between a definite description which does not give a synonym but merely fixes a reference, as well as the distinction between epistemology and metaphysics. The statement

(8) Stick S is one metre long at t_0

has the metaphysical status of a contingent truth because the stick’s length might have been different at t_0 . Its epistemological status is that of an a priori truth. Kripke further argues that, although (8) may look like a definition of the term ‘one metre’, it is only a definition in a special sense. That is, the term ‘one metre’ is not synonymous with the reference-fixing description ‘the length of S at t_0 ’. As Kripke puts it in *Naming and Necessity*, the person who performs such a ceremony is

...using this definition not to *give the meaning* of what he called the ‘meter’, but to *fix the reference*. (For such an abstract thing as a unit of length, the notion of reference may be unclear. But let’s suppose it’s clear enough for the present purposes). He uses it to fix a reference. There is a certain length which he wants to mark out. He marks it out by an accidental property, namely that there is a stick of that length... (p55)

This is a curious passage in that the person who introduces the term ‘one metre’ into the language seems to know what a metre is *before* he discovers the stick that is to

serve as the standard. There is no question of this happening, and the fact that Kripke thinks that it could shows that his example is probably misconceived from the start. But we must not be too hasty. Kripke's claims here are not obviously specious and we must examine them in detail.

Once again we begin with the distinction between descriptions that fix a reference and those that give a synonym, and once again Kripke must show that this method of introducing a term into the language furnishes us with knowledge that we did not already possess. For example, we already know a priori that stick S had whatever length it had at t_0 , even though it might have had a different length from its actual length at that moment. What Kripke must do is show why introducing the term 'one metre' in the way he suggests gives us any knowledge beyond that.

The Type 3 examples have certain advantages over the other types as well as suffering disadvantages. To begin with, the Existential Complaint does not seem to be as strong with the Type 3 cases since there is no reasonable doubt as to the existence of our standard. Also, there is no doubt that we are *en rapport* with the object (to use Kaplan's term) and so, if we have knowledge here, it would be knowledge *of* some object. Donnellan's criticisms of Type 2 cases would therefore not apply. Furthermore, it would seem that we cannot argue that the contingency involved here is merely superficial - i.e. that it does not depend on the way the world is, which would rule out Evans' critique. The major disadvantage of the Type 3 cases, and one which Kripke seems to acknowledge (1980,p63n) is that, even if we are dealing with a contingent a priori truth here, it would be relativised to a particular agent at a particular moment in time. No-one not present at the naming ceremony would know a priori that S is a meter in length at t_0 - although they might know that it *was* a metre in length at t_0 , after the event.

Perhaps a more serious problem with the Type 3 examples (and again one which Kripke concedes) is that we could seriously question whether we have acquired any genuine knowledge by merely stipulating that a certain length is to be a metre, or a certain mass is to be a kilogram. We certainly have not discovered anything here, and this in itself may be enough to dismiss the Type 3 cases as of no real consequence. The fact that a statement like (8) seems to be guaranteed to be true probably works against the Type 3 cases rather than for them. Wittgenstein's point

here could be that, if there is no possibility of a mistake here, then we cannot talk about knowledge. At any rate, I shall argue that Wittgenstein's analysis of the standard metre example, strange though it may seem to many, is in fact correct.

Another interesting difference between the Type 3 cases and the Type 2 cases lies in the way they have been treated by commentators. Where the treatment of the Type 2 cases is, on the whole, characterised by sophisticated and rather ingenious mechanisms for explaining the phenomenon of the contingent a priori, the treatment of the Type 3 examples betrays mostly confusion on the part of philosophers. Not only confusion as to how a name functions, but confusion as to the very concept of measurement.

Not surprisingly the confusion begins with Kripke in his original (1980) example. As already mentioned, Kripke seems to think that the person who fixes the reference of the metric system by stipulating that 'one metre' is to refer rigidly to the length of S at t_0 , knows what a metre is before he comes across the stick that is to serve as the standard. There is no question of this being the case. The closest we could come to such a situation would be if people are already using something called a metre but the length varies slightly in different parts of the country. Our agent merely wants to standardise the unit of length, and so he has *some* idea of the length he is looking for. He does not, indeed cannot, know exactly what a metre is prior to choosing the standard.

Van Brakel (1990,p300) highlights another major confusion, which originates with Kripke, but also pervades the commentaries of many other philosophers, such as Salmon (1986); Harrison (1987); Kennedy (1987); and Bostock (1988, Op.Cit.) The confusion involves the difference between knowing that S is exactly a metre in length and knowing that it is approximately a metre in length. If we are to know a priori that S is a metre in length then surely we must mean exactly a metre in length, not approximately. Measurements on the other hand are merely approximations. There is always a degree of tolerance in a measurement, which rules out our knowing that S is a metre in length by measurement because it surely makes no sense to say that the standard metre is a metre in length plus or minus a certain degree of tolerance. This rules out Salmon's (1986) argument where he claims that

we could only know the length of S by measurement. It also highlights a confusion in Kripke's original account where he asks "If the stick is, for example, 39.37 inches long, why isn't it one metre long?" If we are to go down this road then we introduce a posteriori knowledge into the puzzle because we can only know that S is 39.37 inches long by measurement.

Bostock (1988) fares no better. He claims (p373) that Kripke's metre example is not a case of a contingent a priori truth because someone could understand the description 'the length of that stick' and also know what is meant by 'one metre', based on a different standard from S, and yet not know that S is a metre long. This is a bad analysis of the puzzle for various reasons. First of all, we cannot have two standards for a system of measurement. Secondly, even if we count this second stick as a standard, we cannot know that it is exactly the same length as S - only approximately. And thirdly, the introduction of this other stick simply shifts the problem. That is, how do we know that the other stick is a metre in length?

Van Brakel's main thesis in his (1990) is an interesting one. It is that the statement

(9) One metre is the length of S

is a priori but necessary. He begins (p305) by dismissing the attempt to relativise the claim that S is a metre long to a particular moment in time as a useless way of defining a unit of measurement. We cannot restrict our unit to a single moment in time. He then goes on to claim that a statement like (9) is necessary because, not only is the term 'one metre' rigid, but so also is the reference-fixing description 'the length of S' (p307). According to Van Brakel such descriptions refer rigidly to a natural kind property such as

solid metallic rods of constant temperature and submitted to constant external forces have constant length. (p311)

Consequently, "Definitions of units of measurement are like Kripke's examples of theoretical identities, except that the identity of two rigid designators is stipulated a priori instead of discovered a posteriori" (p307,308). Thus Van Brakel uses

Kripke's own ideas on the necessity of scientific identities against his claims that statements such as (8) are contingent a priori.

In my opinion Van Brakel's analysis is a convincing one but it leaves untouched the question of whether our agent could know a priori that S is a metre long at t_0 . (Presumably Van Brakel would say that he could but it would be of no consequence) Also, Van Brakel, along with just about everyone else, has misunderstood Wittgenstein on the question of whether or not we can attribute length to the standard metre. On page 302 of (1990), while castigating Kripke for misunderstanding Wittgenstein on this point, Van Brakel also seems to miss the point when he says: "Wittgenstein means that in the practice of measuring in metres it does not make sense to start *measuring* S". I think it clear from the original passage in *Philosophical Investigations* that Wittgenstein means something more than this. What he means is that it does not make sense to ascribe length to the standard metre *in the metric system* because of its special role in that system. We can of course attribute length to it in some other system, but certainly not a priori.

There is more to the puzzle of the standard metre, then, than whether or not we can know its length by measurement. Even Salmon concedes this in his (1988) where he makes some concessions to Wittgenstein while realising that his earlier (1986) treatment of the problem may have been inadequate. In an appendix to his (1986) Salmon thinks it straightforwardly the case that to know that the standard metre is a metre in length is a posteriori - in fact he describes it as a paradigm of a posteriori knowledge, discoverable only by measurement (1986, p142). In the (1988) paper Salmon abandons this view, conceding that physical measurement is not only unnecessary but "is in some sense inapplicable to this case" (p207). Salmon then goes on to claim that a hypothetical agent who has randomly chosen S as his standard for measurement would know that S is exactly one metre long "simply by looking at it" (Ibid). This makes S "epistemically unique" (p209) and leads, according to Salmon, to an epistemological paradox as follows.

...as soon as we say that the reference-fixer knows that S is one metre long, we are embroiled in a paradox. The language-game of measuring with a metre rule involves a simple criterion for knowing how long something is. In order for the reference-fixer to know how long anything is, he must be able to

specify its length in metres *and* he must know how long the Standard Metre is. Saying that he knows that S is exactly one metre long attributes to him knowledge of exactly how long the Standard Metre is. But he could not have acquired this knowledge through measurement. If he has such knowledge, he can only have acquired it by simply looking at S. This would require S to be what it cannot be: knowable in a unique way in which no other object is knowable and in which it itself would not be knowable if it had not been arbitrarily selected as the standard. These considerations invite the sceptical conclusion that the reference-fixer does not know after all that S is exactly one metre long. This, in turn, leads to an even stronger sceptical conclusion. For if the reference-fixer does not know how long S is, he cannot know, and cannot even discover, how long *anything* is. Measuring an object's length using S only tells him the ratio of that object's length to the length of S. (p210)

The key word here is 'only', as in "only tells him the ratio of that object's length to the length of S". This passage shows that Salmon, in common with perhaps the vast majority of philosophers, simply does not understand the concept of measurement. This lack of understanding is presaged in a passage on p208 where Salmon argues:

...knowing that a given object's length is exactly n times that of another object (the standard) cannot give one knowledge of how long the first object is unless one already knows how long the second object is. If one knows only that the length of the first is n times that of the second without knowing how long the second object is, one knows only the proportion between the lengths of the two objects without knowing how long *either* object is.

What Salmon fails to understand here is that measurement simply consists in determining (within a degree of tolerance) the ratio of one object's length to the length of some standard. To say that an agent *only* knows the proportion of the length of two objects (one of which is a standard), or that he can *only* determine the ratio of an object's length to the length of S, is to fail to grasp the concept of measurement. Measurement consists in nothing more than the comparison of the object of measurement with some standard. The very question

How long is x ?

presupposes a unit of length and a standard of length. If there were no standards of length we could not ask how long anything is. (This standard need not be a very

accurate one like the standard metre stick. It can actually be quite crude. For example, if I say that a certain object is about three times my height then I use my body as a crude standard in order to give people some idea of the length of the object in question. The standard metre stick may be more accurate as a standard but the principle is basically the same). The question 'How long is x?' *literally means* 'How does x compare with some standard (within a certain degree of tolerance)?' The answer to the question, viz.

x is n units long

literally means

x compares n times with some standard (within a degree of error)

This standard is arbitrarily chosen and agreed upon by the community. Only practical considerations bar us from using anything at all as a standard. If we had chosen a different stick as our standard metre, as we might have done, then a metre would be a different length. This arbitrary nature of standards of measurement seems to be lost on many philosophers.

