KIT FINE

modality and tense

PHILOSOPHICAL PAPERS

MODALITY AND TENSE

This page intentionally left blank

Modality and Tense

Philosophical Papers

KIT FINE

OXFORD
UNIVERSITY PRESS

Preface

THIS volume collects together my published papers on tense and modality up to the present time. It contains two reviews, since the issues they discuss are still of interest; and it also contains a much expanded version of my paper, 'The Reality of Tense', now under the title 'Tense and Reality', and two previously unpublished papers. I have not included my technical papers on modal logic, even when they have contained philosophical material or have had an obvious bearing on philosophical questions; and nor have I included any of my philosophical or technical papers on essence, even when they have dealt with the connection between essence and modality. I have added an introduction to the volume, outlining the central content of each paper and bringing out certain issues and themes that may not be evident from the papers themselves.

There are various people who have helped me in one way or another in the preparation of this volume. They include the many philosophers with whom I have discussed the topics of tense and modality over the years—the UCLA 'crowd', earlier pioneers in modal logic, such as Saul Kripke and Hans Kamp, and several of my students. They also include Peter Momtchiloff, who first urged me to publish some of my work with Clarendon Press, an anonymous referee for the Press, who suggested many valuable improvements, and Ruth Chang, whose steady encouragement made the possibility of this volume an actuality. My greatest debt of gratitude is to Arthur Prior. He was the person who introduced me to modal logic and first got me to think seriously about it, in both its technical and philosophical aspects. He was a tutor during my last two years as an undergraduate at Balliol College, Oxford, from 1965 to 1967, and he helped supervise my research in the following two years, from 1967 until his death in 1969. I cannot think of a more wonderful mentor to have had. He was always supportive, always generous with his time, and always ready to talk logic. He seemed to have the intellectual virtues to which so many of us aspire but so rarely attain. What he had, above all, was an unlimited interest in the subject itself, one that knew no bounds of either an intellectual or personal sort. I feel very fortunate to have begun my academic life under his guidance; and it is with the greatest affection and admiration that I dedicate this volume to his memory.

K.F.

Contents

Origin of the Papers

Papers are listed in chronological order.

Review of David Lewis's *Counterfactuals*, *Mind* 84 (1975), 451–8.

Review of Alvin Plantinga's *The Nature of Necessity*, *The Philosophical Review* 85 (1976), 562–6.

'Prior on the Construction of Possible Worlds and Instants', Postscript to *Worlds, Times and Selves* (with A. N. Prior) (London: Duckworth, 1977), 116–68.

'Reference, Essence and Identity'. Previously unpublished and written up in the spring of 1984 as a talk for the conference 'Themes from Kaplan'. Chris Peacocke was the commentator.

'Plantinga on the Reduction of Possibilist Discourse' in *Alvin Plantinga*, ed. J. E. Tomberlin and P. van Inwagen, Reidel Profiles Series 5 (Dordrecht: Reidel, 1985), 145–86.

'The Problem of *De Re* Modality', in *Themes From Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein (Oxford: Oxford University Press, 1989), 197–272.

'Quine on Quantifying In', in *Proceedings of the Conference on Propositional Attitudes*, ed. C. A. Anderson, J. Owens (Stanford: CSLI, 1990), 1–26.

'The Problem of Possibilia', in *Handbook of Metaphysics*, ed. D. Zimmerman (Oxford: Oxford University Press, 2002), 161–79.

'The Varieties of Necessity', in *Conceivability and Possibility*, ed. T. S. Gendler and J. Hawthorne (Oxford: Oxford University Press, 2002), 253–82.

'Tense and Reality', an earlier, much shorter version of the paper was published under the title 'The Reality of Tense' in *Synthese* (2005).

'Necessity and Non-existence' (2005). Previously unpublished.

This page intentionally left blank

Introduction

It is an oddity of current thinking about modality that it has been heavily influenced, one might even say dominated, by two extreme and highly implausible views. The first of these, associated with the name of Quine, is that modal notions are lacking in sense. There is no intelligible distinction to be drawn between what is necessarily and what is contingently the case or between an object's essential and accidental features. The second of these two views, associated with the name of David Lewis, is that the possible and the actual are on an ontological par. Other possible worlds and their inhabitants are just as real as the actual world and its inhabitants; and there is no difference between them in regard either to the degree or to the kind of reality that they possess.

Given the great implausibility of these views, it is worth considering what kind of intellectual environment might have led philosophers to take them so seriously. Part of what is involved, of course, is the adoption of a theory-driven methodology, one that favours considerations of a broadly theoretical sort over strong and seemingly compelling intuitions. But also partly involved is the adoption of a broadly empiricist point of view. Empiricists have always been suspicious of modal notions. For them, the world is an on-or-off matter—either something happens or it does not; and there appears to be no room in their on-or-off world for a distinction between what happens of necessity and what only happens contingently or between the essential features of an object and those that are only accidental.

There can be no doubt that Quine's misgivings about modality were to some extent fuelled by empiricist concerns. But the same can be seen to be true of David Lewis, notwithstanding the lavish extravagance of his ontology. For empiricists, in so far as they have been able to make sense of modality, have tended to see it as a form of regularity; for something to hold of necessity is for it always to hold, and for something to hold possibly is for it sometimes to hold. But if there is not enough going on in the actual world to sustain the possibilities that we take there to be, then one strategy for the empiricist is to extend the arena upon which the possibilities are realized to include what goes on in each possible world. Of course, such a view is compatible with a moderate realism in which possible worlds, and

what goes on in them, are taken to have a different ontological status from the actual world and what goes on in it.[1] But combine the regularity view of modality with a nominalism about what there is and we end up with a position very like Lewis's. Indeed, it might be argued that, *au fond*, Lewis is as sceptical of modal notions as Quine. Neither can understand modality except as a form of regularity; and the only difference between them lies in the range of the regularities to which their respective ontologies allow them to appeal.

To a large extent, my own thinking about modality has been sustained by a deep animosity to these two views. It has seemed to me almost axiomatic that there is an intelligible distinction between what is necessary and what is contingent and that there is an ontological difference between actual objects and merely possible objects—between actual people and actual cities on the one hand, and merely possible people and merely possible cities on the other. We might call someone who takes modality seriously a *modalist* and some-one who takes actuality seriously an *actualist*. My position is therefore a form of modal actualism.

Much of my work can be seen as an attempt to defend and elaborate a viable form of modal actualism; and so it may be helpful for me to discuss my work under each of these heads.

A prime task for any modalist is to defend the intelligibility of modal notions against the arguments of Quine. Several philosophers (they include Barcan Marcus [1990], Kaplan [1986], Kripke [1980], and Plantinga [1974]) have taken on this task; and my own attempt is made in two companion papers, 'The Problem of *De Re* Modality' and 'Quine on Quantifying In' (Chs. 2 and 3 of the present volume). The second of these was intended to provide an abridged version of the first but ended up containing a great deal of additional material. The first derived in its turn from an earlier unpublished paper, 'Reference, Essence, and Identity', which opens the present collection.

Quine, of course, had arguments against modality both in its application to sentences (modality *de dicto*) and in its application to objects (modality *de re*). The focus of Chs. 2 and 3 is on the second class of arguments. We follow Quine in assuming, if only for the sake of argument, that modality has intelligible application to sentences. Our question is whether it also has intelligible application to objects. Is its intelligibility *de dicto* compatible with its intelligibility *de re*?

It is now generally acknowledged that Quine had two arguments against the intelligibility of *de re* modality, one broadly metaphysical and the other broadly logical in character. The focus of the first is on the intelligibility of a certain notion, of an object's necessarily being a certain way, while the focus

---

[1] As in Stalnaker [2003], ch. 1.

of the second is on the intelligibility of a certain quantificational locution, one in which we say of some object x that necessarily it is a certain way (something we might symbolize as ∃x□Fx). These two conclusions are independent of one another. For one might combine the intelligibility of the notion with the unintelligibility of the locution by maintaining that the notion, though intelligible, is not to be expressed by means of the locution. And (somewhat less plausibly) one might combine the intelligibility of the locution with the unintelligibility of the notion by maintaining that the locution, though intelligible, is incapable of expressing the notion.

The 'metaphysical' line of argument goes as follows. One cannot make sense of an object's necessarily being a certain way independently of how it is described. When described as the number 7 + 2, the number 9 will necessarily be greater than 7; and when described as the number of planets, the number will not necessarily be greater than 7. But no sense can be attached to the claim that the number *itself* is necessarily greater than 9, independently of how it is described.

In evaluating the argument, it is important to be clear on which notion of necessity or possibility is in question. Quine often has the logical or analytic modalities in mind (necessity in virtue of logical form or of meaning). But in these cases, it seems to me, we can make sense of what it is for a certain object or for certain objects necessarily to be a certain way, independently of how they are described. Indeed, it seems to me that our understanding of their *de re* application is already implicit in our understanding of their *de dicto* application; we cannot fail to understand the one if we already understand the other.

My argument for this claim is based upon a certain view of logical form. We normally think of logical form as exclusively an attribute of sentences. However, the notion may also be taken to have application to objects. The argument from P to P is valid since the same sentence occurs as premiss and as conclusion. But if logical form can take account of the repeated occurrence of a sentence, it should also be able to take account of the repeated occurrence of an object. Thus it should be part of the logical form of two objects, if they are the same, that they are the same. But this then means that we can sensibly say that the condition 'x = y' is satisfied by two identical objects as a matter of logical or analytic necessity. For it will be part of the logical form of the trio consisting of the condition and of the objects, that the objects are the same; and so it will follow from the logical form of the objects and of the condition that the condition is indeed satisfied by the objects.

What may make Quine's conclusion seem irresistible is that the resulting notion of *de re* necessity appears to privilege certain descriptions over others. For given two objects that are in fact the same, it privileges their being the same and takes it to be a matter of necessity, while failing to privilege other features or other relationships between two objects. But even if the notion

itself is invidious in this way, our understanding of the notion is not. It is not that we understand the *de re* notion through arbitrarily declaring certain descriptions to be privileged. Rather, we understand the *de re* notion, like the *de dicto* notion, in terms of logical form; and it is then a consequence of this understanding that identities will hold as a matter of necessity.

I therefore conclude that the *de re* application of the logical and analytic modalities is in good conceptual order. But the situation is rather different in the case of the metaphysical modalities. The metaphysical notion of necessity, in contrast to the logical or analytic notions, is capable of discriminating in an interesting way between different objects. The number 9, for example, is necessarily a number though not necessarily the number of the planets; and Socrates is necessarily a person though not necessarily a philosopher. From whence derives these differences in the necessary features of an object?

The Quinean will respond that it derives from our privileging certain descriptions over others. It is built into our very understanding of *de re* metaphysical necessity that certain descriptions, as opposed to others, will be regarded as revealing the essence of an object. Now, in the case of the logical and analytic modalities, I was able to counter the Quinean response; I was able to argue that our understanding of the *de re* application of the modalities was already implicit in our understanding of their *de dicto* application, without the need to privilege certain descriptions. But I know of no convincing argument of this sort in the case of the metaphysical modalities. This is not to say that we should give up the idea of metaphysical necessity *de re*. It is just that I know of no convincing way to argue from its intelligibility *de dicto* to its intelligibility *de re*. In at least this respect, then, Quine's arguments against *de re* modality still have some force. Even if we do not share his strict standards for what it would be to make sense of the *de re* notion, we can still recognize that some kind of conceptual leap is required to bridge the gap between our *de re* and our *de dicto* understanding of the notion.

Let us now turn to Quine's 'logical' argument. This is an argument against the intelligibility of quantifying into modal contexts. It is observed that the occurrence of singular terms within modal contexts may not be 'open to substitution'. A sentence such as 'necessarily, $9 > 7$' may be true while the sentence 'necessarily, the number of planets is greater than 7' is false, even though it is obtained from the first sentence by substituting the coreferential term 'the number of planets' for the original term '9'. From this it is meant to follow that quantification into the resulting context is unintelligible; we cannot make sense of the sentence 'for some x, necessarily $x > 7$'. For in order to make sense of the sentence, we would have to understand what it was for the condition 'necessarily $x > 7$' to be satisfied by an object, independently of how it is described; and the failure of substitution shows that no such understanding can be attained.

The strategy of Chs. 2 and 3 is to break down the inference from the failure of substitution to the unintelligibility of quantifying into modal contexts into four main steps. The first is from the failure of substitution to the 'irreferentiality' of the term; given this failure, the term '9' in 'necessarily, $9 > 7$' cannot be understood as having the sole linguistic function of picking out an object for the rest of the sentence to say something about. The second step is from the irreferentiality of the term to the irreferentiality of the variable; given that the term '9' in 'necessarily, $9 > 7$' does have the sole linguistic function of picking out an object for the rest of the sentence to say something about, the variable 'x' in 'necessarily, $x > 7$' will likewise not have the sole linguistic function of picking out a value for the rest of the condition to say something about. The third step is from the irreferentiality of the variable to the breakdown in objectual satisfaction; given that the variable does not have the sole linguistic function of picking out a value for the rest of the condition to say something about, we cannot make sense of what it would be for the condition 'necessarily, $x > 7$' to be satisfied by the assignment of an object to the variable. The fourth and final step is from the breakdown in objectual satisfaction to the unintelligibility of quantification into modal contexts.