At this point we can give an argument in support of Wittgenstein by asking if it really makes any sense to ask how the Standard Metre stick compares with itself - within a certain degree of tolerance. It would seem that it does not and so it makes no sense to say of the standard metre either that it is or that it is not a metre in length.

Not only is the standard of measurement arbitrarily chosen but standards of measurement are 'artificial' concepts - as opposed to 'natural' concepts like 'tiger' or 'dog'. Although we invented the word 'tiger', we did not invent the concept of a tiger. This is something we discovered when we discovered tigers and decided to name them. Not so with the concept of a metre or a kilogram. Although we discovered the concepts of *length* (and mass) we *invented* the concept of a metre for our own convenience; as a means of making judgements about length which we could record and/or communicate to others. The length 'one metre' was not waiting

to be discovered by us in the way that tigers were. We simply chose a length that we found convenient and *called* it a metre. That is all there is to choosing a standard of measurement. This means that whatever we call a metre, or would have called a metre, is/would have been, a metre. Contrast this with a natural concept like 'tiger'. A tiger is not whatever is called a tiger. A metre *is* whatever is called a metre.

This foregoing leads to a consequence that is in many ways even stranger than Wittgenstein's claim. First of all, if we grant that it is legitimate to say that S is a metre in length at t_0 , but contingently so, then it is contingent for a rather surprising reason. It is not contingent because S might have had a different length at t_0 , but because S might not have been chosen as the standard metre. Furthermore, until S was actually chosen as the standard, it would not have been a metre in length - *even if it had exactly the same length that it had at t_0 !* This is a consequence of the artificial nature of the concept of a metre. Whatever we call a metre is a metre. If the metric system had never been invented then nothing would have been a metre in length because nothing would have been judged to be a metre in length. If this seems strange then consider what is happening at the moment t_0 when S is chosen as the standard metre. Is S already a metre in length at t_0 ? And was it, or could it have been, a metre in length before t_0 ? The answers to these questions may seem obviously 'Yes'; but that is a mistake. If S was a metre in length at t_0 when it was chosen as the standard, or even before t_0 , as most people would probably claim, **then it must have been the most remarkable coincidence in human history when our agent chose S as the standard metre.** Just think, of all the sticks he could have chosen as the standard for the metric system, he just happened to choose one that was exactly one metre in length. Astonishing! Astonishing, and of course nonsense. (It is certainly very difficult to square this with the arbitrary way in which standards are chosen). Not only is nothing *called* a metre without the existence of the metric system, but nothing actually *is* a metre prior to the establishment of the metric system. Again this supports Wittgenstein because at the moment that the metric system is established -i.e t_0 - we could question whether it makes sense to say that anything is a metre in length.

There are several reasons, then, to conclude that the Type 3 cases do not give us examples of contingent a priori truths. First of all, in common with the Type 2 cases, the Type 3 examples such as

Stick S is one metre long at t_0

depend on the distinction between definite descriptions which are synonymous with a term and those that merely fix reference. We have seen that this distinction cannot be made. Even if we were to grant the distinction, however, there are other reasons for rejecting Kripke's claims. There is Van Brakel's argument that (8) is necessary because it involves two rigid designators, not one. Then there is the infamous claim by Wittgenstein that we cannot say of the standard metre either that it is a metre in length or that it is not. This has far more plausibility than it may appear at first sight; especially if we regard S as a criterion by means of which we judge whether or not an object is a metre long. *It surely makes no sense to apply a criterion to itself, at least in the present case, and so it makes no sense to judge or assert that S is a metre in length or that it is not a metre in length.*

We must also ask what is to be *understood* by the term 'metre'. Once again, if this is to be a word in any language, it must contribute to what speakers of the language understand by any sentence in which it appears. The only thing that could be understood by the word 'metre' is the description 'The length of S at t_0 (or in standard conditions)' which is supposed to merely fix reference without giving a synonym. As we have seen there is no such thing as fixing a reference without giving a synonym. Consequently the term 'One Metre' must be synonymous with the above description.

To reiterate, the question:

How long is x?

literally means

How does x compare with some standard (within a certain degree of

tolerance) ?

and nothing more. Likewise, the statement

x is n units long

literally means

x compares n times with some standard (within a degree of tolerance)

The question “Is the standard metre a metre in length at t_0 ?” presupposes the metric system and literally means:

How does the standard metre compare with itself (within a degree of tolerance)?

This does not make sense. Consequently, the person who asks the question:

How long is the standard metre?

does not understand the concept of measurement. It is not a proper question in the way that ‘What time is it on the sun?’ is not a proper question if we use the sun to tell the time. There is no answer to an improper question and so we cannot say either that S is a metre in length nor that it is not.

Even if we grant that it does make sense to say that S is a metre in length, the statement

S is a metre in length at t_0

would be necessary a priori because of the meaning of the statement. The statement would literally mean

S compares with itself (the standard) at t_0

This is undoubtedly a priori but also necessary. Once again we have failed to find an example of a contingent a priori truth.

CHAPTER SIX

THE EPISTEMOLOGY OF NECESSITY

II

ESSENTIALISM

We turn finally to perhaps Kripke's most interesting claim - the claim that there are necessary a posteriori truths and that science can discover real essences empirically. This in itself is a bold and highly controversial claim. What is even more controversial is Kripke's contention that this essentialism can be established from a semantic theory about how names function in natural languages, viz, the theory that proper names and names for natural kinds (such as 'gold' or 'water') are direct, rigid designators. Thus, as we have already seen, statements such as:

- (1) Cicero is Tully
- (2) Hesperus is Phosphorus

will be necessary even though they were empirical discoveries. More interestingly, and more contentiously, so-called 'scientific identities' such as

- (3) Water is H₂O
- (4) Gold is the element with atomic number 79
- (5) Heat is the motion of molecules

will be necessary a posteriori (See Naming and Necessity, 1980, p122 ff.) As if this were not enough Kripke also adds essentialist claims about the origin of individuals. For example, if a certain object is made of ice, then it is essentially made of ice; or if a certain individual developed from a particular pair of gametes, then it is necessary

that that individual came from just those gametes (1980, p110-113). Putnam (1973) makes similar essentialist claims about natural kinds, based on his view that the reference of natural kind terms are fixed indexically and socially and not by means of a concept in the minds of speakers of any language which sets conditions for reference in terms of phenomenological properties. Putnam's views are supported by Donnellan in two unpublished papers (1973 and 1974) - according to Salmon (1981, p163ff). On the other hand, there is a considerable body of literature opposing these essentialist claims, not least of which is Salmon's (1981), even though Salmon is himself a strong advocate of the direct reference theory.

Before evaluating the various arguments for and against this Kripkean essentialism we must be clear as to what exactly is involved in the essentialist claims. Some obvious misgivings are easily dealt with by Kripke - for example, the claim that natural kinds like gold and water do not exist in all possible worlds, and so we cannot be dealing with necessary truths here. Kripke makes it quite clear that he is dealing with a particular notion of necessity - what he terms 'weak necessity': i.e. truth in all possible worlds *where the relevant object or kind exists*. This is sometimes known as *metaphysical necessity* by Kripke and Putnam. Another early qualification that must be made is an epistemological one in answer to the possible criticism that, for example, water cannot be necessarily H₂O because it might turn out not to be H₂O at all. I.e. we could have made a mistake in our empirical investigations. Again Kripke deals with this objection briskly, by arguing that of course the claim that a statement such as 'Water is H₂O' is necessary depends on its being true at all - and of course it may not be. In other words, *if* water is H₂O then it is necessarily so. This is of course because 'water' and 'H₂O' are rigid designators and therefore refer to the same substance in all possible worlds, where the relevant substance exists. (The expression 'relevant substance' is of course crucial to the debate). The notion of necessity with which we are concerned, then, is a qualified one. We are not concerned with worlds where the relevant object or kind does not exist, and we must grant that we could be mistaken about what we take to be empirical discoveries about the world. This latter point means that no scientist could ever actually be in the position to say that he has discovered the essence of a kind because he must concede that he might have made a mistake in his observations: but

this does not seem to trouble Kripke. (We shall see that these qualifications should trouble Kripke. He should be particularly troubled by a further qualification, the one mentioned in chapter 3 to the effect that we really do not need proper names or natural kind terms such as ‘water’ and ‘gold’ in our language. We could use descriptions instead. What this means for the doctrine of Essentialism is that statements like (1) - (5) above will be necessary only if we decide to admit names like ‘Hesperus’, ‘Phosphorus’, ‘water’, ‘gold’ etc. into our language. This is a further weakening of the notion of necessity being employed here. It is a further qualification that Kripke must make, and I contend that it is a qualification too far)

As with the alleged examples of the contingent a priori in chapter 5 we will divide the proposed examples of the necessary a posteriori into different types. Type 1 will consist of identity statements involving proper names such as ‘Hesperus is Phosphorus’ or ‘Cicero is Tully’. The Type 2 cases will be the scientific identities mentioned above [(3) - (5)] ; and the Type 3 examples will be the ones involving claims about the essentiality of origin. The three types are united in that they all involve what Kripke claims are rigid designators and what could broadly be regarded as identity statements. There are important differences, however. First of all, in the Type 1 examples, we are dealing with what seem to be straightforward identities between objects - the kind dealt with in Chapter 4. In the Type 2 cases by contrast, we are dealing with kinds rather than individuals - i.e. general essences rather than individual essences [See Parsons, 1969, in Linsky, 1971,p73ff] Even though we could perhaps regard kinds as ‘higher order’ objects (e.g. Marcus, 1976) it is not obvious that we are dealing with straightforward identity statements in the Type 2 cases. In a statement like

“Heat is the motion of molecules”

for example, we could be dealing with the ‘is’ of composition rather than the ‘is’ of identity. This would have advantages and disadvantages for Kripke, as we shall see.

The other major difference between the different types is that, where we deal solely with proper names in the Type 1 cases, we deal also with descriptions in the other two types - i.e. descriptions that are supposed to be rigid according to Kripke.

This brings problems for the Kripkean view in that it is not always obvious that a description could be rigid, even if we accept the rigidity of names (which we may not). The problem is that, *pace* Almog, while a description like “The square root of 2” may be rigid, or even “The motion of molecules”, to claim that the description “The offspring of gametes G” is rigid is surely to beg the question at issue.

In this final chapter we will mostly be concerned with the Type 2 cases which seem to be the most interesting (and of course the most contentious). We shall say something about the Type 3 examples but very little about the Type 1s because of what was said about identity statements in chapter 4. In this chapter, remember, the claim that identity statements of the form $a=b$ are necessary was seriously challenged, and found wanting. Obviously, if identity statements are not necessary, then we have dealt with the Type 1 cases. We may also have dealt with the Type 2 cases, but we must be careful here since they do not seem to be quite the same sort of claim as the Type 1s. (Although, interestingly, Kripke seems to think that they are).

What also counts against the necessity of the Type 1 examples is the theory of meaning developed in chapter 3, whereby, even if proper names are rigid, they must have a sense given by a set of definite descriptions in order to contribute to what it is that speakers understand by the name. Thus, a statement like:

“Hesperus is Phosphorus”

is equivalent to something like:

“The heavenly body commonly believed to occupy such-and-such a position in the evening sky is identical with the heavenly body commonly believed to occupy such-and-such a position in the morning sky”

On Kripke’s own premises, this would be contingent. The only way in which such a statement could be necessary would be if the descriptions involved were always to be taken referentially (in Donnellan’s sense) and never attributively. But this would

be equivalent to rigidifying all definite descriptions, which would simply destroy the distinction between rigid and non-rigid designators.

We must also be clear about what kind of language we are dealing with. As mentioned in chapter 3 we are dealing with what we have called The Scientific Language - i.e. the language of scientists and philosophers, or indeed anyone who uses language in a serious way to explore the world or describe the world. We are not interested in the language of the man in the street because of its sloppiness. We may have given up long ago on the idea of the logically perfect language but we must be as strict as possible in constructing our semantics, especially when we remember what is at stake.

Of course the theory of meaning developed in chapter 3 also affects the Type 2 and Type 3 cases, as we shall see below. In the meantime it will be as well to note the magnitude of the task Kripke and Putnam have set themselves. Problems with essentialism are well known - the most common being the argument that what we regard as necessary or contingent is relative to our interests or our way of describing the situation at hand. Thus, Quine, in his (1960) famously makes sport with essentialism with his paradox of The Mathematical Cyclist. I.e. if it makes sense to say that a mathematician is essentially rational but not essentially bi-pedal; and a cyclist essentially bi-pedal but not essentially rational, what are we to make of a mathematician who is also a cyclist? He would seem to be essentially rational and not essentially rational. Or is he essentially rational but not essentially two-legged - or vice versa? (1960, p199) We do not seem to be able to make any sense here at all.

Even mathematical essentialism has been attacked by some philosophers. Although it may seem obvious that the number 9, for example, is essentially composite (in a *de re* sense) this has been challenged. Harman (1970), for example, argues against arithmetical essentialism in the following way. First of all, numbers can be identified with sets, and so the number 9 can be identified with a set (in fact numerous sets). But, according to Harman, “.. being a composite number is not an essential property of any set. Therefore if numbers can be identified with sets, and *de re* necessity is in question, no number is necessarily a composite number. Being a composite number is not an essential property of any number” (1970, p184)

One more example, and perhaps the most famous, is Quine's treatment of the problem in "Reference and Modality" (1953, 1961). In this celebrated paper the point is again about whether we can make sense of the claim that objects have essential properties irrespective of how they are described. Specifically, Quine is concerned with the problem of quantification into what he terms 'referentially opaque' contexts. For example, from the premise

(A) 9 is necessarily greater than 7

and

(B) The number of planets = 9

by substitution of co-referential terms we get

(C) The number of planets is necessarily greater than 7

which is false according to Quine. The problem is that we cannot go by existential generalisation from (A) to

(D) $(\exists x)$ (x is *necessarily* greater than 7)

because we cannot specify whether 'x' is to be 9, which would make (D) true, or 'The number of planets', which would render it false. In other words, the way in which the number 9 is specified determines which properties are to be regarded as necessary or contingent. As Quine says in "Reference and Modality", quantified modal logic requires that

An object, of itself, and by whatever name or none, must be seen as having some of its traits necessarily and others contingently, despite the fact that the latter traits follow just as analytically from some ways of specifying the object as the former traits do from other ways of specifying it. [In Linsky, 1971, p30]

Thus Quine concludes that the only way in which we can adopt an essentialist position here would be if we were to indulge in an unjustifiable 'favouritism'

whereby we only consider the object under a certain specification while ignoring others.

There have of course been many counters to these anti-essentialist views. Plantinga, for example, in his (1974, Ch.2) takes issue with Quine's Mathematical Cyclist as well as Harman's set-theoretic arguments. As for "Reference and Modality", Smullyan (1948) famously offers a Russellian reply, which Quine really should accept - since he is himself a Russellian - while Marcus (*Mind*, 1960), also in a Russellian vein, claims that (B) is not a genuine identity, but a weaker equivalence relation, and so does not admit of substitution of 'The number of planets' for '9' in (A). For Kripke, of course, the alleged paradoxes of people like Quine are dissolved by the notion of rigidity. It is only rigid designators that can be substituted in intensional contexts. '9' is rigid, 'The number of planets' is not; and so there is no problem.

Regardless, however, of the merits or otherwise of Marcus, Smullyan and Plantinga, or indeed of the replies to these by Quine (See *Words and Objections*, 1967), there are still problems for essentialism concerning natural kinds. A notable anti-essentialist of recent times is John Dupre who argues against essentialism in his (1993) *The Disunity of Science*. In this book the point is made that what is to count as a natural kind depends on how we decide to characterise things. This is of course the point made by Quine but in a less abstract way. Interestingly, despite this, Dupre claims to be a realist about natural kinds - but only in a special sense - what he calls "promiscuous realism". For Dupre the naturalness of kinds is a matter of degree (p63). Cedars, for example, are not closely related, according to Dupre. The classification of various trees as cedars has more to do with the use the wood is put to by carpenters than any biological taxonomy (p31). Nevertheless, cedars form a natural kind even if they are classified in this way, but such a kind would not be as natural as some others, like oak or beech, which are easier to classify (Ibid.). As Dupre explains:

Certainly I see no possible reason why commitment to many overlapping kinds of things should threaten the reality of any of them. A certain entity might be a real whale, a real mammal, a real top predator in the food chain,

and even a real fish. Many, perhaps all, of these designations might be the appropriate characterisation of that object for some legitimately scientific purpose (Op. Cit, p262)

Such 'promiscuous' realism would undoubtedly undermine the doctrine of essentialism, but that is Dupre's point. Or at least it is part of his point. The other part of Dupre's thesis concerns the role that a real essence would play in determining the properties and behaviour of the objects possessing them (p63). For Dupre, we cannot explicate such a role because there are no kinds the members of which are identical with respect to *all* their properties. Thus real essences are not necessary for natural kinds (Ibid.). Furthermore, even if a kind *is* determined by a real essence, the discovery of the essence *presupposes* the discovery of the kind (Ibid.). As Dupre himself argues:

Only the most extreme reductionist could suppose that examining a particular individual would allow one to determine to what kind it belonged apart from the prior recognition and at least partial characterisation of that kind (Ibid.)

Thus Dupre concludes:

...discovering kinds does not involve discovering essences; and so, given that there is no other way of discovering them, nothing does. (p64)

It is against this background of scepticism about essentialism that Kripke and Putnam must argue. Their claims, particularly Kripke's, are bold because they believe they have developed a powerful theory of how natural kind terms (and others) are to be explicated. Specifically, they are concerned with how the extension of a natural kind term is fixed. Armed with the concepts of rigid designation, necessity of identity and the indexical nature of natural kind terms whose reference is fixed causally by ostensive definition of a paradigm case, Kripke and Putnam feel confident of establishing necessary a posteriori truths.

As we have seen Kripke's notion of rigidity has many problems. However, an interesting development of the rigidity issue is to be found in Danielle MacBeth (1995). In this paper MacBeth makes a distinction between what she terms 'semantic' and 'epistemic' rigidity. Natural kind terms are not semantically rigid

because they are predicates (p260), and as such have essentially associated principles of inclusion (p277). But, argues MacBeth, empirical investigation does not require that a principle of inclusion be given in advance. Rather, we can *stipulate* that a paradigmatic instance is to be an instance of some kind (p275). For example, we can choose a paradigm of a cat and stipulate that it is a cat such that:

In such a context the assertion ‘Felix is a cat’ is not a claim that is true or false depending on the facts about Felix and the principle of inclusion for ‘is a cat’ but is rather a matter of fixing on a particular as a paradigm instance of the kind so that the principle of inclusion for that kind can be investigated (p275)

This is the notion of epistemic rigidity which is supposed to allow us to investigate natural kinds.

The similarities with Putnam’s method in “The Meaning of Meaning” are obvious and unfortunately beset with the same problems. The main problem as I see it with this method of choosing a paradigm in order to avoid any principle of inclusion or a set of descriptive conditions for membership in the kind is that we must be able to *re-identify* the paradigm, especially in the context of a scientific investigation where experimental results must be *repeatable*. The only way in which we can re-identify the paradigm or repeat our experimental results is by giving descriptive conditions or a principle of inclusion for the kind. This is borne out by Mellor (1977) who points out that the above method we have attributed to MacBeth and Putnam is certainly not how biologists actually proceed. For Mellor any paradigm cases are “causally downwind” of the usage they are supposed to constrain (In Pessin and Goldberg, 1995, p74). The archetypes are chosen to fit botanical and genetic knowledge, not the other way round (Ibid.) In fact, in Mellor’s view it is a consequence of Putnam’s thesis that our most authoritative specimens of a kind may not even be of that kind, since we cannot be sure if our current usage stems from a paradigm that is indeed of the appropriate kind. [A similar point is made by Zemach in his, 1976, p123-124; and by Dupre Op.Cit., p63, where, as we have already seen, he contends that even if a kind *is* determined by a real essence, the discovery of such an essence *presupposes* the discovery of the kind. It is also interesting to note that Kripke sees no problem with the possibility of a kind being discovered and

rediscovered many times so that we may not know what the original sample of the kind was]

Even a defender of Putnam like Kim Sterelny recognises a role for descriptions in what he terms the 'grounding' of natural kind terms (1983, in Pessin and Goldberg, Op. Cit.). This is in response to the criticism of causal theories (e.g. Mellor, Op. Cit.) that some natural kinds have no paradigm instances - for example, theoretical particles or elements high in the periodic table. For Sterelny, different natural kind terms are grounded in different ways, depending on the way in which we causally interact with the kinds involved. With something like gold, for example, the grounding of the term is relatively straightforward and similar to the method advocated by Kripke and Putnam - i.e. a speaker is introduced to an actual sample of gold and his use of the term 'gold' grounded in causal contact with the paradigm sample. His subsequent uses of the term will, according to Sterelny, *re-ground* the term because "For many of his uses of that ...term both samples of the substance and his ability to refer will be immediately involved. Such tokens will be semantically linked to gold in the same way that his grounding use of "gold" is so linked" (p102, 1995).