Although these various steps might appear to be reasonable, even truistic, I believe that each of them can fail; and a large part of the two chapters is concerned to detail the various ways in which this can happen. In regard to the third step, for example, it is argued that satisfaction can be objectual; we can make sense of what it would be for the condition to be satisfied by the assignment of objects to the variables, even though the variables are not referential, i.e. not solely used to pick out their values. This is because the variables themselves, in addition to their values, might be relevant to the semantic evaluation of the condition. We thereby obtain a new form of 'literalist' quantification, which is objectual but not referential.[2] Or again, in regard to the first step, it is argued that referentiality is compatible with a failure of substitution. Quine assumed that 'Giorgione' in 'Giorgione is so-called because of his size' was not referential, since the substitution of the coreferential term 'Barbarelli' turned the sentence from a truth into a falsehood. But it may plausibly be argued in this case that the term 'Giorgione' *is* referential; for the failure of substitution may be attributed to a difference in the reference of 'so' rather than a failure in the referentiality of the subject-term.[3] The discussion is, if you like, a microscopic examination of some of the basic concepts of philosophical logic—termhood, referentiality,

---

[2] The relational treatment of variables in Fine [2003*b*] provides further reasons for questioning this step in the argument.

[3] See Forbes [1997] for an application of this kind of semantic mechanism to the use of names in belief reports.

objectuality, satisfaction, and quantification—and reveals anomalies in their behaviour that are not apparent under a more macroscopic view.

But the critical question is whether the steps go through in the modal case; and it seems to me that they do not. The failure occurs at the first step. Quine wishes to argue from the failure of substitution to the irreferentiality of the term '9' in 'necessarily, 9 > 7'. Now let us concede that if the term '9' was here being used in the same kind of way as the term 'the number of planets' in 'necessarily, the number of planets > 7' (either both referentially or both irreferentially), then it would indeed follow from the failure of substitution that both terms were being used irreferentially, since their both being used referentially would not be compatible with the failure of substitution. But what if the term '9' were not being used in the same kind of way as the term 'the number of planets'? Nothing would then follow concerning the irreferentiality of either term. Indeed, it seems to me that the most plausible view is that the use of the name '9' is referential, while the use of the description 'the number of planets' is not.

This line of thought also puts the second step in doubt. It is a reasonable requirement on our understanding of the quantified sentence 'for some x, necessarily x > 7' that it should derive from our understanding of an instance 'necessarily, t > 7', for some particular closed term t. Our understanding of the instance should conspire with our general understanding of the quantifier so as to yield an understanding of the quantified sentence. But what is a relevant instance? If the use of different terms t in the context 'necessarily, t > 7' is not uniform, then we have some choice as to what the appropriate use of the term should be taken to be. Thus we may concede that the use of the term 'the number of planets' is not referential in 'necessarily, the number of planets > 7' and yet take our understanding of the quantified sentence to derive from the referential use of '9' in 'necessarily, 9 > 7'. A referential understanding of the quantifier is thereby secured, even though referentiality may not be preserved upon substituting a description for the variable of quantification.

We might see both of Quine's arguments as resting upon certain background assumptions in the respective areas of metaphysics and the philosophical logic. The metaphysical argument presupposes an empiricist criterion of intelligibility under which essence must have its source in meaning. And the logical argument presupposes a naive view of singular terms under which no significant distinction is to be drawn between the use of names and descriptions. Without these background assumptions, the arguments lose all their force.

It is not enough for the modalist to defend the intelligibility of modality against attacks. He should also provide a positive account of its sense. Even if the notion cannot be defined in other terms, he should still attempt to make clear how it is to be understood. The logical and analytic modalities

are relatively unproblematic in this regard. But the metaphysical modalities are not. What is it for a truth to hold or for a feature to be had of *metaphysical* necessity? Philosophers have not given this question the attention it deserves; they have simply taken for granted that there is a single coherent notion that goes by this name.

My own view is that metaphysical necessity is to be understood in terms of its distinctive source. A logical necessity has its source in logical form; it is true, or necessary, in virtue of its logical form. An analytic necessity has its source in meaning; it is true, or necessary, in virtue of the meaning of its terms. I wish to claim, in an analogous manner, that a metaphysical necessity has its source in the identity of objects; it is true, or necessary, in virtue of the objects with which it implicitly deals. This account is meant to cover not only the obvious cases of *de re* necessity but also the less obvious cases of *de dicto* modality. Thus the necessary truth that all bachelors are unmarried has its source in the concept of being a bachelor just as the necessary truth that Socrates is a man has its source in the identity of Socrates.

If this account is to be of any explanatory value, we must distinguish between the identity or 'essence' of an object and the properties that it has as a matter of necessity. I have argued for the distinction between the two in some other papers, not included in the present volume.[4] But there is one chapter in the volume, 'Reference, Essence and Identity' (Ch. 1), that does allude to the distinction, though in a very rough and rudimentary form.

Once we have pinned down the notion of metaphysical necessity, there still arises the question of its relationship to the other modalities. Some forms of necessity are clearly species of others. A mathematical necessity, for example, is presumably a metaphysical necessity which happens to be a proposition of mathematics. Other forms of necessity are clearly relative versions of others. A technological necessity, for example, is a natural necessity relative to the current state of technology. But suppose we look at modalities that cannot be explained, in this or in any other way, in terms of other modalities. What basic forms of necessity remain?

This is the topic of Ch. 7, 'The Varieties of Necessity'. It is claimed that there are three basic forms of necessity—the metaphysical, the natural, and the normative. Each has its own distinctive source: metaphysical necessity in the identity of objects; natural necessity in the 'fabric' of the universe; and normative necessity in the realm of values and norms. Thus assuming Aristotle is right, it is a metaphysical necessity that Socrates is a man; assuming Newton is right, it is a natural necessity that distant bodies attract one another; and assuming Kant is right, it is a normative necessity that lying is wrong.

On the face of it, no one of these notions of necessity is subsumed under any other. Thus the Aristotelian metaphysical necessity is neither a natural

---

[4] Most notably Fine [1994, 1995].

nor a normative necessity, the Newtonian natural necessity is neither a
metaphysical nor a normative necessity, and the Kantian normative necessity
is neither a metaphysical nor a natural necessity. However, various contem-
porary philosophers, heavily enamoured of the notion of metaphysical
necessity, have argued against this intuitive position. They have argued
that natural necessity or that normative necessity is to be subsumed under
the more general rubric of metaphysical necessity. The issue is one with far-
reaching implications for our conception of science and of ethics; for the
status of their principles, the way we take them to bear on the world, and the
way we take ourselves to know them is intimately tied to the kind of
necessity that we take them to possess.

Chapter 7 is in large part an attempt to show that these philosophers are
mistaken. The first class of subsumptionists have been impressed by the
failure of certain counterexamples to the claim that every natural necessity
is a metaphysical necessity. It is sometimes been thought that the 'inverse
cube law', for example, is a metaphysical possibility; bodies might be
attracted to one another inversely to the cube, not the square, of the distance
between them. The negation of this law would then be a natural necessity
though not a metaphysical necessity. But it has been objected that all we can
properly conclude from this putative counterexample is that the world might
contain a different kind of object, schbodies rather than bodies, which behave
according to the inverse cube law rather than the inverse square law. Thus
there is no counterexample to the inverse square law since this concerns the
behaviour of bodies, not schbodies; and similarly, it is maintained, for any
other putative counterexample to the claim of subsumption.

There is something deeply suspicious about this line of defence. Instead of
considering the merits of particular counterexamples to the claim of subsump-
tion, let us simply ask: is every metaphysical possible world a natural possibil-
ity? The answer seems clearly to be 'no'. Surely, among the wide range of
metaphysically possible worlds, some are simply excluded as genuine possibil-
ities according to natural law. There will be no genuine possibility of a world in
which there are schbodies, for example, behaving according to the inverse cube
law rather than bodies behaving according to the inverse square law. Thus the
most that this line of objection can show is that the putative counterexamples
to the subsumption claim have been misdescribed. They do not concern the
necessity of these laws, as these are usually stated, but the necessary non-
existence of certain kinds of thing, such as schbodies, or the necessity of laws
under a broader construal of their range of application.

The second class of subsumptionists have likewise been unimpressed by
the putative counterexamples to the claim that every normative necessity is a
metaphysical necessity. It has been thought to be a normative necessity that
lying is wrong, for example, but not a metaphysical necessity. But these
philosophers have wanted to distinguish in a familiar way between the

property of being wrong (what wrongness is) and the *concept* of being wrong (what we understand in understanding 'wrong'). Now normative necessity concerns the application of the property, rather than the concept; and it may well be a metaphysical necessity that lying has the property of being wrong even though it is not a metaphysical necessity that lying will fall under the concept.

I do not object in principle to drawing a distinction between the concept and the property of being wrong. But it seems to me that the usual naturalistic ways of drawing the distinction will have implausible epistemological consequences. Suppose, for example, that the concept of wrong is the concept of a property that is generally disvalued. Then in order to know that something is wrong I will need to know that it is generally disvalued. But this is not something I need to know in order to know that it is wrong; and I argue that there is no reasonable way in which the naturalist might overcome difficulties of this sort.

Even if the subsumption claim fails, it may still be possible to see natural or normative necessity as relative forms of metaphysical necessity: relative necessity will be metaphysical necessity relative to the natural laws; and normative necessity will be metaphysical necessity relative to the normative laws. The problem with this proposal is that it fails to provide an adequate account of the necessity of the laws. There appears to be a significant sense in which the laws themselves are necessary. But according to the proposal, the necessity of the laws can amount to no more than their being entailed by the laws, i.e. to their self-entailment. However, every proposition entails itself; and so, in regard to their status as necessary truths, the laws are incapable of being distinguished from any other truth.

The earlier literature on modality, arising from the work of Quine, was characterized by an unwarranted contempt for modal notions. The subsequent literature, arising from the work of Kripke, has been characterized by an unwarranted enthusiasm. This enthusiasm has taken two different, though related, forms. The first, which we may call 'modal mania', is a matter of seeing everything as modal; every notion which is somehow associated with modal features is itself taken to be modal. The second, which we may call 'modal myopia', is a matter of seeing all modality as metaphysical; every modality is somehow to be understood as a form of metaphysical modality.

The failure to distinguish between the identity or essence of an object and its necessary features is an instance of modal mania. Another, that I have discussed elsewhere,[5] is provided by the standard modal characterization of supervenience: one class of propositions is taken to 'supervene' on another if

---

[5] In §5 of Fine [2000*b*]. A similar point is made in connection with the concept of referentiality in §3, step (2), of Fine [1989] (Ch. 2 below).

it is necessary that for every true proposition from the first class there are true propositions from the second class that entail it. But this is to ignore the explanatory aspect of supervenience. Not only must the propositions from the second class *entail* the proposition from the first class, they must be propositions *in virtue of which* it is true. There is an explanatory connection here that cannot be captured in purely modal terms.

The tendency to conflate or to collapse the different forms of necessity is an instance of modal myopia. Another is provided by the tendency to use the notion of metaphysical necessity as the modality of choice in discussions of semantics. The content of a sentence, for example, is often identified with the set of possible worlds in which it is true, where the possible worlds in question are taken to be those that are *metaphysically* possible. But this has the awkward consequence that all metaphysically necessary truths will have the same content. A much more satisfactory account of content may be obtained by appealing to *analytically* possible worlds (as was customary in an earlier semantic tradition). The content of sentences which are metaphysically necessary but not analytically equivalent may then be distinguished. In general, there is no reason to expect that the concept of necessity that is most pertinent to the study of metaphysics should be capable of doing double duty as the concept that is also most pertinent to the study of semantics.

Philosophers with a new concept are like children with a new toy; their world shrinks to one in which it takes centre stage. Now there can be no doubt that the preoccupation with metaphysical modality has had a beneficial effect on the discussion of a number of philosophical topics. But there is a danger of its becoming a new restrictive orthodoxy, with metaphysical modality supplanting logical modality as the arbiter of intelligibility for all things modal.

The question of how we should understand metaphysical modality is further pursued in the most recent of the papers from this volume, 'Necessity and Non-Existence' (Ch. 9). It is there argued that there are two fundamentally different ways in which a proposition may be metaphysically necessary: it may be a *worldly* necessity, true *whatever* the circumstances; or it may be a *transcendent* necessity, true *regardless* of the circumstances. The circumstances are constituted by how things might turn out; and, in the first case, the circumstances are relevant to the truth-value of the proposition but in such a way as to render it true whatever they might be while, in the second case, the circumstances are not even relevant to the truth-value of the proposition. There is therefore no possibility of them either rendering the proposition true or rendering it false. The proposition that Socrates exists or does not exist is necessary in the worldly sense since, whatever the circumstances, they will either include his existence or his non-existence and hence will be such as to render the proposition true. On the other hand, the proposition that Socrates is identical to Socrates will be transcendent since

there are no circumstances concerning Socrates or the rest of the world that have any bearing on his self-identity.

We might understand the distinction by analogy with the familiar distinction between sempiternal and eternal truths. A sempiternal truth is one that is true whatever the time while an eternal truth is one that is true regardless of the time. Thus the proposition that Socrates was or is now alive is a sempiternal truth while the proposition that the Battle of Hastings took place in 1066 is an eternal (though not a transcendent) truth. Substitute worlds for times, confine the content of worlds to how things turn out, and we obtain the corresponding distinction in the modal sphere.