Things are obviously more complicated when no paradigm sample is available; such as the case of a theoretical particle. In such cases Sterelny envisages two possibilities. First of all, a kind term can be introduced by descriptions that pick out the *causal powers* of the kind. For example, kind K is that kind which has such-and-such effects. Such descriptions would be used referentially not attributively to avoid having to conclude that the term refers to nothing if we do not get the original description exactly right. In the second type of case the causal powers of the speculated kind are not known - for example, the case of the neutrino. In such a case the descriptions used to introduce the term are initially use attributively, but once the effects and interactions of the kind are detected experimentally, samples of the kind can be designated by referential descriptions. Thus, according to Sterelny, a causal theory of natural kind terms can be given which is more sophisticated than the early theories of Kripke and Putnam, but which concedes nothing to the Fregean - even though descriptions play an important role.

Unfortunately for Sterelny his arguments seem prone to the same problems as the previous anti-Fregean theories. To begin with, it surely makes no sense to talk of the re-grounding of a term like 'gold' without some means of re-identifying samples. This must involve recognition of properties in some way. Even if it is only the experts who are required to be aware of the recurring identifying conditions, we are still dealing with properties. As Mellor puts it in his (1977) "...no doubt the labour of reference is divided, as Putnam says, but it may be a Fregean labour for all that (1995, p73). Sterelny replies (Ibid. P101) that we cannot be dealing with a fully Fregean labour here because there is no Fregean account of reference-borrowing available. I suggest that the arguments put forward in chapters two and three of the present work, especially the theory of meaning developed in chapter three, show that a description theory of names does not require any account reference-borrowing.

As for Sterelny's account of kind terms where no paradigm is available, it would seem to be susceptible to the same counters already made against the Kripkean distinction between descriptions that are synonymous with a term, and those that merely fix its reference. As we have already seen, this distinction is untenable because of the need to specify what is *understood* by a term. In addition to this, Sterelny seems to overlook the possibility of two kinds having the same effects in the context in which they are first identified. In such a situation would we succeed in fixing the reference of the term?

So far, then, the anti-essentialist appears to be winning the argument; but there are defects here too. For example, we could question Dupre's claim that kinds involve degrees of naturalness. Classifying different trees as cedars, for example, because they all have a common use as timber, does not seem to be a natural way of classifying at all. Even if we accept that distinguishing species can be difficult, as with the approximately 290,000 species of beetle, for example, the problem is surely not insurmountable. Similarly with Dupre's claim that we cannot explicate a role for essence in determining properties and behaviour because there are no kinds the members of which are identical with respect to all their properties. This is too strong a claim to make on the basis of such a trivial point. Of course no two members of a species will be absolutely qualitatively identical, but we can surely live with this.

The main point to remember here is that all the normal members of a species will be identical with respect to the *type* of properties each of them has. This is something that certainly could be explained by appeal to the notion of essence.

The degree of naturalness alluded to by Dupre is even less obvious if we deal with non-biological kinds, like gold or water, or even heat; and even Dupre must accept that many biological kinds are easily identified. Actually it is a defect of Dupre's thesis that he only deals with biological kinds. Non-biological kinds seem to be relatively easy to classify; although there could be a problem with isotopes, as even Putnam admits (1990, Chapter 4). But even if there is a problem here it is nothing like the difficulty presented by the classification of biological kinds. Nevertheless, Putnam does see difficulties in dealing with two samples of iron, one of which is mono-isotopic, the other 'natural' iron which contains more than one isotope (Ibid, p68) Rather surprisingly, Putnam concludes that whether or not we classify the two samples as consubstantial *may depend on our interests!* (Ibid.)

Putnam may be right about this point, but there are other ways of looking at the problem. If mono-isotopic and poly-isotopic iron, or whatever, have a different micro-structure and behave differently - albeit only slightly and in highly specialised circumstances - then there seems little doubt that, strictly speaking, they are different substances. The fact that the layman, whose interests do not coincide with those of the scientist, does not make the distinction, or is unaware of the distinction, is not relevant. Strictly speaking they are different substances because they are constituted differently and behave differently. What more could we wish for in our attempt to distinguish different kinds of thing?

Of course these conditions would only provide criteria for sameness of substance in the actual world. The key question now is whether sameness of microstructure together with same behaviour provide identity conditions for kinds across possible worlds, which is what is required by essentialism. Interestingly, we can find a negative answer to this question from Putnam himself in (1990, Chapter 4) where he begins to question his essentialist catechism. In Putnam's later view microstructure is no longer the sole criterion for consubstantiality (even in the actual world). If it makes sense to talk of laws of nature then we must make use of laws in our attempt to define consubstantiality (1990, p68,69). Thus Putnam concludes that "A and B

are the same substance if and only if they obey the same laws” (Ibid.) But, argues Putnam, we cannot use this as a criterion for substance identity across *possible worlds* because there could be possible worlds **where the laws of nature are different** (Ibid.) This seems correct because, if there is a possible world where H₂O exists but because the laws of nature are different it does not behave like the substance we call ‘water’ - i.e. it is not drinkable, does not boil at 100 degrees, dissolve salts etc. - it is far from obvious that this H₂O would be water. Consequently, and crucially, Putnam distances himself from Kripke with the words:

I now think that the question “What is the necessary and sufficient condition for being water *in all possible worlds?*” makes no sense at all. And this means that I now reject “metaphysical necessity” (1990, p70)

This is a crucial admission by Putnam and it would be of great interest to hear Kripke’s response to it. Significantly, perhaps, Kripke has failed to comment in public - as far as I know. The only thing he could do is appeal to intuition and claim that he has an intuition to the effect that it *does* make sense to ask for the necessary and sufficient conditions for substance identity across possible worlds. Such an intuition would be severely tested by the admission that the laws of nature can vary across possible worlds, so that we could have worlds where H₂O has *none* of the properties we associate with water and some other substance, like Putnam’s XYZ, has *all* the properties we normally associate with water (except perhaps the property of being separable into hydrogen and oxygen in the ratio 2:1). Is it really a straightforward case to say which of these substances is water?

On the other hand, even if we accept that obeying the same laws is a criterion for substance identity in the actual world (and even here the notion of ‘same laws’ could be vague as Putnam concedes [p68]), this may not be a criterion for consubstantiality across possible worlds. So even if there are metaphysically possible worlds where the laws of nature differ from the actual world, perhaps Kripke can still maintain that anything that is H₂O in these worlds is water.

Where does this leave the doctrine of essentialism about natural kinds, then; or the Type 2 cases as we have chosen to call them? It seems that Putnam is correct in

his claim that there will be possible worlds where the laws of nature are different. At the very least this puts a severe strain on the claim that anything that is H₂O in such worlds is water. It may even bring about the downfall of metaphysical necessity. It is ironic perhaps, but Putnam's argument here is actually more powerful than any of the anti-essentialist positions mentioned above. Putnam has not given up on necessity, however. He still accepts that a statement like 'Water is H₂O' is physically necessary, and that the laws of nature exist objectively, and that science can discover these laws and this kind of necessity (1990, p68). If Kripke does not accept Putnam's argument here he would do well to consider the arguments against the project to derive non-trivial essentialism from the theory of direct reference to be found in Salmon's *Reference and Essence* (1981) where the charge against Kripke et. al. is once again that of begging the question.

In this very thorough and well-argued book Salmon deals with the attempt by Putnam and Donnellan to derive non-trivial essentialism about natural kinds from DRT, and with Kripke's attempt regarding individual essence. Salmon characterises the Putnam/Donnellan attempt to derive non-trivial essentialism in the following schema, which he terms (after Donnellan) a 'mechanism' for 'generating' necessary a posteriori truths.

The mechanism for water is as follows (using Salmon's original numbering):

(13') It is necessarily the case that: something is a sample of water if and only if it is a sample of *dthat* (the same substance as *this* is a sample of)

(18') *This* (liquid sample) has the chemical structure H₂O

(19') Being a sample of the same substance as something consists in having the same chemical structure

Therefore

(23') It is necessarily the case that: every sample of water has the chemical structure H₂O

More generally:

- (15') It is necessarily the case that: Something is a (bit of) v if and only if it is an instance of *dthat* (the same K -kind that *this* is an instance of)
- (25') *This* has the ζ - property of being ξ
- (26') Being an instance of the same K -kind as something consists, at least in part, in having the same ζ property that the given thing has
- (28') Therefore it is necessarily the case that: every (bit of) v is ξ (p166-169)

Where v is a common noun, e.g. 'water'

K is a generic natural kind predicate

ξ expresses a hidden structural property

ζ expresses the property of being a hidden structural property of a certain sort

dthat is Kaplan's indexical operator which transforms any given singular term into a rigid designator (In French, et.al.,1979)

According to Salmon a similar mechanism suggests itself for generating necessary a posteriori truths concerning individuals - what he calls the **I** - mechanism. For example, the **I**-mechanism for generating the necessary a posteriori truth that a certain table, if it is in fact made of wood, is essentially made of wood, is as follows (again using Salmon's original numbering):

- (30') It is necessarily the case that: Woody is *dthat* (the table located *here*)
- (31') The table located here was originally constructed from hunk of wood **H**
- (32') Being the very same table as something consists, at least in part, in having

the same material composition

Therefore

(34') It is necessarily the case that: Woody, if it exists, was originally constructed from hunk of wood **H**

where 'Woody' is a proper name and rigid designator of a particular table. This generalises in a similar way to the natural kind case.

So according to Salmon, the programme to derive non-trivial essentialism from the theory of direct reference takes the following form. We begin with a premise stating that a natural kind term, or a name for an individual, is a rigid designator. This is followed by an empirical discovery that members of the kind have a certain chemical composition, or the individual has a certain origin. Thirdly, we have the premise which states that being the same *x*, or whatever, consists in having the same chemical composition, or the same material origin, or in the case of a biological species, in being of the same biological class as the *x*. All of which allegedly leads to the conclusion that the *x* is essentially composed of whatever it has been found to be composed of, or that it is essentially part of the biological class (e.g. mammal) that we have discovered it to be part of. The crucial premise is the third, which Salmon claims is a non-trivial essentialist assumption - a metaphysical assumption not derived from the theory of direct reference or any semantic theory. It is because of this that the Putnam/Donnellan program fails. In Salmon's view it is quite simply question-begging.

Interestingly, according to Donnellan, the third premise is not a metaphysical assumption, but a scientific discovery, and hence, a posteriori. He reasons that, since there was a time when nothing was known about, say, chemical composition, and hence it could not have been known that consubstantiality consists in having the same chemical composition, that our current knowledge concerning consubstantiality is a posteriori, since it depends on the empirical discovery that substances do in fact have a chemical structure. So the claim to essentialism concerning natural kinds has two sources of a posteriority according to Donnellan.

(1981, p.165,166) I think Donnellan must be wrong here. Of course the third premise depends on the empirical discovery that substances do in fact have a chemical composition, but once this is discovered, it requires a priori reflection (or stipulation) to conclude that consubstantiality consists in sameness of composition. So if Salmon's characterisation of the program to derive non-trivial essentialism from the theory of direct reference is correct it would indeed seem that the program fails. It fails because the alleged necessary a posteriori truths generated are not generated solely from the theory of direct reference in conjunction with empirical premises free from non-trivial essentialist import. As Salmon says: "Our essentialist intuitions aside, the theory of direct reference seems perfectly compatible with the claim that examples such as 'water is H₂O', 'Tigers are mammals', and 'This very table originated from hunk of wood H'...are contingent a posteriori rather than necessary a posteriori" (1981, p217)

Kripke's arguments for the essentiality of origin are similarly rejected by Salmon. These arguments take the form of essentialist claims about the origin of artefacts, such as a table, or the biological origin of an individual, like Nixon, for example. Could a certain table which is in fact made of wood have been made from a block of ice? asks Kripke in "Identity and Necessity". Similarly, could Nixon, or anyone, have developed from gametes other than the ones from which he in fact developed? Kripke answers these questions in the negative. Salmon does not of course necessarily disagree with these essentialist conclusions - only with the claim that they can be derived from the theory of direct reference in a non-question-begging way. Once again, according to Salmon, Kripke assumes a non-trivial essentialist premise something like:

(V) If it is possible for a table, x, to be originally constructed from a hunk of matter y, then necessarily, any table originally constructed from hunk y is the very table x and no other

which is something an anti-essentialist would deny. Now it seems that an anti-essentialist would indeed deny (V), but actually I think Kripke would also deny it because this is not what he is claiming at all. Of course a lot depends on what is meant by 'hunk of matter y'. If it is big enough then surely more than one table

could be made from it and no-one would want to assert (V). What Kripke means here is that if a certain table comes from a certain hunk of matter, then it is necessary that *it* could not have come from any other hunk of matter, regardless of how many *other* tables could have come from it.

In fact it seems that Kripke may be on stronger ground in his claims about individual essence than with his claims about kinds, although it remains to be seen if even these claims stand up to scrutiny. As far as the Kripkean claims to non-trivial essentialism concerning natural kinds are concerned, however, it seems that Salmon is right. The crucial third premise in the schema to derive non-trivial essentialism about kinds does indeed seem to be an unjustified, and in the present context, unjustifiable metaphysical assumption. After all, as Dupre points out, there could be kinds that encompass radically different structures (1993, p26). The contention that consubstantiality across possible worlds consists in sameness of microstructure would seem to be no more than a prejudice on Kripke's part. In short, DRT does not so much establish essentialism about natural kinds as presuppose it - a much-vaunted criticism.

Returning to the Type 3 cases, the individual essences, it could be that Kripke is indeed on firmer ground here. Not only is he not relying on the necessity of identity but neither does he deal solely with names. In the case of the table, for instance, indexicals are the order of the day, and indexicals seem to be uncontroversially rigid. The question for Kripke here is 'Could *this* table, which has its origin in a certain piece of wood, have had its origin in something else?' He still uses examples involving proper names, but as with the 'table' example, we are again involved with necessity of origin rather than straightforward necessity of identity. The most famous example, perhaps, is:

Nixon is the offspring of gametes G

This is a posteriori, according to Kripke, yet necessary due to the rigidity of 'Nixon' and 'the offspring of gametes G'

To deal with the latter first, it obviously suffers from the defects of DRT already discussed. 'Nixon' is a proper name and we have seen that it is highly doubtful

whether proper names are rigid (Kripke certainly has not established this). And to say that the description 'the offspring of gametes G' is rigid is simply to beg the very question at issue. Furthermore, it was not an empirical discovery that Nixon or anyone else developed from some particular gametes, in the sense that no-one has observed the actual ones. What was an empirical discovery was that human beings in general developed from a pair of gametes, and Nixon is presumably no exception to this. What we must question here is whether it is necessary that human beings developed from one sperm and one egg. Surely the world could have been such that people developed from several ova and several sperm which all fused together at some point, each contributing some genetic material. The question now becomes: "Could this be a world in which Nixon exists?" It is certainly not obvious that Nixon would not exist in such a world unless it is necessary to the existence of human beings, as human beings, that they develop from just two gametes like in the actual world. This would seem not be so. Surely the entire human race could have developed from more than two gametes - and this includes Nixon. So the biological origin of an individual is not necessary for that individual to be that very individual. Of course a certain genetic code may be necessary for Nixon to be Nixon, regardless of the origin of the code, but that is another question.

The example of the wooden table is perhaps more interesting in its use of indexicals rather than names. Indexicals are undoubtedly rigid (if anything is) and so the expression 'this table' will designate the same object in all possible worlds (where the object exists). What is interesting here, though, is that even the rigidity of the indexical involved does not establish that the table is made of wood in all possible worlds. Slote (1974), for example, challenges Kripke's position here by claiming that the table could have originated in a block of wood *w*, which before being made into the table, changed into silver then back into wood 'steadily and lawfully' both ways. If the table was made during the silver stage and later turned into wood, then surely, argues Slote, it only accidentally originates from wood (1974, p8). Perhaps a more palatable possibility would be to imagine the block of wood turning into charcoal rather than silver. This is certainly possible if the block was caught in a fire and roasted (There could of course be metaphysically possible worlds where wood turns into something other than charcoal when roasted). In this

case the table would have been made of charcoal, which is not wood - although it is derived from wood. Kripke's only reply here would seem to be to say that the wood still has the same origin in every possible world - i.e. the same tree, the same sapling, and ultimately the same seed. But this is to shift the problem from determining the origin of the table to determining the origin of the block of wood, and ultimately the tree from which the wood came. This just brings its own problems. For instance, the tree could have been turned into charcoal before the block of wood was hewn from it. And as for the tree having its origin essentially in a particular seed, this brings us back to the problem of the essentiality of the origin of individuals in a certain pair of gametes. Along similar lines, the tree need not have its origin in a single seed at all. The genetic material required to make the tree could have converged from several sources. Even if we decide that this genetic material is necessary for this tree to be *this* tree we are a long way from establishing that a certain table is essentially made of wood! There is also a problem with the claim that genetic code is essential to an individual or a species. For instance, if the genetic code involved in producing Nixon had been slightly different (so that the resulting individual had a different colour of hair or eyes, for example) would the result be a numerically distinct individual? Or if the DNA of tigers had been such that tigers did not have stripes, would we be dealing with a numerically distinct species? It would seem not. In fact it is difficult to see how genetic code can constitute the essence of an individual or a species since it can change, or part of it can change. What kind of essence has parts, and what kind of essence can change? What is clear is that defining the essence of a living thing is very different from the question of what constitutes the essence of an inanimate kind like gold. This is a distinction that Kripke does not recognise.

Kripke's only possible response to the case where the table turns into charcoal would be to say that we no longer have the same table - because the table designated as *this* table is a wooden one. Now this may again be question-begging on Kripke's part, but even if it is not, we should consider the following. First of all, the table is the same object (numerically) even after it has been turned into charcoal (or whatever) in the same way that Nixon, who was once a cherub-faced child, would be the same object (numerically) even when he is a wizened old man. Secondly, if

we designate the wooden table as *this* table then it is *this* table that was caught in the fire and turned into charcoal (or whatever). This is a consequence of the rigidity of *this*. If not then we must question the rigidity of the indexical here. If the word 'this' is a rigid designator - and it would seem to be - then ironically it is this rigidity which helps to show that the table is not made of wood in all possible worlds.

So it seems that it is not necessary that an object have the origin that it in fact has; even if we accept the doctrine of rigid designation! What is even clearer is that a species, like the tiger for example, need not have the same origin in all possible worlds; although, curiously, Linsky (1977) thinks otherwise. According to Linsky, if astronauts were to visit Mars and discover creatures that looked and behaved exactly like tigers, and even had the same internal structure as tigers, nevertheless they would not be tigers because they could not have had the same origin as our tigers (1977, p80). Now it is just wrong to say that the Martian creatures could not have had the same origin as our tigers. Why could they not have had the same origin? But even if they do not have the same origin, if they look like tigers, behave like tigers, have the same internal structure and DNA as tigers, and could interbreed with tigers, then why are they not tigers? The difference in origin is surely irrelevant here. If it is not irrelevant then do we conclude that these creatures are not mammals because they do not have the same origin as our mammals; even though they are warm-blooded and give birth to live young? Surely not. Once again it is properties that determine what is a tiger or not. Origin is in no way necessary.

Finally we return to the theory of meaning expounded in chapter 3 and how it affects the claims of Kripke et. al. regarding the necessary a posteriori - especially the Type 2 cases. The new theory has some very interesting consequences for the Type 2 examples. It means for instance, that even if terms like 'water' or 'gold' or 'tiger' are rigid, they must have a sense given by a set of definite descriptions or properties which are commonly believed to belong to the kinds in question (and which speakers of the language use to refer each other to what they believe to be natural kinds). It follows from this that, even if it is necessary that, for example, water is H₂O, it will be just as necessary (in some sense of the word) that water has the phenomenological properties it has in the actual world (in standard conditions).

So, in general, if microstructure is necessary, so is macrostructure. This is just what we should expect when dealing with scientific discoveries since it is a basic tenet of science that microstructure determines macrostructure in a law-like manner; or that macrostructure is a function of microstructure. And yet to say that macrostructure is just as necessary as microstructure (in some sense of 'necessary') would be anathema to Kripke and Putnam (at least the early Putnam). This merely shows how little DRT understands the concept of a natural kind.