Just as there is a distinction between necessary and transcendent truth, so there is a distinction between necessary and transcendent existence. A necessary existent is one that exists *whatever* the circumstances while a transcendent existent is one that exists *regardless* of the circumstances. Perhaps the circumstances in the sense of the circumstances-whatever-they-might-be is a necessary existent while ordinary abstract objects, such as sets and numbers and the like, will be transcendent existents. Thus necessary existents are in the world while transcendent existents enjoy a form of existence outside the world, just as sempiternal existents are in time while eternal existents enjoy a form of existence outside time.

An interesting application of the distinction is to the status of hybrid abstract objects, such as singleton Socrates, which contain concrete constituents even though they themselves are abstract. If we are actualists, then we will deny that there is such a thing as singleton Socrates in a possible world in which Socrates does not exist. However, we will still want singleton Socrates to be like any other set in enjoying an extra-worldly form of existence. Thus there will be a sense in which we can maintain that it is possible both that the set exists and that there is no such set!

Another application is to the status of sortals. I believe that 'substance' sortals—such as *man* or *explosion* or *number*—are extra-worldly in their application. They apply to their objects regardless of the circumstances. Although there may be something about how the matter of Socrates turns out that is relevant to its *constituting* a man, there is nothing about how Socrates himself turns out that is relevant to his *being* a man. If I am right, then this means that philosophers have been mistaken in thinking that Socrates cannot be a man unless he exists, that existence must precede essence. Socrates must already be a man, if I may put it that way, before the question of how things turn out for him can even arise.

Some further applications of the distinction are explored in the chapter; and I feel that I have only just begun to understand the ramifications that the distinction might have on a wide range of topics.

I turn now to the second plank of my position, the actualism. The actualist favours the actual over the possible. But there are two rather different forms

that the favouritism might take; one ontological and the other metaphysical. The ontological form of actualism has already been mentioned; it is the position that takes only actual objects to be real. And similarly for the ontological form of presentism; it is the position that takes only present objects to be real.

The ontological actualist and the ontological presentist face a challenge. For, on the face of it, talk of possible objects and talk of past and future objects makes perfectly good sense. We may correctly say, for example, that there are many possible people who never have and never will be born; and I hope we may correctly say that the total number of people in the world— past, present, and future—exceeds one billion billion. But how are we to make sense of such claims if possible objects or past and future objects are not real?

This is a topic that I discuss in three of the chapters that make up Part II of this volume: 'Prior on the Construction of Possible Worlds and Instants' (Ch. 4), 'Plantinga on the Reduction of Possibilist Discourse' (Ch. 5); and 'The Problem of Possibilia' (Ch. 6). Each approaches the topic from a somewhat different point of view. The first is a quasi-formal and somewhat condensed account of how various kinds of possibilist discourse might be translated into the kind of language acceptable to the actualist; it may be skipped or skimmed by the reader who is not so interested in technical detail. The second contrasts my approach to possible objects with that of Plantinga, another modal actualist, and therefore serves to highlight what I regard as most distinctive about my own approach. The third was intended as a more informal treatment of the topic and may usefully be read as an introduction to the other two chapters; it also contains some additional material and an extended discussion of modal fictionalism.

The key idea behind my approach is to treat ordinary quantification over possible objects as a special way of quantifying over actual objects. Roughly speaking, to say that some possible object is a certain way is to say that possibly some object is that way. Thus to say that there is a possible child of whom it is possible that J. Edgar Hoover was the father is to say that it is possible that there is a child of whom J. Edgar Hoover is the father. Here the embedded quantifier 'there is a child' should be taken to be actualist rather than possibilist; it should be taken to range, in each possible world, over the actual objects of that world rather than over every possible object.

However, this account only works when the condition attributed to the possible object is itself modal. For example, to say that some possible object is not identical to any object (actualist 'any') is not to say that possibly some object is not identical to any object. Embedding the actualist quantifier within the scope of the possibility operator has two effects, one desired and the other undesired. The desired effect is to extend the effective range of the quantifier to all possible objects; the modal-quantifier combination

'◇∃x' looks, in effect, for some actual object of some possible world, i.e. for some possible object. The undesired effect is to extend the circumstances in which the condition might be satisfied from the actual world to any possible world; in saying ◇∃xA(x) one is saying that some possible object possibly satisfies A(x), not that it actually satisfies A(x). We therefore require some way of undoing the undesired effect while leaving the desired effect alone; and there are various more or less natural ways, all acceptable to the actualist, in which this may be achieved. Instead of saying that it is possible that an object is a certain way, for example, we might say that the circumstances are such that it is possible that an object is a certain way in those circumstances. Reference to the circumstances then brings the 'target' of the modal-quantifier combination back home to the actual world.

It is important to appreciate that the proposed analysis is not a form of proxy reduction (what Lewis [1986], ch. 3) calls 'ersatzism'). There are no objects that do duty for the possible objects. If it is asked, 'what, under the proposed analysis, do I take a possible object to be?', then no sensible answer can be given. In talking of possible objects—of possible people, say, or possible facts—one is talking of actual objects—of actual people or actual facts—but under the rubric of what is possible.

Many actualists have wanted to identify a possible object with some sort of actualist substitute. It has been supposed, for example, that each possible object has an individual essence, an essence possessed by that object alone. The essence of a possible object will still be actual even when the object itself is not; and so the essences of possible objects may be used as actualistically acceptable substitutes for the objects themselves.

I doubt that any such form of proxy reduction can succeed. The basic problem is that the possibilist's ontology may outrun the resources by which the actualist is capable of discriminating between its objects. There may be two possible electrons, for example, that are completely indiscernible from the actualist point of view, any actualistically acceptable property of the one is a property of the other; and so there will be no individual essences, or the like, by which the objects might be distinguished. Of course, the electrons will not be indiscernible *tout court*; for one of the electrons, e, will have the property of being identical to e while the other will not. But it is unclear how the actualist might be justified in admitting these identity properties into his ontology if he is not also justified in admitting the objects by which they are given.

Part of what may have made the strategy of proxy reduction seem so attractive is the view that there is a categorial difference between actual and possible objects. Actual objects are somehow concrete or substantial while possible objects are somehow abstract or insubstantial. Moreover, this difference appears to be one in kind; it is of the nature of actual objects to be relatively concrete or substantial and of the nature of possible objects to

be relatively abstract or insubstantial. It is therefore natural to seek an account of possible objects that will reflect this difference. And how better to do this than to identify possible objects with the abstract counterpart of the actual objects—with the possible essences of an object, for example, or with the possible ways an object might be? Indeed, some philosophers may even have been attracted by the view that this is what a possible object *is*, and not just something with which it may reasonably be identified.

Tempting as such a view may be, it is hard to see how it can be sustained. For if it is of the nature of a possible object to be abstract, then this is presumably a property that it must have in any possible circumstance in which it is actual. But in such a circumstance, it is an actual object and therefore also concrete. Thus the properties of being abstract and concrete, whatever they might be, are not incompatible with one another and so cannot, after all, give rise to a difference in kind. Indeed, if I am correct in thinking that attributions of kind are unworldly, that they hold independently of the circumstances, then it cannot even be allowed that a merely possible object might be actual, and hence concrete in some possible world, and yet denied that it is concrete in the actual world.

To this objection, the actualist might respond that it is only in a manner of speaking that a possible object should be said to be possibly actual. Considered as an object in its own right, it is already actual. But considered as a surrogate for a possible object, it is only to be said to be actual when it *corresponds*, in the appropriate way, to an actual object. Thus we may safely say that it is possibly actual, without thereby implying that it suffers any difference in kind.

The problem with this response is that it is so utterly implausible. I may say that:

(1) possibly some object is the actual son of J. Edgar Hoover.

From this it follows that:

(2) some possible object is possibly the actual son of J. Edgar Hoover.

In (1), the use of 'actual' is the literal use while, in (2), it is merely a manner of speaking. But surely what is said to be possible for the possible object under (2) is the same as what is said to be possible for the actual object under (1).

If I am right, then the difference between possible and actual objects is not correctly regarded as a difference in kind. It is a difference in what one might call ontological status, of what it is for the object to be. This is not to dispute that possible objects are somehow lacking in substantiality. But the lack of substantiality resides in what it is for there to be such objects rather than in the objects themselves.

I turn to the other way in which the actualist might favour the actual over the possible. This concerns not a question of ontology, of what there is, but a

question of metaphysics, of what is the case. The issue is whether reality, or the 'facts', is biased towards one particular world, the actual world, or is neutral between one possible world as opposed to another. Given that the facts are 'worldly', ones that can sensibly be said to obtain at one world and fail at another, then should we take reality to be constituted by the facts that obtain in the actual world or to be somehow constituted by the facts that obtain in any possible world? Similarly in the tense-logical case. The issue is whether tensed reality is biased towards one particular time, the present, or is neutral between different times. Given that the facts are tensed, ones that can sensibly be said to obtain at one time and fail at another, then should we take reality to be constituted by the facts that obtain at the present time or to be somehow constituted by the facts that obtain at any time?

Now it may well be thought that there is no real issue here. Certainly, if one thinks of the facts as unworldly or as tenseless, then they will not be biased towards one world or one time as opposed to another. But once one grants that the facts may be worldly or tensed, then how could they *fail* to be biased towards a particular world or a particular time? Must not the facts that constitute reality be ones that simply obtain? And how could we think of them as simply obtaining unless there is a privileged standpoint—the actual world or the present time—from which they can be regarded as obtaining?

There are, I believe, two general assumptions about the nature of reality that have stood in the way of seeing how one might combine a worldly or tensed conception of the facts with an unworldly or tenseless conception of the standpoint from which they obtain. The first is that reality is absolute; the facts that constitute reality are those that simply obtain, they do not obtain relative to this or that standpoint. The second is that reality is coherent; it is 'of a piece' and will not contain facts that are incapable of jointly obtaining at a single standpoint. By challenging either of these assumptions, one can thereby embrace a position that accepts the worldly or tensed facts but does not accept a privileged standpoint from which they obtain.

The resulting form of realism about the worldly or the tensed facts I dub 'non-standard'; and it can take either of two forms depending upon which assumption is challenged. Thus it can either take reality to be relative to an external standpoint or it can take reality to be absolute, but fragmented—not 'of a piece'. Now it has to be conceded that either form of non-standard realism is most implausible in the modal case. For, *pace* Lewis, it is hard to believe that the facts that obtain in other possible worlds are just as real as the facts that obtain in the actual world. But non-standard realism has much more plausibility in the tense-logical case and there is nothing at all outrageous in the view that the tensed facts that obtain at the present time are no more real than the tensed facts that obtain at any other time.

'Tense and Reality' (Ch. 8) is largely an exploration of the implications of adopting a non-standard form of realism about tense. I spend a good deal of

time trying to state what the position is and, on this point, there is a significant overlap with an earlier paper of mine, 'The Question of Realism' (Fine [2000*b*]), and with contemporary discussion of McTaggart's argument against the reality of time. In the earlier paper, it is argued that the concept of reality underlying the issue of realism should be taken as a primitive and that clarity in understanding the issue is to be achieved not through further analysis of the concept but in attempting to see how it is to be applied. This conceptual stance is then taken as the starting point for a discussion of what might be meant by realism in the case of tense. By taking the concept of reality more seriously as an integral element of McTaggart's argument, it is possible, I believe, to achieve a deeper understanding of what the argument is and how it might be met.

Although I do not take a stand on whether to be realist about tense, I do argue that the non-standard forms of realism are more plausible, in several key respects, than the standard forms. They are better able to make sense of the passage of time, better able to account for the link between the truth of a tensed utterance and the reality upon which it is meant to bear, and better able to make allowance for the special theory of relativity. In working through these various considerations, we are led to a view that is radically different from the standard view. There is no privileged NOW, but a succession of NOWs, each equally real; the truth of an utterance may shift with its time, not because of a shift in the context, but because of a shift in the reality with which it deals; and, given the truth of special relativity, the physical world must be taken to be endowed with a plurality of space-time structures. Despite its radical character, I believe that the view constitutes the only viable way in which a form of realism about tense might be sustained.

Non-standard realism constitutes some sort of compromise between the standard realist and anti-realist positions. There are two competing perspectives that have been taken to reflect the genuine nature of temporal reality. Under one, we take ourselves to be in time while, under the other, we take ourselves to be out of time; and the facts that constitute reality, in either case, are taken to be the facts as they appear to someone from that perspective— either as tensed when the perspective is from within time or as tenseless when the perspective is out of time. But from the point of view of non-standard realism, each perspective is legitimate in its own way. The internal perspective is appropriate to the nature of the facts themselves, to their being tensed, while the external perspective is appropriate to the nature of reality as a whole, to there being no privileged standpoint from which the facts should be taken to obtain. It would be comforting to think that a large part of the controversy over the reality of tense has been the result of failing to distinguish between these two different ways in which a perspective may be true to how things are.

I

Issues in the Philosophy of Language

This page intentionally left blank

1
Reference, Essence, and Identity

There are three main concerns within current thinking on modality. One relates to the problem of essentialism, of making sense of *de re* modal discourse. Another relates to the problem of transworld identification, of individuating objects across possible worlds. The third relates to the problem of direct reference, of whether any terms can refer to their bearers independently of how they are described.