To see why DRT does not understand the concept of a natural kind or the meaning of a natural kind term, consider the examples mentioned above. That is:

Water is H₂O

Gold is the element with atomic number 79

Heat is molecular motion

For Kripke and Putnam these are necessary truths (in some sense of 'necessary'). But the question we must ask here is not just whether it is necessary that water is H₂O etc. but whether it is *sufficient* for anything to be water that it be H₂O. It seems that as far as Kripke and Putnam are concerned it must indeed be sufficient. That is, for anything to be water in any possible world it is necessary and sufficient that it be H₂O. The same goes for gold. For anything to be gold in any possible world it is necessary and sufficient that it be the element with atomic number 79. This must be the case for two reasons. First of all, if we really are dealing with identities here then whatever is on one side of the identity sign must be necessary and sufficient for whatever is on the other side. Secondly, if it is not sufficient for the kind to have the appropriate microstructure, then something else must be necessary. But what else could be necessary on the Kripke/Putnam view? Nothing, it would seem. If we did decide that something other than the microstructure was necessary for anything to belong to a natural kind it could only be the phenomenological properties of the kind (plus perhaps its law-like behaviour); but according to Kripke and Putnam, these properties are neither necessary nor sufficient for anything to belong to a kind; or, at least, for the application of the kind term. As far as Kripke and Putnam are concerned, the phenomenological properties of a natural kind merely serve as a

guide to what kind we are dealing with. They serve as what Putnam calls “operational definitions” (1975, p232). Similarly, for Kripke, these properties can be used to fix the reference of a natural kind term, without giving a synonym for the term (1980, p131)

What this means for the Kripke/Putnam view is that phenomenological properties are neither necessary nor sufficient for something to be a member of a natural kind. This seems to have two extremely counterintuitive consequences. First of all, if the phenomenological properties of a natural kind are neither necessary nor sufficient for the application of the appropriate term then how can these properties provide *any* kind of indication that we are dealing with the particular kind we believe ourselves to be dealing with? Secondly, and more bizarrely, if it is necessary and sufficient for anything to be water or gold, or whatever, that it have a certain microstructure, it would not be necessary that it had *any* macrostructure, or phenomenological properties. This is unintelligible, and cannot be what Kripke and Putnam intend; but it does seem to be a consequence of their view - a consequence they have not foreseen.

Another consequence is that the so-called ‘scientific identities’ may not be identities at all. This is because what is on the right hand side of the identity sign is at best only necessary for what is on the other side, and not sufficient as an identity would seem to require. So a statement like ‘Water is H₂O’ is not a necessary identity between two rigid designators, since it is not any kind of identity. Of course we still seem to be dealing with a genuine statement here, and so the question remains as to whether we can discover necessity empirically.

The best that Kripke and Putnam can argue here is that it is not necessary that a natural kind have the macrostructure that it has in the actual world (This does seem to be what they claim). They cannot argue that a natural kind need have no macrostructure at all. So it is necessary that a natural kind have some macrostructure - but what macrostructure? Could water have had a completely different macrostructure so long as it is H₂O? Could it have been a blue solid that was poisonous to humans (where humans are constituted in exactly the same way as at present)? On the Kripke/Putnam view that would indeed be a metaphysical possibility. In fact, according to this view, natural kinds could have had absolutely

any macrostructure whatsoever. But this is at least very difficult to accept if we are dealing with scientific identities. No scientist would accept that water or gold could have had absolutely any phenomenological properties whatsoever, with the same microstructure, for the simple reason that science believes that microstructure *determines* macrostructure - in a law-like manner. Water and gold have the macrostructure they have *because* of their microstructure. Macrostructure is a *function* of microstructure as far as science is concerned. The strange thing here is that macrostructure is a function of microstructure as far as Kripke and Putnam are concerned. As Putnam says in (1983,p73)

...The use of a word such as 'gold' depends on our possessing *paradigms*, standard examples that are agreed to be model members of the kind. ...What makes something gold is having the same nature as the paradigms; in current physical theory this is unpacked as having the same composition, since it is the atomic structure that determines the law-like behaviour of a substance.

Now if macrostructure is a function of microstructure there must be some kind of necessary connection between them. This need not be what Kripke calls 'metaphysical necessity', but it must be something like physical or causal necessity. If there is no necessary connection between microstructure and macrostructure then how can it be a law of nature that microstructure determines macrostructure? It is ironic, but in trying to establish that science can discover the true nature of natural kinds, Kripke and Putnam undermine the basic tenets of that science! They don't seem to have realised this, but it makes a considerable difference to the claim that statements such as :

Water is H₂O

Gold is the element with atomic number 79

Heat is the motion of molecules

are necessary. Even if they are necessary - i.e. even if it is necessary that a natural kind have a certain microstructure - *it will be equally necessary that it has the macrostructure it in fact has*. This is the only way in which we can account for the relation between the microstructure and the macrostructure. The only way that

Kripke and Putnam can deny that it is necessary that kinds have the phenomenological properties they have in the actual world (in standard conditions) is to deny that macrostructure is a function of microstructure. This they cannot do without alienating the scientific community they claim to be aiding. If anyone wants to object that it is *metaphysically* possible that our natural kinds should have different phenomenological properties in different possible worlds we need only reply that the meanings of our natural kind terms could have been different, because they must have some meaning.

Another point that must be dealt with here is the way in which a natural kind term such as 'water' is introduced as a rigid designator of a certain substance. As we have seen this is done by choosing a paradigm sample and performing a kind of baptism whereby we stipulate that *this* stuff here (the paradigm sample) is water, for example. Anything else will be water if and only if it is the same substance as the paradigm. As Putnam says in "The Meaning of 'Meaning'", any prospective sample of water must bear the relation same_l to the paradigm sample. This not only applies to the actual world but also across possible worlds. Now as far as the claim to the necessary a posteriori is concerned we have already seen that Putnam, at least, no longer accepts that sameness of microstructure can be regarded as a criterion of consubstantiality across possible worlds, but there is another way in which we can question the claims to necessity here, and that is to question whether it is even necessary that the paradigm sample has the microstructure it in fact has. We do this by constructing an argument similar to the one whereby we questioned whether the table was essentially made of wood. Along similar lines we can question whether the paradigm sample of water is necessarily H₂O - assuming that it really is a sample of water and that it really has chemical structure H₂O. I.e. in the same way that the table could have undergone some kind of transformation so that it was no longer made of wood, the water in our paradigm sample could have been transformed into a substance which does not have chemical composition H₂O. For example, there are metaphysically possible worlds where the hydrogen in the sample is fused together to form helium and so we have a mixture of helium and oxygen instead of a mixture of hydrogen and oxygen. The laws in this world could even be such that this mixture looks and tastes like water, but this is not important. The main

point here is that H₂O might not have been H₂O because it could have undergone a transformation. We are still dealing with the same sample however. The rigidity of the indexical 'this' guarantees it. So *this* sample of water, which is in fact H₂O, might not have been. If the paradigm sample is not necessarily H₂O then it can hardly be a necessary truth that water is H₂O.

The meaning of natural kind terms is of crucial importance in the present context - i.e. in deciding whether or not we have discovered necessary truths empirically. According to Kripke and Putnam meaning is given by reference, to paradigm instances, which leads to the conclusion that there are indeed necessary a posteriori truths. As we have seen this is untenable. Meaning must be given in terms of what speakers of the language *understand* by the term in question, which does not seem to lead to any essentialist conclusions. The distinction between the two theories of meaning is obvious when applied to most of the previous examples, but there are some cases where Kripke seems to be on much firmer ground with regard to the necessary a posteriori. These are such identities as

(P) Heat is the motion of molecules

(Q) Light is a stream of photons

It is more difficult to argue against Kripke here because it may not be clear what is to be understood by 'heat' or 'light' in the same way that we can specify what is to be understood by 'water' or 'gold'. Furthermore, there does not seem to be a straightforward distinction between microstructure and macrostructure in (P) and (Q). It is obvious what is to count as the macrostructure of water or gold, but what is the macrostructure of heat or light? In *Naming and Necessity* (p131,132) Kripke is quite confident that although we pick out heat by the fact that it causes us to experience a certain sensation, nonetheless the term 'heat' cannot be *synonymous* with a phrase such as 'that which causes sensation S' because heat might not have caused this sensation. The distinction between descriptions that give a synonym and those that merely fix reference seems to be secure here because it is surely correct to claim that human beings might not have been affected by heat (or light) in the way

in which we in fact are. So does Kripke have an example of the necessary a posteriori in a statement like “Heat is the motion of molecules”?

Actually he does not, and the reason for this brings us to the very heart of the problem of establishing the necessary a posteriori. First of all, even if it is not obvious what a term like ‘heat’ or ‘light’ is to mean, what is clear is that *something* must be understood by them - otherwise they are simply not part of any language. A term like ‘heat’, therefore, must have a sense which speakers grasp when they understand any sentence containing the word; and the only candidate for this sense is the description

That which causes sensation S

Or, more precisely,

That which is commonly believed to cause sensation S

If this is accepted then the statement

Heat is the motion of molecules

reduces to

That which is commonly believed to cause sensation S is the motion of molecules
which is not a necessary truth.

Now this foregoing is just what Kripke is at great pains to deny. He would object that ‘heat’ cannot be synonymous with the description ‘That which causes sensation S’ because what we call ‘heat’ might not have had this effect on us. Now it is undoubtedly true that heat might have affected us differently, or not at all; just as water or gold might have affected us differently. But, I contend, this is irrelevant to the present debate. The fact is that heat causes us to experience a certain sensation,

and so, as a matter of fact, the word 'heat' (or whatever word we use) has to mean 'That which causes sensation S', or 'That which is commonly believed etc.' *because there is nothing else for it to mean.* The Kripke/Putnam requirements on meaning are too strict. They constantly argue that natural kind terms cannot be synonymous with a list of properties because it is logically possible that the kind in question should lack some or even all of the properties it has in the actual world. But this is irrelevant. There is no need to deal with what is logically possible here because when we say that, for instance, tigers are striped, we are not dealing with a logical truth. No-one wants to claim that 'Tigers are striped' or 'Gold is yellow', or whatever, are logical truths; so what is the point in talking about what is logically - or even metaphysically - possible here? What is relevant here is that it is a fact (as far as we can see) that natural kinds have certain properties. Furthermore, it is a fact that natural kind terms must be understood in a certain way, and this can only be in terms of the properties they are believed to possess. Thus in some sense of 'synonymous' natural kind terms are synonymous with lists of properties, contrary to the claims of the direct reference theory.

However, Kripke may have one card left to play, and that concerns the necessity of self-identity. There is a way of reading Kripke which leads to the conclusion that the necessity of identities such as 'Hesperus is Phosphorus' or 'Heat is the motion of molecules' is to be derived from the necessity of self-identity. In "Identity and Necessity", for example, we find the following:

If names are rigid designators then there can be no question about identities being necessary....in every possible world a and b will both refer to the same object x, and to no other, and so there will be no situation in which a might not have been b. That would have to be a situation in which the object we are now calling 'x' would not have been identical with itself (In Moore (Ed.) 1993, p181)

Later in the same paper we find Kripke discussing the necessary a posteriori with regard to heat and the motion of molecules...

...as Bishop Butler said “everything is what it is and not another thing”. Therefore, “Heat is the motion of molecules” will be necessary not contingent.... (Ibid., p188)

Yet another indication that Kripke derives his essentialism about natural kinds from the necessity of self-identity is to be found on page 133 of (1980) with the words:

We have discovered a phenomenon which in all possible worlds will be molecular motion - which could not have failed to be molecular motion, because that's what the phenomenon *is*

It seems clear, then, that Kripke derives his essentialism from the necessity of self-identity and that he regards the denial of the necessity of scientific identities as tantamount to the denial of the necessity of self-identity. (Salmon makes the same point in chapter 3 of *Reference and Essence*). In the ‘heat’ example it is easy to sympathise with Kripke on this point. After all, if heat is molecular motion in the actual world, then how can it fail to be molecular motion in any possible world if it is to be that very thing in those worlds?