It has commonly been supposed that these various problems are connected and that a solution to the one will push us in a certain direction in regard to another. But I shall argue that, once the problems are properly understood, it will be seen that they are quite distinct and that the supposed connections among them are illusory.

Let us first consider essentialism. Different philosophers, and perhaps sometimes the same philosopher on different occasions, have meant different things by the term. But I mean something rather specific, what I have elsewhere called '*de re* anti-scepticism' (Fine [1978]).[1] The *de re* sceptic, or anti-essentialist, characteristically maintains that no object has an essential property independently of how it is described. He will maintain, for example, that Aristotle is essentially a person because the description associated with Aristotle includes the property of being a person; and he will maintain that Aristotle is only accidentally a philosopher since the description associated with Aristotle fails to preclude his not being a philosopher. There is, if I may put it this way, nothing in the object itself to sustain a distinction between its accidental and essential properties.

The reader should be reminded that the paper constituting this chapter was originally written, in very rough form, as the text for a talk given in 1984. Indeed, the front piece to the paper contains the remark that 'it was prepared according to the precept "Write in haste, retract at leisure" '; and, in order to preserve its improvisatory feel, I have subjected it to only the lightest editing. I am grateful to Chris Peacocke whose original remarks on the paper made me realize various ways in which it might be improved. All the footnotes have been added more recently for the purpose of the present publication.

[1] In more recent work (e.g. Fine [1994]), I distinguish essence from *de re* modality. No such distinction is intended here.

It is important when formulating essentialism in this way to guard against a certain misunderstanding. In saying that no object has any essential properties independently of how it is described, I am not suggesting that the term actually used to describe the object need be the *source* of the relevant description. I may say 'the object I was just talking about is essentially a person'. Now there is nothing in this description to imply that the referent is a person. But still it may be true, given that the object I was talking about is Aristotle, that it is essentially a person. We may say, if you like, that no object has essential properties independently of some *canonical* description of the object. But there is no need for the description by which I refer to the object to be the canonical description.

At the heart of *de re* scepticism is a certain metaphysical doctrine about the nature of necessity. It is the doctrine that all necessity is ultimately general. In the fundamental formulation of modal claims, no reference need be made to any individuals.

The original formulation of *de re* scepticism can then be seen as a plausible consequence of this metaphysical doctrine. Given that all necessity is ultimately general, all singular or *de re* necessities need to be explained away. Now although it does not strictly follow, it seems plausible to suppose that this is to be done by reducing all singular necessities to general necessities. And again, although it does not strictly follow, it seems plausible that all singular necessities are to be reduced to general necessities by associating descriptions with the objects involved.[2]

I therefore propose to identify *de re* scepticism with the thesis that all necessity is ultimately general.

It has been more usual to take the issue of essentialism to be one concerning intelligibility. The central question has been: is *de re* modal discourse intelligible? Now certainly we may distinguish between what I have called the 'soft' and the 'hard' *de re* sceptic, with the soft sceptic finding *de re* modal discourse intelligible (though reducible) and the hard sceptic finding it unintelligible (Fine [1978]). But compared to the difference between the sceptic and the anti-sceptic, the difference between soft and hard scepticism strikes me as relatively unimportant. Both the soft and the hard sceptic agree that all necessity is ultimately general. They differ on the question of whether *de re* modal discourse can be reconstructed. But this would appear to be more a matter of degree and not to involve any significant issues of principle. If this is right, then the central question concerns, not the intelligibility of de re modal discourse, but its ontological ground.

---

[2] It is usually supposed that descriptions are associated with the objects one at a time. But one might also associate the descriptions with several objects at a time. This move has significant implications for the problem of transworld identity discussed at the end of the chapter.

De re scepticism, as so understood, is an instance of what I call *generalism*. Generalism holds, in regard to a particular sphere of reality, that all of its facts are ultimately general. It therefore follows that all the putative singular facts (from the given sphere) are to be explained away and, although this is not part of the doctrine proper, it is usually supposed that the putative singular facts are to be reduced to the general facts by associating the objects involved with appropriate descriptions.

Generalism is a very pervasive doctrine within philosophy, and it will be illuminating to give further examples. The problem of *de re* belief, like the problem of *de re* necessity, can be regarded as a generalist issue. In this case, the sceptic or generalist holds that all beliefs are to be reduced to general beliefs through the association of objects with concepts.

Likewise, the celebrated issue over whether existence is a predicate may be seen, on one of its many interpretations, as a generalist issue. The generalist, in this case, maintains that the facts of existence are ultimately general ('existence is a second-level concept') and he is likely to hold that putative singular facts of existence are to be explained in terms of the instantiation of a concept associated with the object.

Another example comes from the metaphysics of cause. Here the generalist holds that all causal facts are ultimately general. He may hold, for example, that cause is primarily a relation between event-types and that it only applies to event-tokens in so far as they have been associated with appropriate event-types. This kind of view is usually associated with Humeans, since it is only between event-types that regularities can properly be said to hold. But it is a view that can be held by theorists of very different persuasions and that might even be held by those who see an irreducibly nomological element in causal connections.

A related example comes from the foundations of probability theory. The generalist, in this case, will hold that all probabilistic facts are ultimately general. He may hold, for example, that probability is primarily a relation between event-types and that it only applies to event-tokens in so far as they have been suitably associated with event-types. But again, it is a view that might be held by theorists of different persuasions and even by those who see probability as an objective dispositional trait of the world.

A final example comes from ethics. The principle of universalizability, on one of its many interpretations, holds that all normative facts are ultimately general. It might be held, for example, that all singular obligations have their source in general obligations—that it is obligatory for me to help other people, for example, only in so far as it is obligatory that anyone in my situation help other people.

In all of these examples, it is important to guard against unwanted cases of generality or singularity. With *de re* belief, for example, the singularity that arises from the individual who has the belief is of no account; it is the

generality of the content of the belief that matters. Another case is more insidious. Suppose one is a bundle theorist; one takes it to follow from the nature of individuals that any fact concerning the individuals will resolve into general facts. Then one will automatically become a generalist in all other areas as well. For example, one will deny that any proposition of the form □S is singular simply on the grounds that the proposition that S is not ultimately singular.

But this is not the intended sense of generalism in the other areas. The modal generalist wants to deny that there are any genuine singular necessities, not because the nature of individuals demands it, but because the nature of necessity demands it. It is therefore important, in formulating the generalist doctrine for a particular area, to indicate the *source* of the generality. The point might be put counterfactually: the modal generalist would want to deny that there are singular necessities *even if* there were genuinely singular facts.

Let us now turn to the theory of direct reference. This theory is standardly taken to maintain that certain singular terms, the 'genuine names', directly refer to their bearers. But what is meant here by 'directly refer'? There are perhaps two main accounts to be found in the literature. According to one, genuine names are rigid designators in the sense of Kripke [1975]. But it is clear—intuitively, from hints in Kripke ([1975], n. 21), and from arguments of Almog and Kaplan and others—that such an account is neither correct nor on the right track.

According to the other account, genuine names are explained in terms of the theory of propositions. Following Russell, it is supposed that propositions may contains objects as individual constituents. The genuine names are then those terms whose semantical role is to put the objects into the propositions (Kaplan [1975]). So what would make 'Aristotle' a genuine name, for example, is its role in enabling a sentence such as 'Aristotle is a philosopher' to express a proposition containing Aristotle.

Now it seems to me that this account of genuine names commits one to much more semantical theory than is strictly required. Let us distinguish between *singular* and *structurally singular* propositions. A structurally singular proposition is one that contains an object as an individual constituent. A singular proposition, on the other hand, is merely one to the effect that an object x has a certain property. Our acceptance of singular propositions could be expressed by our willingness to quantify into contexts governed by the operator 'the proposition that'; where x is Aristotle, for example, we would be prepared to talk of the proposition that x was a philosopher.

For the purposes of direct reference theory, it would be sufficient to claim that genuine names enable one to express singular propositions, without making any commitment on the question of whether the propositions are *structurally* singular. Frege and Russell were both structuralists; they

believed that their propositions had a well-defined internal structure.[3] Russell was an objectualist, he believed that objects could be individual constituents of propositions; Frege was not. It has been usual to detach structuralism from the Fregean position but, for some reason, the structuralism of Russell's position in regard to singular propositions seems to have stuck. However, it is with equal plausibility that it can be detached. So just as the neo-Fregean might identify necessarily equivalent propositions, so might the neo-Russellian. But he had better not suppose that they still have individual constituents, since he will then find himself attributing the same individual constituents to the proposition that Plato is Plato as to the proposition that Aristotle is Aristotle.

But even the commitment to propositions seems to me to be unnecessary. One might distinguish in the usual way between a *de dicto* and a *de re* use of the 'says that' operator. On the *de dicto* use, one could only say such and such, for such and such a sentence; but on the *de re* use, one could also say *of* an object that it had a certain property. Note that the use of the 'says that' operator need not be taken to commit one to propositions. It could now be maintained that the genuine names were those terms that enabled one to make *de re* sayings, to say *of* an object that it had a certain property. Thus what would make 'Aristotle' a genuine name, on this account, is its role in enabling one to use such a sentence as 'Aristotle is a philosopher' to say of Aristotle that he is a philosopher.

However, even this account is not quite satisfactory; for it requires us to insist on a strict reading of *de re* saying, one according to which the Fregean could not properly claim that one might say of an object that it had a certain property. But how is this strict reading to be made out? (To be scrupulous, there is a similar problem for the earlier propositional accounts, since they admit of a weak reading under which even a Fregean might admit that one can express singular propositions. But the problem does not seem so acute in this case since the unintended reading is so artificial.)

We have here a problem very similar to the one of distinguishing the *de re* sceptic from the anti-sceptic. For the sceptic, *de re* attributions of modality are always mediated through a description. In the same way, for the Fregean, *de re* sayings will always be mediated through a description or the like; to say of Aristotle that he is a certain way is to say, for a suitable description D or the like, that the D-er is that way.

I suggested that at the heart of *de re* scepticism was the metaphysical doctrine that all necessity is ultimately general. In the same way, I would like to suggest that at the heart of the Fregean position is the doctrine that all saying is ultimately general or, to put it the other way round, at the heart of the

---

[3] It might be more accurate to say that there are strong structuralist tendencies in Frege's thought.

Russellian position is the doctrine that some sayings are ultimately singular. The point can be made even more simply in terms of reference. Facts of reference for the Fregean are general; to refer is to stand in a certain relationship to a concept that is uniquely satisfied. For the Russellian, some facts of reference are irreducibly singular; irreducibly involved in the fact that I refer to a certain object is the object itself.

What is the bearing of direct reference theory on the thesis of essentialism? It has been common to suppose that a positive stand on direct reference implies, or at least makes plausible, a positive stand on essentialism. For let us grant that necessity is a coherent predicate of sentences; and let us assume that the direct reference theory is correct. Then it seems to follow that we can make sense of *de re* modal claims independently of any appeal to descriptions or concepts. For suppose we wish to say *of* Aristotle that he is necessarily a person. Then we need only take a genuine name of Aristotle— 'Aristotle', say—and using that name, say that the sentence 'Aristotle is a person' is necessarily true.

   I have here talked of a transition from genuine names to *de re* modality. But I could equally well have talked of a transition from singular propositions to *de re* modality. In that case, it would have been necessary to start off with necessity as a predicate of propositions. I could then have made the *de re* claim about Aristotle by saying that the proposition to the effect that Aristotle is a person is necessarily true.

   Such a line of argument can, I think, be discerned in the work of Plantinga, Kaplan, and Kripke. It is implicit in Plantinga's reduction of *de re* to *de dicto* modal discourse ([1974], 27–43). Kaplan in his paper 'How to Russell a Frege-Church' (Loux [1979], 218–19) explicitly argues from the acceptance of singular propositions to Haecceitism with respect to transworld identity, which in its turn is meant to imply *de re* anti-scepticism. Kripke seems similarly to argue ([1980], 49) that 'it is *because* we can refer (rigidly) to Nixon, and stipulate that we are speaking of what might have happened to *him* (under certain circumstances), that "transworld identifications" are unproblematic in such cases'; and his constant appeal to facts of rigidity in establishing essentialist claims would appear to suggest that the intelligibility of those claims, at the very least, could be made to rest on the existence of the appropriate rigid designators. Even when philosophers have not been so explicit in making the transition, I think it is fair to say that they have often felt that the use of genuine names removes the old Quinean difficulties over *de re* modality.

   One possible objection to this line of argument goes as follows. Look, just because I accept necessity as a predicate applicable to certain sentences of my language, it does not follow that I accept it as a predicate applicable to them all. Suppose I start out life as a Fregean and am ready to accept necessity as a universal predicate of sentences. Later, as a result of reflections

in the philosophy of language, I come to endorse a Russellian position and, as a consequence, admit into my language sentences containing genuine names. I am not then under any obligation to extend the application of my necessity predicate to this enlarged class of sentences.

However, this objection strikes me as very weak. Necessity, if applicable to sentences at all, appears to be universally applicable; there appears to be no good reason to allow its application to certain sentences and not to others.[4] There is also a special reason for supposing it to be universal in its application. Given a sentence S, either it or its negation −S is true. If S is true, then surely S is possibly true; and, by the same token, if −S is true then surely −S is possibly true. So possibility is applicable to any sentence or its negation; and it is hard to see how this could be so unless necessity had universal application.

The weakness of this objection underscores the strength of the original line of argument. Given a Russellian position, there really appears to be intellectual pressure on us to extend the notion of necessity to sentences containing genuine names and thereby endorse the legitimacy of *de re* modal discourse.

All the same, it seems to me that there are genuine weakenesses in the transitionalist's position. Suppose I start out life as a *de re* sceptic; I believe that objects only have essential properties in virtue of associated descriptions. Later I become converted to Russellianism and so am prepared to apply necessity to singular propositions or to sentences containing genuine names. I am still going to believe that those *de re* propositions or sentences are necessary in virtue of the appropriate association of objects with descriptions.

But still, it may be protested, does not the Russellian stand at least give us the *intelligibility* of *de re* modal discourse? Even if it does not settle the issue between the sceptic and the anti-sceptic, does it not at least force the sceptic into a 'soft' position?

I have already tried to indicate why I do not think that the question of intelligibility is the critical issue. But even here, it seems to me, the argument does not really work. Suppose that this time I start out life as a hard sceptic; I do not believe that there is a unique or reasonably specific way of associating canonical descriptions with objects so as to yield *de re* modal claims of determinate truth-value (certain trivial cases aside). Then my hard-line attitude will persist after my conversion to Russellianism. Of course, I will admit the grammatical correctness of the application of the necessity-predicate to the names of sentences containing genuine names or to the names of singular

---

[4] Oddly enough, in a more recent paper (Ch. 9 of this volume), I give reasons, tangential to the issue at hand, for thinking that the necessity-predicate might selectively apply to some sentences and not others.

propositions; but that is not to say that I will have any good reason for thinking that the resulting modal claims will have a determinate truth-value.

But am I not being unfair to the transitionalist? If there is a well-determined notion of necessity that applies to general sentences or propositions, then, if there are indeed singular sentences or propositions, shouldn't it extend to them in a way that does not call for special explanation in terms of associated descriptions? Shouldn't the extension to those sentences or propositions already be implicit in our understanding of necessity?

The *de re* sceptic might even be willing to concede that this is so and he might even agree with his opponent on the general sentences or propositions to which the base notion of necessity will apply. His disagreement will then turn on how he thinks this base notion is to apply to singular propositions. His opponent will presumably have a metaphysical notion of necessity in mind, one that is capable of applying differentially to different individuals. But the *de re* sceptic will find himself incapable of understanding such a notion. For him, the notion of necessity can hold only in virtue of general connections and so its application will be incapable of discriminating between different individuals.

It is here important to distinguish between someone who endorses a metaphysical concept of necessity but who holds, as a substantive metaphysical principle, that the application is blind to individual differences, and someone who is only prepared to endorse a concept of necessity that holds in virtue of general connections and who is thereby *obliged* to accept that its application is blind to individual differences. Both philosophers accept the same *de re* modal claims. But the former will admit that he is in genuine dispute with his opponent; while conceding that different individuals could in principle differ in their qualitative essential properties, he will still maintain that the nature of such individuals is that they never do. On the other hand, the latter philosopher will not recognize that there is anything sensible to dispute. It follows, not from the nature of the individuals, but from the very concept of necessity, that no two individuals can differ in their purely qualitative essential properties.

But why, it may be retorted, all the fuss about the *metaphysical* concept of necessity? It has been conceded that a Russellian position may force one to extend a concept of necessity from general to singular truths in a way that is not dependent upon associated descriptions. So what does it matter whether or not this is a concept of metaphysical necessity?

The answer is that the issue of *de re* scepticism is of interest only for the concept of metaphysical necessity or for concepts of necessity that are capable of discriminating among different individuals. Take a concept, such as that of logical necessity, that is incapable of discriminating among different individuals; if one individual necessarily has a certain qualitative

property, then so does any other. For such a conception of necessity, there is—or never should have been—any substantive problem about its *de re* application.[5] If we ask, for example, 'when is it logically possible for a given individual to have a certain property (one not involving the object itself)?', then we may simply look to see whether the property itself is logically consistent.

The point is a general one in regard to the issue of generalism. What the hard-line generalist denies is that there is any intelligible singular application of the given notion in the *intended sense*, not that there are any intelligible singular applications at all. We can imagine a frequentist, for example, who sees no sense in what people are trying to say then they say that the probability of *this* penny turning up heads right now is one half. But he can admit that this penny being tossed and this penny turning up heads right now are degenerate cases of event-types and that, so regarded, the statement that the penny will turn up heads will have probability 1 if the penny turns up heads and probability 0 otherwise. But this is not the intended or even an interesting concept of probability; and it is in the same kind of way, it seems to me, that the Russellian may be willing to recognize an unintended and uninteresting concept of *de re* modality.

It has to be said that there has been a failure on the part of philosophers to appreciate that there is a *distinctive* problem of the *de re* for the concept of metaphysical modality. It often seems to be assumed that the problem is a general one and that to allay doubts in regard to one concept of necessity is to allay doubts in regard to them all. An example of this sort is perhaps afforded by Kripke's Introduction to 'Naming and Necessity' ([1980], 16–20). He there points out that there is no difficulty in talking of the *de re* possibilities for two dice, quite independently of how the dice might be described. He then goes on to suggest that there is likewise no difficulty in talking of the *de re* possibilities of the universe. But to this, the *de re* sceptic might object: the possibilities for the dice (at least in so far as I can understand them) are logical possibilities and so, of course, there is no difficulty in taking them to be *de re*; but the possibilities for the universe are meant to be metaphysical possibilities and so I still do not see how such possibilities might be capable of discriminating among different individuals unless it is by reference to some implicitly associated descriptions.

Another example is provided by contemporary puzzlement over Russell's attitude towards existence claims. Why did he say that the sentence 'Socrates exists' is meaningless? Why could he not see with Moore and the rest of us

---

[5] Of course, even in the case of logical necessity, we will want to say that it is logically necessary that Aristotle is Aristotle though not logically necessary that Aristotle is Plato. This means that in reducing singular to general necessities we must at least be taken to be cognizant of when two individuals are the same. The issue of making sense of the *de re* application of logical and analytic necessity is further discussed in Ch. 2 and 3 below.

that, for 'Socrates' a genuine name, the sentence 'Socrates exists' is true, and contingently true at that?

But suppose, as is very plausible, that Russell was only prepared to accept the *de re* for the concept of logical necessity at best, and not for the concept of metaphysical necessity. We could then have two possible explanations of Russell's position. According to the first, 'x exists' is taken to mean '$\exists y(x = y)$'. So in the only applicable sense of necessity, the sentence 'Socrates exists' is necessary and is therefore *empirically* meaningless. According to the second (more plausible) explanation, 'x exists' is to have a sense according to which existence claims can be contingent. But since in the only applicable sense of necessity, the sentence 'Socrates exists' is necessary, 'exists' cannot bear its intended sense and so the sentence is *literally* meaningless.[6]

If I am right, there is no plausible route from Russell to Aristotle, no plausible way of getting from direct reference to essentialism. But what of the reverse direction? Is the Russellian attitude somehow necessary for anti-scepticism? Can one adopt a Fregean position in the philosophy of language and yet still be anti-sceptical in one's attitude towards *de re* modality?

Here again, it strikes me that there is no plausible route from the one position to the other. Suppose one is a committed Fregean, but wants to say of Aristotle, quite independently of how he is described, that he is necessarily a person. Then it won't do to say that the sentence 'Aristotle is a person' is necessary or that the proposition expressed by this sentence is necessary; for the necessity will then attach to the sense of the name, if I may put it that way, and not to the bearer. The name will be analogous in its use to a definite description; and so one will not be saying of Aristotle that he is necessarily a person, but only that it is necessary that the D-er (whoever he might be) is a person. So the *de re* claim cannot be expressed in the obvious way by attributing necessity to a sentence or proposition; and it is this that makes the transition from the Fregean position to *de re* scepticism seem so compelling.

But this is by no means the end of the matter. For one might suppose that just as there is a concept of necessary truth, so there is a concept of necessary

---

[6] In his response to my paper, Chris Peacocke pointed out that Russell's reasons for rejecting the meaningfulness of individual existence claims are 'stateable without any implicit rejection of metaphysical necessity'. This is perhaps sometimes true. But not always. For example, in the discussion of lecture V of the Lectures on Logical Atomism (Russell [1956], 241), he says, 'There is no sort of point in a predicate which could not conceivably be false. I mean, it is perfectly clear that, if there were such a thing as this existence of individuals that we talk of, it would be absolutely impossible for it not to apply...'. Russell is here presupposing that for an individual existence claim to be meaningful, or to have 'point', it must be possible for it to be false; and he is therefore implicitly rejecting a metaphysical notion of possibility under which the claim would be false.

satisfaction (or possession) that corresponds to the simple non-modal conception of satisfaction (or possession). One would then have a relation that held between an object and a condition (or an object and a property) just in case the object necessarily satisfied the condition (or necessarily satisfied the property). And armed with such notions, there would then be no difficulty in expressing *de re* modal claims without making any reference to associated descriptions. We could say that Aristotle was necessarily a person, for example, by saying that the relation of necessary satisfaction holds between Aristotle and the condition 'x is a person' or that the relation of necessary possession holds between Aristotle and the property of being a person.[7]

Somewhat surprisingly, the objection works equally well against the form of transitionalism presented by Kaplan [1975] (Loux [1979], 218–19). He there argues as follows. Suppose the identity of individuals across possible worlds is given (this corresponds to *de re* anti-scepticism). Then we can make sense of singular propositions, at least on their possible worlds representation; for a singular proposition can simply be identified with the set of possible worlds in which the given object has the given property. But our objection now takes the following subtle form. For a Fregean who endorses essentialism, a possible world will not correspond to a set of propositions, namely, the set of propositions true in that world, but to a set of object-property pairs, namely, those in which the object possesses the property in the given world. But when possible worlds are so construed, a set of worlds cannot properly be taken to represent a proposition at all, let alone a singular proposition.

We therefore see that there is no route from either direct reference theory to *de re* anti-scepticism or in the other direction. As far as I can see, the only connection between the two issues is that one's position on direct reference may act as a constraint on how one is to express *de re* claims of necessity, should one be an anti-sceptic. If one is not a Russellian, it is not adequate to suppose that necessity is sentential in its role, serving either a predicate of sentences or propositions or as an operator on sentences. One must suppose that necessity serves in a more direct way to link an object to what is predicated of it.

It has to be recognized that there are two quite distinct issues here: one in the philosophy of language; and the other in the metaphysics of modality. There is the question of whether the mechanism of reference requires a descriptive intermediary, of whether one can refer to anything independently of how it is described; and there is the question of whether the mechanism of necessary attribution requires a descriptive intermediary, of whether one can attribute a necessary property to an object independently of how it is

---

[7] This corresponds, of course, to Quine's [1956] relational treatment of *de re* constructions.

described. The difference might be put in terms of 'quantifying in': is there irreducible (objectual) quantification into contexts governed by the proposition-forming operator?; and is there irreducible quantification into contexts governed by the necessity operator? Or the difference might be put in terms of generalism: are there irreducibly singular facts of reference?; and are there irreducibly singular facts of necessity? There is a manifest absurdity in conflating other cases of the issue of quantifying in or of generalism, such as those concerning belief and obligation; and there is something of an equal absurdity in conflating the present two cases of the issue.

Given that the two issues are independent, why have philosophers so persistently tried to relate them? One specific reason has to do with the narrow terms in which the problem of *de re* modality was sometimes conceived. One was supposed to start off with necessity as an operator on sentences or perhaps as a predicate of sentences or propositions and, upon this basis, to explain the notion of *de re* necessity. With the problem conceived in these terms, the Fregean position would appear to constitute an insuperable difficulty, since the sense of a name would get in the way of the proper expression of the *de re* modality; while with the adoption of the Russellian position, the difficulty would disappear, since a proper name would have no sense and so there would be no sense that might get in the way.

Another, more general, reason has to do with the way in which discussion of *de re* modality was originally oriented towards issues in the philosophy of language. None of the earlier parties to the dispute—Carnap, Church, or Quine—really cared about *de re* metaphysical necessity. What they cared about were general issues in the philosophy of language: the use of extensions versus intensions; the desirability of having an extensional language; the proper understanding of quantification. As a consequence, the topic of modality became merely a battleground upon which to dispute over the various issues in the philosophy of language, and the distinctively metaphysical issues about modality were largely ignored. Although philosophers have now become more attuned to *de re* metaphysical necessity as a concept of interest in its own right, the feeling has persisted that the issues to which it gives rise must somehow be related to those in the philosophy of language.

We come to the last locus of discussion in the debate over *de re* modality, the issue of Haecceitism. Very roughly, we may say that this is the issue of whether individuals can be identified in terms of their qualitative properties and relations. However, it is important to distinguish between two forms of anti-Haecceitism, the modal and the metaphysical, that differ according to the ground for the identification.

Modal anti-Haecceitism is a doctrine about the nature of necessity. I take it, by definition, to be the counterpart within the possible worlds framework

to the doctrine of *de re* scepticism. So to work out what it is, we need only see how sceptical concerns regarding *de re* modality translate into concerns regarding the structure of possible worlds.