Now I do not wish to deny the necessity of self-identity, but that is not the point at issue here. What is at issue is whether we have discovered empirically something which could not have been otherwise. So what exactly did we discover when we discovered that heat is molecular motion? We certainly did not discover the self-identity of heat - i.e. we did *not* discover that heat is heat or that molecular motion is molecular motion, just as we did not discover empirically the self-identity of the planet Venus when we discovered that Hesperus is Phosphorus. These things were known a priori. In fact they can *only* be known a priori. We simply cannot discover self-identity empirically, so how can Kripke, or anyone, derive the necessary a posteriori from the necessity of self-identity? To attempt such a thing is like trying to square the circle.

So what did we discover when we discovered that heat is the motion of molecules? It seems that the only thing we could have discovered empirically is that the phenomenon that causes us to experience a certain sensation turns out to be molecules in motion. There is nothing else to be discovered empirically here. This fits exactly with what was said previously about the meaning of a term like ‘heat’, and so, on Kripke’s own premises, we have not discovered anything necessary here.

The main problem with the necessary a posteriori regarding natural kinds is how to *individuate* kinds. The Fregean does so according to phenomenological properties, which does not lead to de re necessity. Kripke and Putnam claim to have refuted this position and to be able to individuate kinds by reference to causal contact with paradigm instances. These instances are, according to Putnam

...given existentially and not by criteria.... Actual things, whatever their description, which have played a certain causal role in our acquisition and use of terms determine what the terms refer to.... (1983, p73)

However, as we have seen, this is a seriously mistaken view. The key phrase here is “Actual things *whatever their description*” because what this implies is that, if it does not matter how a type of thing is described, then it does not matter how it is *experienced*. But it does matter how we experience things in the world because we must experience them in order to pick them out. The Kripke/Putnam view whereby we fix the reference of a natural kind term by choosing a paradigm sample of the kind regardless of how it is experienced or described is actually a naive view for two reasons. First of all, we can only be in a position to choose a paradigm because we recognise, or believe we recognise, certain properties. Secondly, we must be able to *re-identify* the paradigm, and this can only be done in terms of properties. This is a very important point when we are involved in a scientific investigation where experimental results must be communicated to other scientists, and, crucially, experiments must be repeatable. We can hardly publish a paper in a scientific journal proudly announcing that we have discovered that *this* stuff is H₂O, or whatever. We must say what it is that we have discovered to be H₂O and this involves us in the meanings of our terms - meaning which can only be given in terms of the properties we believe the kinds to possess.

The failure of the Kripkean essentialist program has many sources. First of all, the failure to appreciate that referring terms, be they singular or general, must have a sense given by a set of definite descriptions or a list of properties that are commonly believed to belong to the kind in question. Secondly, the failure to realise that we

cannot discover necessity empirically, if by 'necessity' we mean the necessity of self-identity. This is something which just cannot be discovered empirically. Thirdly, the failure to be clear about what is *meant* by an identity statement. Fourthly, the failure to realise that, even if we choose paradigms in the way advocated by Kripke and Putnam, we must be able to re-identify the paradigm, which involves us in descriptions. And fifthly, the failure to recognise that even the paradigm may not have its microstructure necessarily - if we mean metaphysically necessary. (All this is in addition to the problems mentioned previously due to Dupre, Salmon, and Putnam (1994)). The problem for the direct reference theory is that it takes identities like:

Hesperus is Phosphorus

Heat is molecular motion

Water is H₂O

to be basically statements of self-identity (if it says anything at all about their content) Or at least their core content is a proposition expressing the property of self-identity. Hence the necessity. To see that self-identity does not lie at the core of these statements we need only look at what must be *understood* by these sentences. They cannot possibly be understood as statements of self-identity because they can be the subject of speculation, conjecture and debate - debate which can be resolved by empirical investigation. This proves that these statements in no way involve anything as simple and trivial as self-identity because self-identity is not something that could be the subject of such speculation, conjecture and debate (controversial debate at that). Nor could a debate about self-identity be resolved empirically. The content of such statements must be given in terms of what speakers of the language must understand by the appropriate sentences. What must be understood by these sentences will depend on the situation being described; but whatever it is, it will not be anything necessary.

Our scientific identities may of course be physically necessary, but that is something very different from the metaphysical necessity mooted by Kripke and the early Putnam. In fact physical necessity may not be any kind of genuine necessity at

all, since the laws of nature are discovered by inductive generalisation, which brings its own problems. The so-called Laws of Nature may be mere descriptions of apparent regularities devoid of any normative force whatsoever. On the other hand, we do seem to be reaching a deeper level of reality when we discover chemical composition or DNA structure - a level previously hidden from us - and so perhaps there is something to the realist position after all. What is clear, however, is that whatever we discover empirically, it will not be necessary in the metaphysical sense. What is even clearer is that essentialism cannot be derived from any semantic theory. As we saw in chapter 3, if the sense of a natural kind term must be given in terms of what speakers of the language believe, then questions of necessity would hardly seem to arise for what we have called the 'Scientific Language'. We must also remember that, strictly speaking, we could devise a language which only used descriptions and no proper names or natural kind terms at all. So Kripke would have to qualify his thesis to the effect that a statement like

Water is H₂O

is necessary a posteriori, *if* it is true, and where necessity is taken in the weak sense. And, *if* we accept terms like 'water' and 'H₂O' into our language. This I suggest is a qualification too far.

It could be that no metaphysical doctrine can be derived from a semantic theory - or at least from a theory of naming. In the words of Joseph Almog, "Naming is naming and necessity is necessity".

CONCLUSION

The Direct Reference Theory (DRT) states that proper names, and possibly natural kind terms, refer directly to their designata and not by means of any Fregean descriptive sense which determines their reference. Rather, reference is determined by a causal chain leading from use of a term back to an initial baptism of an individual, or, in the case of a natural kind term, a paradigm sample of the kind. For Kripke names are not only direct designators, but rigid designators, which leads him to conclude that there are necessary a posteriori truths and contingent a priori truths.

We have seen that these claims have not, and cannot be upheld. Although DRT does have some cogent points to make against the traditional theory (or what it takes to be the traditional theory) some kind of description theory must be accepted - which destroys the claims to the necessary a posteriori and the contingent a priori.

The chief advocate of DRT is Kripke in his (1972, 1980) where he proposes three arguments against the description theory - The Modal Argument, The Epistemological Argument, and The Semantic Argument. To illustrate, consider the example:

(*) Aristotle taught Alexander

where the name 'Aristotle' just means 'The teacher of Alexander'. In Kripke's view this would have the consequence that (*) is necessary, a priori and analytic; whereas it would seem to be almost a paradigm case of a statement that is contingent, a posteriori and synthetic. This seems to be correct and provides at least a strong *prima facie* case against the description theory.

However, as we have seen, the Modal Argument trades on an illicit shift of scope. Kripke's strongest argument is the Semantic Argument which he illustrates with his 'Godel' example. This forces us to modify the senses of names but not to abandon them. From now on the senses of proper names and natural kind terms must be given in terms of what is commonly believed by the speakers of the language. (It is difficult to say whether this belief is *de dicto* or *de re*. Probably we should say that

it is de dicto in the first instance, and may turn out to be de re on closer examination. Or perhaps we simply cannot talk about de re beliefs, just as it is very difficult to talk about de re necessities). This leads us to conclude that sense determines reference in that it is speakers who refer by evoking concepts in the minds of other speakers. I.e. speakers refer *each other* to what they take to be things in the world. This is all there is to reference.

DRT makes many fundamental mistakes. It seems to be question-begging in several ways. Salmon (1981) highlights an example of this in DRT's attempt to derive non-trivial essentialism from a semantic theory. And Almog (1986), correctly it would seem, points out that the crucial notion of rigidity actually presupposes essentialism. Then there is Kripke's basic 'intuition' that proper names behave differently from descriptions in modal contexts (and probably non-modal contexts), from which he concludes that they cannot be synonymous. This is question-begging, because *if* names are synonymous with descriptions then they do not behave differently in these contexts.

There is a serious 'use/mention' confusion in DRT - especially regarding proper names. Examples can be found in Putnam (1975), Kripke (1980), Marcus (1976), and Evans (1973).

As far as the claim to the necessary a posteriori is concerned Kripke and Marcus do not understand what is meant by an identity statement. Marcus' (1947) 'proof' of the necessity of identity must be judged a failure. (See Ch.4) In addition to this there is a way of regarding identity statements as contingent.

DRT's most serious mistake is to fail to deal with the notion of understanding. Understanding is without doubt the most important concept in the philosophy of language because without understanding there is no language. A referring term like a proper name or natural kind term can only refer if it is part of a language. This means that it must be understood in some way; which means that it must have a sense which speakers can grasp. It follows from this that sense is at least necessary for reference. The best that Kripke and Putnam can do is show that sense is indeterminate. This means that reference is indeterminate, so we are hardly in a position to talk about metaphysical necessities being derived from a theory of how

names refer. In fact strictly speaking, we may not know what we are talking about most of the time!

Sense is also sufficient for reference since it is speakers who refer by referring each other to the world (or what they take to be the world). Thus, sense determines reference, and names are synonymous with descriptions given in terms of what is commonly believed by the speakers (It also means that there is a sense in which meanings are in the head, but not in the way advocated by Searle). This profoundly affects the project to establish the existence of necessary a posteriori truths and contingent a priori truths in the following way. First of all, there is no distinction between descriptions which fix reference and those that give a synonym. If a name like 'Neptune' is to be a part of a language it must be synonymous with the description that Kripke claims to be merely fixing reference. This puts an end to the claim that there are contingent a priori truths involving proper names (Although Kaplan's examples involving indexicals may be on more solid ground). Secondly, if terms like 'water' or 'gold' or 'heat' are synonymous with descriptions then statements like

Water is H₂O

Gold is the element with atomic number 79

Heat is the motion of molecules

are either definitely not necessary, or the question of their necessity does not arise since we are always dealing in beliefs of speakers.

To say that speakers refer is not to make a distinction between speaker's reference and semantic reference. Speaker's reference just is semantic reference. I.e. it is language that refers but only in the sense that it can be used by speakers in a certain way.

There is no distinction between referring and using a term correctly. If we use a referring term correctly we will be referring. And if we are referring we must be using the term correctly. Meaning is use. Also, reference is use.