The underlying ontological position that all necessity is ultimately general corresponds to the claim that all possible worlds must ultimately be given in purely qualitative terms. It must be possible to specify the worlds without making reference to any individuals. This latter requirement has a rather nice technical formulation. Suppose we have a model or representation $\mathscr{A}$ of the set of possible worlds. Then in such a model we will be forced to settle questions of 'external' or cross-world identity; for given an individual x from one world and an individual y from another world, either x will be identical to y or x will be distinct from y. But if the possible worlds are given in purely qualitative terms, such external identities should be incidental to the representation. It therefore follows, if $\mathscr{B}$ is 'locally isomorphic to $\mathscr{A}$, i.e. differs from $\mathscr{A}$ only in the identities of the individuals that figure in the different worlds, then $\mathscr{B}$ should serve as an equally good representation of the underlying modal reality (cf. Kaplan [1975], 221, and Fine [1978], 126, 136).

The problem of making sense of *de re* modality in terms of *de dicto* modality also has a nice counterpart within the possible worlds framework. What is required is that we make sense of the attribution of a property to the very same individual from one world to another; and what this seems to require is that the individual not only have an 'internal' identity within each world but also an 'external' identity across worlds. We are thus led to the celebrated problem of transworld identity, according to which the identity of an individual, as given in one world, must be extended to all other worlds in which it might exist.

In contrast to the modal doctrine, metaphysical anti-Haecceitism is a doctrine concerning the identity of individuals. It states that the identity of individuals—or, at least, of certain individuals—is to be explained in terms of their purely qualitative features or in terms of their qualitative relationships to other individuals. Very roughly, metaphysical anti-Haecceitism is a version of the bundle theory ('there is nothing to a particular over and above its properties'), while metaphysical Haecceitism is a version of the doctrine of 'basic particulars'.

The metaphysical form of anti-Haecceitism is quite different from *de re* scepticism and hence from the modal form of anti-Haecceitism. Indeed, it seems to me that the two positions are completely independent of one another, with all four combinations of them being coherent. One combination, of modal and metaphysical Haecceitism, is unproblematic. It represents the 'naive' view, according to which neither the identity of individuals nor the application of *de re* necessity to individuals stands in special need of explanation. However, the other combinations are all problematic in their own way and worthy of more detailed study.

Consider the combination of metaphysical and modal anti-Haecceitism. On this view, the identity of individuals is to be explained in qualitative terms and the necessities concerning individuals are to be explained in terms of general necessities. We may think of the proponent of this view as facing two problems. The first is to identify the individuals of the actual world (however it might turn out) in suitably non-modal qualitative terms; the second is to explain how the notion of necessity applies to these individuals, as so individuated. Suppose, for example, that our proponent holds a bundle theory of personal identity. Then to solve the first problem, he would have to say which bundles of experiences constituted a person; and to solve the second problem, he would have to explain how to make sense of *de re* modal assertions concerning persons, as so construed. If he were unable to solve the second problem (which is by no means trivial), then he would not think 'so much the worse for the concept of a person' but 'so much the worse for the concept of *de re* necessity'.

Now consider the combination of metaphysical Haecceitism and modal anti-Haecceitism. On this view, the identity of individuals is simply given, but *de re* necessity needs to be explained in terms of general necessity. The proponent of this view does not face the problem of explaining the identity of individuals (unless it is to explain how their identity might just be given), but he does face the problem of explaining how *de re* modal claims might be made concerning the primitively given individuals.

But isn't there a difficulty here? The metaphysical Haecceitist demands that the actual world (however it might turn out) be described in terms of the given individuals, whereas the modal anti-Haecceitist demands that the possible worlds, including the actual world, be described in qualitative terms. But isn't this to place conflicting demands on the actual world?

The solution to this difficulty lies in the seeming paradoxical claim that, for the purposes of describing the respective forms of Haecceitism, the actual world should not be treated as a special case of a possible world. When I describe the actual world in the most fundamental terms, I make reference to individuals. But this does not mean that when I describe the possible worlds in the most fundamental terms, I must also make reference to individuals; for as a *de re* anti-sceptic, I will think that ultimately there are no *de re* possibilities. When I describe reality, I say that there are such and such individuals but, when I describe possible realities, I do not mention individuals, not because I do not believe in them but because I do not believe that there are any genuine possibilities concerning them.

This difference becomes especially clear on the Fregean view according to which necessity is a predicate of sense. For when I describe the actual world, I shall use the names of individuals; but when I come to describe the possible worlds, I shall use the names of individual concepts. So in describing the

actual world and the possible worlds, I shall not even be talking about the same things.

To give an adequate formal representation of this and other combinations of the Haecceitist positions, it is better to use a model that is constituted not only by a set, to represent the domain of possible worlds, but also by a distinguished member of that set, to represent the actual world. Under the modal construal, anti-Haecceitism would then reside in the possibility of altering the pattern of cross-world identities within any one model, Haecceitism in the necessity of keeping it fixed. Under the metaphysical construal, on the other hand, anti-Haecceitism would reside in the possibility of altering the identity of the individuals in the actual world across models, Haecceitism in the necessity of keeping it fixed. The present combination of positions would therefore consist in our being required to keep the identity of the individuals in the actual world fixed but otherwise being allowed to vary the identities of the individuals across possible worlds.

Consider finally the combination of metaphysical anti-Haecceitism and modal Haecceitism. On this view, *de re* necessities are in no need of special explanation but the identity of individuals must be explained in qualitative terms. However, in contrast to the case of modal anti-Haecceitism, it is no longer sufficient, in explaining the identity of individuals, to explain their identity in the actual world alone; for then the application of necessity to the individuals, so identified, becomes problematic. It is necessary, at the very least, to explain the identity of individuals across possible worlds. Suppose, as before, that one holds to a bundle theory of personal identity. Then in explaining the identity of a person one must not only say when a bundle of experiences constitutes a person, one must also say when person-constituting bundles of experiences from different worlds constitute the same person.

The point might be put this way. For both the metaphysical and the modal anti-Haecceitist, the claim that an individual necessarily has a certain property may be problematic. But for the modal anti-Haecceitist, the problem lies with the notion of necessity, its extension to the *de re* will have been insufficiently specified; while for the current metaphysical anti-Haecceitist, the problem lies with the nature of the individual, its identity across worlds will have been insufficiently specified. It may be genuinely indeterminate, for example, whether it is possible for a table to have been made with one fewer atom. For the modal anti-Haecceitist, this is an indeterminacy in the notion of *de re* possibility; we may be perfectly clear what a table is, it is just that we have not worked out which concept to associate with a table for the purpose of interpreting *de re* modal claims. For our metaphysical anti-Haecceitist, on the other hand, the indeterminacy lies in the identity of the table; it is really not clear what it is.

The present combination of metaphysical anti-Haecceitism and modal Haecceitism gives rise to a problem of identity across possible worlds,

what one might call the anti-sceptical problem of transworld identity. This problem might be illustrated by the case of ordinary material things—chairs, tables, and the like. Suppose one is a metaphysical anti-Haecceitist with regard to such things, one thinks that their identity is to be explained in terms of their relation to matter, or what makes them up. Given that one is also a modal Haecceitist, it would then appear to be necessary, in explaining the identity of such things, to specify not only the conditions under which they continue to exist but also the conditions under which they might have been. One would need to say, for example, what variation was possible in the original constitution of a chair—whether it might have been made from altogether different matter, from slightly different matter, or whatever.

Once translated into the possible worlds framework, this then becomes a problem that might aptly be described as a problem of transworld identity. We suppose that each world is described in terms of the qualitative relations of ordinary material things to their matter; we do not say '*this chair* is constituted in such and such a way over time', but 'there is a chair that is constituted in such and such a way over time'. The problem is then to identify the material thing across possible worlds on the basis of their qualitative relations.

However, this metaphysical version of the problem of transworld identity needs to be sharply distinguished from the traditional or modal problem of transworld identity (cf. Kripke [1975], 50–3 and Plantinga [1976], 88–102). For, first, the new problem is one that presupposes modal Haecceitism; it can therefore arise only when the other is not in question. But, second, the new problem is one that is genuinely about the identity of the individuals; the original problem, despite its name, is not about the identity of individuals, but about the notion of necessity and how it might be extended to individuals whose identity is already assumed to be given.

What makes it so easy to confuse the two problems is that they can both be given very similar formulations. In both cases, we may suppose that the possible worlds are specified in qualitative terms; and in both cases, we are required to identify individuals across possible worlds on the basis of their qualitative features. However, the source of the underlying generalism and the point of the identification are very different in the two cases.

As we have seen, our metaphysical anti-Haecceitist is prepared to concede in principle that there are singular propositions; it is just that he thinks that the nature of individuals is such that singular facts resolve into general facts and hence singular possibilities into general possibilities. The modal anti-Haecceitist, on the other hand, is not even prepared to concede in principle that there might be singular possibilities.

In the same way, our metaphysical anti-Haecceitist regards his individuals as having a genuine transworld identity. It is just that he believes, on

metaphysical grounds that the transworld identity of an individual can be recovered on the basis of its qualitative features in each world. On the other hand, our modal anti-Haecceitist does not regard his individuals as having a genuine transworld identity. He takes the identity of individuals to be given and then thinks of the identificatory paths as mere artificial devices, designed to secure a meaning for cross-world attributions. It is not as if the objects had a modal underside, a 'fifth dimension', and that he was attempting to explain their identity in this dimension, just as other philosophers attempt to explain the identity of objects in space and time. Rather the objects come first, and the modal dimension is something tacked on later. The identificatory paths that emanate from the object are not parts of the object, so to speak, but appendages; they are the picturesque counterparts to the individual concepts that his more pedestrian colleague, the *de re* sceptic, had wished to associate with an object in the first place.

The point becomes especially vivid if we compare the modal anti-Haecceitist with his doxastic counterpart. The doxastic anti-Haecceitist holds that all beliefs are ultimately general and hence faces a problem of identification, similar to the modal anti-Haecceitist's, only with doxastic alternatives in place of possible worlds. Now he is certainly not going to suppose that individuals are extended in doxastic space and that, in constructing identificatory paths across the doxastic alternatives, he is somehow explaining what they are; and no more should the modal anti-Haecceitist.

The failure to distinguish the modal and metaphysical versions of Haecceitism has resulted in a number of confusions. Let me mention four. There is first the misleading practice of using the terminology appropriate to one form of the doctrine in the description of the other. I think that this sort of terminological trespass has been fairly common and, as an example, I might mention the account of Haecceitism given in Kaplan [1975], 217. He there says that, on a Haecceitist view, 'a common "thisness" may underlie extreme dissimilarity' and that 'we can meaningfully speak of a thing in itself—without reference . . . to . . . defining qualities'. (It seems to be in the same spirit that in his paper 'Transworld Heir Lines' (Loux [1979], 99) he associates the relativistic position on transworld identity with the 'bundles of qualities metaphysics'.) It is Kaplan's aim here to characterize the modal form of Haecceitism, corresponding to *de re* anti-scepticism. But if this doctrine is combined with metaphysical anti-Haecceitism, then it will be denied that there is a 'thisness' that may underlie extreme dissimilarity or that individuals can be given independently of their defining qualities.

Another confusion has to do with the possibility of relativizing the doctrine of Haecceitism to a class of individuals. If Haecceitism bears the modal sense, then it is bizarre to suppose that the doctrine applies differentially, to

some individuals and not to others.[8] If the notion of necessity is such that individuals cannot have essential properties independently of how they are described, then this is a consideration that tells equally against one kind of individual as opposed to another.

On the other hand, if Haecceitism bears the metaphysical sense, it is extremely plausible to suppose that the doctrine might apply differentially. We have talked of a world being given in *purely* qualitative terms. But it is much more plausible to suppose that some individuals are just given and that the other individuals are to be identified on the basis of their qualitative relationship to the given individuals. We may suppose, for example, that experiences are given and that selves are to be identified on the basis of their qualitative relationship to the experiences, or that matter, or what constitutes it, is given and that all other material things are to be identified on the basis of their relationship to the matter. Kaplan's remark that 'it would be more exact to speak of Haecceitism *with respect to* a given kind of entity' ([1975], 217) is therefore best understood as applying to the metaphysical rather than the modal form of Haecceitism.

A third confusion concerns the analogy between the identity problems for modality and time. There is a traditional problem of identity over time for various kinds of entity—for material things, for example, or persons. The problem of transworld identity is then thought to be analogous to the problem of identity of time, but with modality substituted for time. In this way, the modal version of the problem acquires a certain degree of innocence by association.

But the analogy is very weak and potentially very misleading. The traditional problem in the area of modality has to do with the notion of necessity, with the basis for its application *de re*. The traditional problem in the area of time has to do with the metaphysics of individuals, with the basis for their identity over time. But if the two are treated as analogous, then the modal form of Haecceitism is likely to be confused with its metaphysical form.

The true temporal analogue of the traditional question over modality is an issue that is very weird. Let us grant, if only for the sake of argument, that there are genuinely singular facts in the present. The issue is then whether the notions of the past and the future are such as to permit genuinely singular facts in the past or future or whether all such apparently singular facts should be explained on the basis of general facts in the past and future. But the thesis that if there are present singular facts then there are past and future singular facts (in so far as there are any past or future facts at all) seems so utterly compelling that it is hard to see why any philosopher would want to think otherwise.

---

[8] An exception could perhaps be made for properties or concepts as opposed to individuals, though even here there is some awkwardness in distinguishing between the two kinds of case.

On the other hand, the modal analogue to the traditional problem of identity over time does seem to be a genuine problem. It is the problem that we previously identified as arising once modal Haecceitism was combined with metaphysical anti-Haecceitism. Given that there are ultimately only facts concerning certain individuals, how are other individuals to be identified on the basis of their relationship to the underlying individuals?

It strikes me that the analogy between these two problems runs quite deep. In each case, we have a problem that, intuitively, is about the identity of an individual. In specifying its identity conditions through time or through worlds, we are saying more exactly what the individual is. In each case, the formulation of the problem has the same general form. At each time or at each world at which the individual exists is associated a certain 'presentation' of the individual, given by its purely qualitative properties or by its qualitative relationship to other individuals. The problem is then to explain upon what basis the two presentations are presentations of the same individual. Finally, the motivation for the problem in each case is the same. It is supposed that features of the objects whose identity is in question are supervenient upon features of the objects whose identity is not in dispute, at least for the purposes of the problem at hand. This then gives rise to the problem of explaining when two presentations are presentations of the same object in terms of the supervenience base; and this, in other terms, is merely the problem of identity. Indeed, under quite reasonable assumptions, the identity problem can be seen to be the *critical* problem to be solved in making out a general claim of supervenience; once it is solved, the truth of the more general claim becomes clear.

Although the analogy between the two identity problems runs deep, it seems to me that very different considerations will be plausible in the two cases. Consider the question of which individuals should be taken as basic in explaining the modal or temporal identity of all other individuals. In the temporal case, there are two views that have some plausibility, even if ultimately they are to be rejected. According to one, the basic individuals are all instantaneous (this is the 'sliced' view of enduring things) while, according to the other, the basic individuals are material packets whose temporal, and perhaps even whose spatial, boundaries are indeterminate (this is the 'stuffy' view of enduring things). However, the modal analogues of these two views have no plausibility at all. The counterpart to the first is that the basic individuals are world-bound; and the counterpart to the second is that the basic individuals are some sort of stuff whose modal boundaries are indeterminate. But the idea that I might be made up of more basic world-bound individuals seems absurd; and what the modal stuff could be defies the imagination. Of course, one might hold a queer essentialist view according to which I am world-bound; but if I am not

world-bound, it is hard to see how my identity might be better understood in terms of objects that are.

The final possible confusion has to do with the question of identification. The problem of making sense of *de re* modality is usually taken, in its possible worlds form, to be the problem of identifying individuals across possible worlds: each world is given in purely qualitative terms; and each individual is then to be identified in any world on the basis of its purely qualitative features.

On this way of understanding the problem, certain cases give rise to special difficulties. We may have two objects which, on independent grounds, we know to be distinct and yet have an equal right to be identified with a third individual. We may have a world w, for example, containing two indiscernible spheres, S_1 and S_2. Suppose now that we take a world w' in which there exists a single sphere just like S_1 and S_2. Since it seems possible that either sphere should exist on its own, just as it is in w, there is the problem of saying which of S_1 and S_2, if either, is to be identified with the sphere in w'.

Various ingenious solutions might be given to this puzzle and to others like it. But it is important to appreciate that the very terms in which the problem is posed are misconceived. It is perfectly appropriate for the *de re* sceptic to postulate a third world w″ qualitatively indiscernible from w' (what we might call a 'duplicate' of w'); S_1 can then be identified with the sphere in w' and S_2 with the sphere in w″.

But isn't this cheating? Mustn't the worlds for the sceptic be given in purely qualitative terms, thereby ruling out the possibility of duplicates? And even if we allow duplicates, such as w' and w″, then on what basis do we identify S_1 with the sphere in w', say, rather than in w″?

However, there is nothing to *stop* the *de re* sceptic from introducing duplicates and identifying individuals with the individuals in them in an arbitrary manner. The aim of the sceptic is to reconstruct *de re* modal discourse; and he can do this in any way he likes as long as the resulting reconstruction is intuitively acceptable. So what is there in this intuitive constraint that might block the use of duplicates?

The problem facing the *de re* sceptic might be posed in the following terms. He starts off with a system of worlds $\mathscr{A}$ in which cross-world identities of individuals are taken to be arbitrary; and from this he must construct a system of worlds $\mathscr{A}^*$ in which the cross-world identities of individuals are taken to be fixed. As far as I can see, there are only two formal constraints that the normalizing function* should satisfy. One is that the structure of $\mathscr{A}^*$ should not depend upon how the cross-world identities in $\mathscr{A}$ happen to be fixed ($\mathscr{A}^*$ should be isomorphic to $\mathscr{B}^*$ when $\mathscr{A}$ is locally isomorphic to $\mathscr{B}$). The other is that for any world in $\mathscr{A}^*$ there should be a duplicate in $\mathscr{A}$ and vice versa ($\mathscr{A}^*$ should be locally

isomorphic to $\mathscr{A}$). The first of these requirements corresponds to the requirement that the basis of the reduction should be *de dicto*; and the second corresponds to the requirement that the reduction should be conservative, that the *de dicto* truths should remain the same after the reconstruction as they were before.

But there is nothing in these formal constraints to bar the presence of duplicates in $\mathscr{A}^*$; and so, if I am right about the adequacy of these constraints, there is nothing in the problem of reconstructing *de re* modal discourse to prevent the sceptic from using duplicates. Of course, he may hold essentialist opinions that prevent him from entertaining duplicates; he may be a metaphysical anti-Haecceitist, for example, and take all singular facts to be supervenient upon general facts. But then it is his metaphysical anti-Haecceitism that is getting in the way of his accepting duplicates, and not his modal anti-Haecceitism.

The formal point might be made even more forcefully if we regard the construction of $\mathscr{A}^*$ from $\mathscr{A}$ in the following way. We first take the domains of individuals from different worlds of $\mathscr{A}$ to be disjoint; and we then add all the worlds that can be obtained by 'permuting individuals'. We thereby obtain a new, very large, system of worlds $\mathscr{A}^+$. The problem of obtaining $\mathscr{A}^*$ from $\mathscr{A}^+$ (and hence from $\mathscr{A}$) can now be regarded as the problem of cutting $\mathscr{A}^+$ down to size, of discarding unwanted worlds. Viewed from this perspective, we don't even have anything that *looks* like a problem of identification; we have, instead, the problem of narrowing down an excessive range of *de re* possibilities.

In regard to the modal Haecceitist's problem of transworld identity, the situation is quite different. For the modal Haecceitist, the *de re* possibilities and hence the possible worlds are just given; there is no more room for filling out modal space with extra worlds than there is for filling out the dimension of time with extra instants. So if the possible worlds can be completely described in terms of the underlying individuals, there is not even the possibility of duplicate worlds. Nor can we readily admit distinct but indiscernible individuals within a world, as with the spheres S_1 and S_2 in w, since, in the absence of duplicate worlds, it is not clear what might constitute the existence of S_1 rather than S_2 within another world.

So, in regard to the modal Haecceitist, we see that there is a genuine issue of transworld identification. Given an individual from one world (qualitatively given) and an individual from another world, we can sensibly ask, if we are metaphysical anti-Haecceitists, whether they are the same or distinct and expect an answer on the basis of the information given. It has been almost universally assumed that the *de re* sceptic faces a similar problem of transworld identification. But this is not so. He can allow both that there might be duplicate worlds and that there might be no qualitative means of tracking an individual from one world to another.

2

The Problem of *De Re* Modality

I

Quine has two arguments against quantifying into modal contexts. Each begins in the same way. It is agreed that, for a quantified modal statement such as $\exists x \Box (x > 7)$ to be meaningful, the corresponding notion of objectual satisfaction must be meaningful; it must make sense to say of an object that it satisfies the condition $\Box (x > 7)$. It is then denied that there is any such notion of objectual satisfaction.

But at this point the two arguments diverge, with each providing a very different ground for the denial. For one, it is taken to be a general requirement on the notion of objectual satisfaction that the variables in a condition to which it is applied should be open to substitution; truth should be preserved upon the substitution of coreferential singular terms. But it is argued that, in the particular case at hand, this requirement is not met and that therefore there is no objectual notion of satisfaction.

According to the other, it is taken to be a consequence of there being an objectual notion of satisfaction for a condition such as $\Box (x > 7)$ that it makes sense to say of an object that it necessarily fulfil the corresponding non-modal condition $x > 7$.[1] But it is argued that this does not make sense— an object does not necessarily fulfil a condition in and of itself, but only relative to a description; and therefore, for this reason, there is no objectual notion of satisfaction.

We may call the problem raised by the two arguments, without discrimination between them, *the problem of* de re *modality*. The problem raised by the first argument alone may be called *the problem of quantifying in*, or, more specifically, *the problem of quantifying into modal contexts*; and the problem raised by the *distinctive* part of the second argument may be called *the problem of essentialism*. I do not know if this is how other philosophers

[1] I am only careful about use-mention conventions when it matters. I use 'condition' sloppily. Sometimes it means open sentence, as here; sometimes it means property; and sometimes it is ambiguous between the two. The context should decide.

use these labels; but it is the way I propose to use them here. The rationale for the present division is that the considerations raised by the common part of the two arguments belong most naturally to the first of them.

The difference between the arguments and the problems to which they give rise might be put in the following way. Both arguments constitute an attack on the notion of necessary satisfaction. But, in each case, this notion is understood in a different way. In the case of the first argument, it is understood to be a species of satisfaction; for an object necessarily to satisfy a certain condition $\phi(x)$ is for it to satisfy the corresponding necessity condition $\Box\phi(x)$. The argument is then an attack on the notion of necessary satisfaction as a *species* of satisfaction. In the other case, the notion is understood to be a mode, though not necessarily a species, of satisfaction; for an object necessarily to satisfy a certain condition is for it to satisfy that condition in a certain peculiar way. The argument is then an attack on the notion of necessary satisfaction as a *mode* of satisfaction.

Some further differences between the two arguments should be noted. The first belongs to the general area of the philosophy of language and relates, specifically, to the question of interpreting satisfaction and quantification. The second belongs to the general area of metaphysics and relates, specifically, to the question of how necessity can attach to objects. The arguments therefore raise very different issues and relate to very different areas of philosophy.

The logical argument is operator-indifferent. It applies equally well to any operator which, like the operator of necessity, creates opaque contexts, contexts containing terms not open to substitution. The metaphysical argument, on the other hand, is operator-specific. Substitute a different operator, such as that for belief, and one gets a different problem.

The focus of the logical argument is on the intelligibility of a certain kind of *expression*. The question is whether there is any meaningful use for free variables within the scope of a modal operator. The focus of the metaphysical argument is on the intelligibility of a certain kind of *idea*. The question is whether it makes sense to say of an object that it necessarily fulfils a condition. One might say that in the one case we are concerned with the intelligibility of a certain kind of expression without regard for what it might express, while in the other case we are concerned with a certain kind of idea without regard for how it might be expressed.

It is perfectly conceivable that one could find the idea of necessary fulfilment intelligible and yet, through acceptance of the logical argument, consider the modal conditions incapable of conveying that idea. Of course, finding the idea of necessary fulfilment *un*intelligible would provide a reason for taking the modal conditions also to be unintelligible. But still, the reason in this case, and in the case of the logical argument, would need to be distinguished. For there is a difference between saying that a form of

expression is unintelligible because it is incapable of expressing any idea and saying that it is unintelligible because there is no idea for it to express. In the one case, the fault lies in what one is attempting to say; while, in the other case, it lies in how one is attempting to say it.

Although the two arguments are distinct, they are in a certain way complementary. We can imagine an opponent of Quine attempting to vindicate the notion of *de re* modality in either of two ways. He can appeal either to the thought itself or the language by which it is expressed. He may say: 'Can't you just see that the notion of necessary satisfaction is intelligible?' Or he may say: 'Doesn't the possibility of quantifying into modal contexts guarantee the intelligibility of the notion?' Quine's two arguments can then be regarded as his response to each of these attempts to vindicate the notion.

What has made it so easy to confuse the two arguments is that the different considerations in either case can often be formulated in very similar terms. Quine's objection to quantified modal logic is often stated in the words: no object necessarily satisfies a condition independently of how it is described. But this may be interpreted either as a problem with the relation of satisfaction, in its application to necessity conditions, or as a problem with the relation of necessary fulfilment, in isolation from its connection with satisfaction.

It may not even help to cite the failure of substitution as the reason why there is no coherent notion of necessary satisfaction. For the failure may be a reason *in general* or a reason in *this special case* for rejecting the coherence of the notion. One may accept as a general principle that the application of the notion of objectual satisfaction requires that the variables in the conditions to which it is applied be open to substitution. Applying this general principle to the particular case of modal conditions, then gives the incoherence of the notion of necessary satisfaction.

On the other hand, one may accept in principle that objectual satisfaction can operate in the presence of substitution failure but have special reasons for thinking that it cannot so operate in this particular case. Presumably these reasons relate to an underlying form of scepticism over the notion of necessary fulfilment. If the application of this notion is relative to a description and if, moreover, the description is given by means of the very term used to refer to the object, then the failure of substitution would show that the relativity mattered, and that an objectual notion of necessary fulfilment was not to be had simply by dropping the reference to the description.

Again, both arguments may appear to be motivated by a common concern. For in both cases, a certain class of statements, the *de dicto* ones, are taken to be unproblematic, and another class of statements, the *de re* ones, are taken to be problematic. The question then is whether the problematic *de re* statements can be 'understood' in terms of the unproblematic *de dicto*

ones. It is this demand for a reduction that may make it plausible that there can be no notion of satisfaction for modal conditions in the presence of a failure of substitution, and that may likewise make it plausible that any notion of necessary fulfilment should be relative to a description.