DRT fails to establish a way of referring to the world which does not involve descriptions, just as it fails to establish the existence of the necessary a posteriori

and the contingent a priori. Perhaps the simplest way to show this is to consider the possibility of a language which does not use proper names or even natural kind terms but uses descriptions in their place. Such a language is certainly possible. The opposite to this - i.e. a language which employed names but no descriptions - is not possible because there would be nothing to contribute to understanding, which means that we would not have any kind of language at all. The failure to deal with what is understood by a name is the biggest mistake made by DRT. It shows that the direct reference theorists really don't understand what language is.

DRT does however force us to re-examine our views about sense, reference and modality which has, hopefully, led us to the true relationship between language and the world.

BIBLIOGRAPHY

- Almog, J (1986) "Naming Without Necessity"; *Journal of Philosophy* 83, pp210-242
- Ayer, A.J. (1975) "Identity and Reference"; *Philosophia* 5
- Aristotle *Analytica Posteriora, De Interpretatione, Categories, Meta-physics.* Oxford: Clarendon Press
- Bach, K (1987) *Thought and Reference,* Oxford
- Bolton, C.J. (1996) "Proper Names, Taxonomic Names and Necessity", *Philosophical Quarterly,* pp145-157
- Bostock, D (1988) "Necessary Truth and A Priori Truth", *Mind* 97, pp343-379
- Butchvarov, P (1979) "Identity", In French et.al., 1979
- Boolos, G (1990) *Meaning and Method,* (Ed.) Cambridge University Press
- Butler, R.J (1967) *Analytical Philosophy.* (Ed.)
- Carter, W.R (1976) "On A Priori Contingent Truths", *Analysis* 36, pp105-106
- Casullo, A (1977) "Kripke on the A Priori and the Necessary", *Analysis* 37, pp152-159
- Chomsky, N (1975) *Reflections on Language* (New York: Pantheon)
- Cowles, D (1994) "The Contingent A Priori: An Example Free From Existential Worry"; *Philosophical Studies* 74, pp137-141
- Crimmins, M. and Perry, J (1989) "The Prince and the Phone Booth", *Journal of Philosophy* Vol. 86, pp685-711
- Davies, M.and Humberstone, L (1980) "Two Notions of Necessity", *Philosophical Studies* 38, pp1-30
- Devitt, M and Sterelny, K. (1987) *Language and Reality,* Oxford: Basil Blackwell
- Donnellan, K (1966) "Reference and Definite Descriptions", *Philosophical Review* 75, pp281-304

- (1974) (a) "Speaking of Nothing", *Philosophical Review* 83, pp3-31
- (b) Unpublished Manuscript, cited in Salmon, 1982
- (1979) "The Contingent A Priori and Rigid Designation", In French et. al. 1979, pp45-60
- Dummett, M (1973) *Frege: Philosophy of Language*, New York
- (1978) "Frege's Distinction Between Sense and Reference", In *Truth and Other Enigmas*, Harvard, pp420-430
Also in Moore (Ed.) 1993
- Dupre, J (1993) *The Disorder of Things*, Harvard
- Evans, G (1973) "The Causal Theory of Names"; *Proceedings of the Aristotellian Society* Vol. 47 pp187-208
Also in Moore, (Ed.)
- (1979) "Reference and Contingency", *The Monist*, 62 pp161-189
- (1982) *The Varieties of Reference*, Oxford University
- Fitch, G.W. (1976) "Are There Necessary A Posteriori Truths?", *Philosophical Studies* 30 pp243-247
- Fodor, J. (1975) *The Language of Thought*, (New York: Thomas Y. Crowell)
- Forbes, G (1989) *Languages of Possibility*, Blackwell
- Frege, G (1892) "On Sense and Reference", in *Collected Papers* (1984)
- French, P (et al.) (1979) *Contemporary Perspectives in the Philosophy of Language*, Minneapolis
- Goldberg, S, with Pessin, A (1995) *The Twin Earth Chronicles*
- Grayling, A (1982, 1990) *An Introduction to Philosophical Logic*, Duckworth
- Harman, G (1970) "A Non-Essential Property", *Journal of Philosophy* 67, p183
- Hacking, I (1976) In *Identity, Necessity and Physicalism*, pp147-159, Blackwell
Korner, (Ed.)
- Harrison, J (1987) "The Confusions of Kripke", *Erkenntnis* 27 pp283-290

- Hughes and Cresswell (1972) *An Introduction to Modal Logic*, London, Methuen
- Hume, D *A Treatise Of Human Nature*, Ed. Selby-Bigge, Oxford
- Kaplan, D (1969) "Quantifying In", In Linsky (Ed.) ,1971, pp112-144
 (1970) "Dthat", In French, et. al., pp383-400
 (1973) "On The Logic of Demonstratives", In French, et. al. pp490-518
- Korner, S.A (1976) *Philosophy of Logic* (Ed.)
- Kripke, S.A. (1963) "Semantical Considerations on Modal Logic", *Acta Philosophica Fennica*
 (1971) "Identity and Necessity", In Munitz (Ed.) , New York
 Also in Moore (Ed.)
 (1972) "Naming and Necessity", In Davidson and Harman (Eds.)
Semantics of Natural Language
 (1980) *Naming and Necessity*, Oxford
- Katz, J. (1990) "Has The Description Theory of Names Been Refuted?"
 In Boolos (Ed.), . pp31-61
 (1994) "Names Without Bearers", *Philosophical Review* Vol. 103
 pp1-39
- Kennedy, R (1987) "Salmon Versus Kripke on the A Priori", *Analysis* 47,
 pp158-161
- Linsky, L (1971) (Ed.) *Reference and Modality*, Oxford University
 (1977) *Names and Descriptions*, University of Chicago
- Locke, J *Essay Concerning Human Understanding*. Ed. Nidditch Oxford
- McCulloch, G. (1989) *The Game of the Name*, Oxford
- McGinn, C (1976)(a) "On The Necessity of Origin", *Journal of Philosophy* 73

- (b) "A Priori and A Posteriori Knowledge", P.A.S, pp195-208
- (1982) "Rigid Designation and Semantic Value", *Philosophical Quarterly* 32, pp97-115
- MacBeth, D. (1995) "Names, Natural Kind Terms and Rigid Designators", *Philosophical Studies* 79, pp259-281
- Mackie, J.L. (1974) (a) "De What Re is De Re Modality?" *Journal of Philosophy*, 71, pp551-561
- (b) "Locke's Anticipation of Kripke", *Analysis* 34, pp177-180
- Marcus, R B. (1947) "The Identity of Individuals in a Strict Functional Calculus of Second Order", *Journal of Symbolic Logic* 12, pp12-15
- (1960) "Extensionality", *Mind* 69, pp55-62
- (1976) In Kormer (Ed.) (1976)
- Margalit, A (1979) *Meaning and Use* (Ed.)
- Mellor, D.H. (1977) "Natural Kinds", *British Journal for the Philosophy of Science* 28, pp299-312
Also in Goldberg and Pessin, Op. Cit.
- Mill, J.S. (1843) *A System of Logic*
- Moore, A.W. (1993) *Meaning and Reference* (Ed.) Oxford University
- Morris, T.V. (1984) *Understanding Identity Statements*, Aberdeen
- Munitz, M.K. (1971) *Identity and Individuation*, (Ed.) New York University
- Oppy, G. (1994) "Salmon on the Contingent A Priori and the Necessary A Posteriori", *Philosophical Studies*, 73, pp5-33
- Parsons, T (1969) "Essentialism and Quantified Modal Logic", *Philosophical Review* 78, pp35-52

Also in Linsky (Ed.), 1971, Op. Cit.

- Perry, J. with Crimmins, M. (1989) "The Prince and the Phone Booth", *Journal of Philosophy* 86, pp685-711
- Pessin and Goldberg, (1995) *The Twin Earth Chronicles*,
- Plantinga, A. (1974) *The Nature of Necessity*, Oxford
- (1978) "The Boethian Compromise", *American Philosophical Quarterly* 15, pp129-138
- Predelli, S (1998) "I am Not Here Now", *Analysis* 58.2 pp107-115
- Putnam, H (1970) "Is Semantics Possible?" In Putnam, 1975(b); Schwartz, 1977
- (1973) "Meaning and Reference", *Journal of Philosophy* 70
Also in Moore (Ed.) Op.Cit.
- (1975) (a) "The Meaning of 'Meaning'", in Putnam, 1975 (b)
- (b) *Philosophical Papers II: Mind Language and Reality*
Cambridge University
- (1981) *Reason, Truth and History*, Ch 1,2
- (1983) *Philosophical Papers III: Realism and Reason*,
Cambridge
- (1988) *Representation and Reality* M.I.T.
- (1990) *Realism With A Human Face*, Harvard, Ch 4
- Quine, W.V.O. (1953) (a) "Reference and Modality", In Quine, 1953(b)
Also in Linsky, 1971, Op. Cit.
- (b) *From A Logical Point of View*, New York
- (1960) *Word and Object*, M.I.T.
- (1969) "Replies" In Davidson and Hintikka (Eds.) *Words and Objections: Essays on the Work of W.V. Quine*
- Ray, G (1994) "Kripke and the Existential Complaint", *Philosophical Studies* 74, pp121-135

- Russell, B (1903) *The Principles of Mathematics*, New York
- (1905) "On Denoting", *Mind* 14, pp479-493
- (1911) "Knowledge By Acquaintance and Knowledge By Description", P.A.S. pp108-128
- Salmon, N (1982) *Reference and Essence*, Oxford
- (1986) *Frege's Puzzle*, M.I.T.
- (1988) "How To Measure the Standard Metre", P.A.S. pp193-217
- Schiffer, S (1979) "Naming and Knowing", In French, et. al., pp61-74
- Searle, J (1958) "Proper Names", *Mind* 67
- (1983) *Intentionality*, Cambridge University
- Slote, M (1974) *Metaphysics and Essence*, Oxford
- Smullyan, A.F. (1948) "Modality and Descriptions", *Journal of Symbolic Logic* 13, pp31-37. Quoted in Linsky, 1971
- Sterelny, K (1983) "Natural Kind Terms", In Pessin and Goldberg,
- Strawson, P.F. (1974) *Subject and Predicate in Logic and Grammar*, London
- Taschek, W (1992) "Frege's Puzzle, Sense and Information Content", *Mind* 101, pp767-791
- Tienson, J (1984) "Hesperus and Phosphorus", *Australasian Journal of Philosophy* 62, pp16-25
- Van Brakel, J (1990) "Units of Measurement and Natural Kinds: Some Kripkean Considerations", *Erkenntnis* 33, pp297-317
- Wiggins, D (1965) "Identity Statements", In Butler (Ed.)
- (1976) *Identity, Necessity and Physicalism*,
- (1980) *Sameness and Substance*, Harvard

Wittgenstein, L (1922) *Tractatus Logico Philosophicus*
 (1953) *Philosophical Investigations*, Oxford