But the mechanism of understanding is very different in the two cases. In the logical case, the mechanism is linguistic; it is the functioning of language that is meant to explain how the *de re* is understood in terms of the *de dicto*. In the metaphysical case, the mechanism is ontological; it is not language that carries the reduction, but the facts themselves (if it may be put this way).

The distinction between the *de re* and *de dicto* should also be construed differently in the two cases. For the purpose of the logical argument, it should be taken to be a syntactic distinction: the *de re* statements are those that contain a free variable within the scope of a modal operator. For the purposes of the metaphysical argument, the distinction should be drawn in semantic terms: the *de re* statements are those whose truth-conditions presuppose the intelligibility of the idea of necessary fulfilment.

The critical difference in the two characterizations arises when 'genuine' proper names are present in the language. Under the syntactic criterion, simple statements of the form $\Box Fa$, for a genuine proper name, will count as *de dicto*. But under the semantic criterion, such statements will count as *de re*, since presupposed in their truth-conditions is the intelligibility of the idea that an object necessarily fulfils the condition of Fx.

The two arguments are confused or, at least, not clearly distinguished in the earlier work of Quine. A crucial case is to be found in the second section of 'Reference and Modality' (Quine [1953]). He there writes as if the unintelligibility of quantification into modal contexts were merely another *symptom* of referential opacity. This would be compatible with the failure of substitution being the sole *reason* for finding the modal contexts opaque in the first place. But it seems that Quine wants to put forward another argument for opacity, one that relies on the incoherence of the notion of necessary fulfilment (p. 149).

Again, it is strange that Quine should think that the physical modalities should call for separate treatment (p. 158). For he has already come by the 'sweeping observation' (p. 159) according to which nontruth-functionality and substitution under logical equivalents imply opacity. Now presumably there is no doubt that the physical modalities satisfy the conditions of this observation. So I can only suppose that it is not the logical argument or its conclusion that is in question, but the quite different consideration concerning the notion of necessary fulfilment.

In the later work of Quine, the distinction between the two arguments is at least implicit. For his differential stand on modality and belief compelled him to recognize that there was a problem for the relational idiom of necessity that was not also a problem for belief.

The distinction has been emphasized by Kaplan in his two commentaries on Quine (Linsky [1971], 143, and Kaplan [1986], 232). But confusion persists. Philosophers are still prone to present one-sided refutations of Quine. So they cite the criticisms of Smullyan, on the one hand, or the criticisms of Kripke, on the other, without realizing that at best only one of Quine's arguments is thereby demolished.

Also, the relevant sense of the *de re/de dicto* distinction or of the mechanism of understanding is not always kept in mind. In Plantinga ([1974], 29–43) for example, we find that genuine proper names are used in the *de dicto* statements of a reduction that is directed against the metaphysical sceptic. Or, in Kaplan (Linsky [1971]), it is hard not to have the sense that the quest for linguistic understanding in the earlier sections (I–III) has been merged with the quest for ontological understanding in the later sections (IV–XII).

II

I want now to evaluate Quine's objections to quantified modal logic, dealing first with the metaphysical and then with the logical argument.

I observed before that the metaphysical argument was operator-specific; for different operators it yields different problems. This observation applies as much to different notions of necessity as it does to notions other than necessity. There is not a single problem of essentialism, but a range of problems, that vary according to the notion of necessity in question.

There are perhaps four principal notions of necessity for which the problem arises; these are, respectively, the logical, the analytic, the metaphysical, and the natural. Of these, the most important is undoubtedly the problem for the metaphysical notion. Indeed, not only is this problem of great importance in itself, but it is central, in my opinion, to any attempt to understand the nature of metaphysics.

However, it is not my intention to discuss this problem here. I wish to follow Quine and concentrate my attention on the logical and semantic modalities.

Our question then is this: can we make sense of what it is for a condition to be logically or to be analytically true of an object? Before attempting to answer the question, it is worth observing that it is one that naturally arises; it is not just forced upon us by the attempt to interpret quantified modal logic. For those very considerations that lead us to suppose that there is a special analytic or logical mode of truth may equally well lead us to suppose that there is a special analytic or logical mode of satisfaction. If a sentence can be *true* in virtue of its meaning or logical form, then why should a condition not be *true of* an object in virtue of its meaning or logical form?

Not only are the notions ones that naturally arise, they may also have some use. For in logic and semantics, the theses propounded include both closed and open sentences. If it is required that the closed theses should be logically or analytically true, why should it not also be required that the open theses be logically or analytically true of the objects with which they deal? Thus the notions of logical and analytic satisfaction may have the same role in characterizing the aim of these disciplines as the notions of logical and analytic truth.

Quine sometimes talks as if it is the treatment of modality as an operator rather than a predicate that leads us to take the plunge into modality *de re*; for one can only quantify into a sentence that is used, not into one that is mentioned (see Quine [1966], 170, 174). But the refusal to countenance any difference in the modal potential of the notions of truth and satisfaction provides another motivation for the plunge, one that in fact applies more naturally to the use of modality as a predicate than as an operator.

Let us now return to our question: does any sense attach to the notions of logical or analytic satisfaction? We may concentrate on the notion of logical satisfaction. For the most part, the considerations concerning the notion of analytic satisfaction will be similar; to the extent that they are not, they may be given separate attention.

If our question itself is to make any sense, then we must be able to say what the putative notion of logical satisfaction is and in a way, of course, that does not already presuppose its intelligibility. It seems to me that the most reasonable way of identifying the notion is as follows. We note that the notion of logical truth is obtained in a certain way from the ordinary notion of truth; there is, if you like, a certain operation that transforms the ordinary notion into the logical notion. The question then is whether this operation yields anything when applied to the notion of satisfaction; is there a notion of logical satisfaction that stands in the same relationship to the ordinary notion of satisfaction as the notion of logical truth does to the ordinary notion of truth?

If asked to explain in general what the operation was, in terms that did not arbitrarily restrict its application, then one would say something like this. Suppose a predicate P of n arguments is given. (Similar considerations apply to expressions of other categories.) Then the operation L (call it 'logicizing') delivers a new predicate $L(P)$ which applies to the n arguments $a_1, \ldots, a_n$ exclusively on the basis of their logical form.[2]

---

[2] The transformation L may have an interest that extends beyond its present use. We are all familiar with the problem of characterizing the logical constants. But there is also a problem of characterizing the metalogical terms; these are not the expressions that determine logical form, but the ones whose application is determined on the basis of logical form. In terms of L (as extended to expressions of arbitrary category), we may define a metalogical expression E as one for which it is true, or perhaps one should say analytically true, that E has the same extension as

It is as if the new predicate $L(P)$ operates through a 'veil of ignorance', a veil that is opaque to all but logical form.[3] Or one can imagine that there is a mechanism, blind to all but logical form, that surveys the arguments in turn. The predicate L(P) is then one that can only operate on the basis of the information supplied to it by the mechanism.

Our problem, therefore, is what sense, if any, attaches to saying: it can be determined that a condition Q is true of the objects $a_1, \ldots, a_n$ on the basis of the logical form of Q, $a_1, \ldots, a_n$ alone. Now when put in this way, a question immediately arises: do the objects $a_1, \ldots, a_n$ themselves contribute to the logical form? Can they be seen through the veil? Or, to put it another way, will our mechanism for discerning logical form even bother to look at them? (I have here presupposed that a uniform decision is made: either the mechanism always looks at the objects or it never does. This is a natural presupposition to make. In terms of the general transformation L, we may suppose that it is determined on the basis of the nature of the predicate P itself which of its arguments are to be looked at.)

Two aspects of logical form need to be distinguished. There is first of all the contribution made by the logical constants themselves. But there is also the contribution that is provided by the pattern of occurrence of the non-logical constants. This is something that can exist independently of the presence of the logical constants. So it is because of this other aspect that it is taken to be part of the logical form of the sentence 'Nixon admires Nixon' that there are two occurrences of the same singular term 'Nixon'.

The objects $a_1, \ldots, a_n$ will not normally make any contribution to the logical form in the first of these two ways. But it is perfectly conceivable that they might in the second way. Whether or not this is so will depend upon what exactly it is that has logical form. Suppose that logical form is taken, in the standard cases, to be primarily a feature of the physical expression-tokens. Then the pattern of occurrence of the non-logical constants would be given by the presence of suitable physical similarities among the sub-tokens, and it would presumably not be in the same way that one could talk of pattern of occurrence among the objects of an abstract sequence $a_1, \ldots, a_n$.

Now suppose, as is much more plausible, that logical form, in the stand-ard cases, is primarily a feature of expression-types or of intensions. Then one may still talk of different occurrences of the same sub-expression or of

---

L(E). The familiar metalogical predicates, such as those for logical consequence or logical inconsistency, would turn out to be metalogical on this definition. So also would familiar syntactic operations, such as those for negation or disjunction.

[3] My adoption of Rawls's phrase is not intended to be frivolous. There are in fact deep analogies between my use of the veil in characterizing logical notions and his use in character-izing the original position.

the same constituent. One may say, for example, that there are two occurrences of the sentence-type '$1 = 1$' in the complex sentence-type '$1 = 1 \supset 1 = 1$'. But occurrences in this sense are not tokens (i.e. physical tokens), for no token can occur in a type or intension. They are something as abstract as the types or intensions themselves.

It seems clear that, in this sense of 'occurrence', an object may equally well have different occurrences within a sequence. It would therefore be quite arbitrary, once the pattern of occurrence within an expression-type or intension was taken in, not also to take in the pattern of occurrence within a sequence. (Of course, what an actual physical mechanism would look at in such a case are not the occurrences themselves, which are abstract, but the representatives of them by tokens. It would no more 'see' the different occurrences of a symbol-type in an expression-type than it would 'see' the different occurrences of the philosopher Socrates in an abstract sequence of objects. But still, the processing of the tokens would merely constitute an indirect means of getting at the logical form of the types.)

It may be noted that one is almost compelled to take this view of logical form on a Russellian conception of propositions. For among the constituents of a proposition will be ordinary objects; and so part of the pattern of occurrence of the non-logical constituents will be the pattern of occurrence among the ordinary objects. It is then a short step from the pattern of occurrence of ordinary objects in a proposition to the pattern of occurrence of those objects in a sequence.

The position that allowed the objects to make a contribution to logical form would therefore appear to be the more plausible. But let us, in a spirit of neutrality, attempt to work out the conception of necessary fulfilment for either position. In case the mechanism is blind to the objects, this would appear to be straightforward. For upon conducting its search, our mechanism will only report back to 'control' on the logical form of the condition; and control is only able to say, on the basis of this information, that the condition is true of the given objects if it is able to say that it is true of those objects whatever they might be. It follows that a condition will be logically true of the given objects just in case its universal closure is a logical truth.

This approach has the interesting consequence that all identities are contingent. For regardless of whether the objects a and b are the same, the condition '$x = y$' will fail to be logically true of those objects, since '$\forall x \forall y (x = y)$' is not a logical truth. On the other hand, all self-identities will be necessary, at least if the identity symbol is a logical constant. For given that '$\forall x (x = x)$' is a logical truth, the condition '$x = x$' will be logically true of any object.

The approach therefore provides some sort of vindication of the contingency theory for identity claims. However, the vindication is not of the usual sort. It is contingently true of 9 and the number of planets that they are the

same, not because of some sensitivity to the way the objects are described, but because of a lack of sensitivity to what the objects are.

To the proposed account of logical satisfaction, it might be objected that it is unable to sustain a reasonable interpretation of quantified modal logic. For suppose that we take a condition such as $\Box\phi(x, y)$ to be satisfied by a pair a, b just in case the embedded condition $\phi(x, y)$ is logically satisfied by the pair a, b. Then $\Box x = x$ will be satisfied by the identity pair a, a while $\Box x = y$ will fail to be satisfied by the pair a, a; and so that familiar instance of substitutivity, $\forall x\forall y(x = y \supset (\Box x = x \supset \Box x = y))$, will turn out to be false.

But against this objection, two points need to be made. The first is that it is no requirement on an acceptable conception of logical satisfaction that it should be capable, in the suggested manner, of sustaining an interpretation of quantified modal logic. The other point is that the given conception is, in any case, capable of sustaining such an interpretation. As we shall see later in the chapter, by carefully distinguishing between the objectual and referential interpretation of the quantifiers, it is possible to show that the object-blind conception of modality is compatible with the objectual interpretation.

The object-sensitive conception of logical satisfaction is more problematic. For given that the mechanism for discerning logical form actually looks at the objects, we have to determine what it sees.

One question is whether it is aware of which of the objects are the same. Is it sensitive to relative identity? Perhaps the most reasonable view is that it is. For surely, when it is looking at an expression, i.e. a sequence of symbols, the mechanism will be aware of which of the symbols are the same; and it seems equally plausible that, when looking at a sequence of objects, it should be aware of which of the objects are the same. Identity of occurrence is a matter of logical form, regardless of which entities it is that occur.

On this view, the necessity of identity will hold, in contrast to the previous approach. For if the arguments a_1 and a_2 are the same, then the mechanism will report back to control that they are the same; and so it can then be determined on the basis of the logical form of the condition $x = y$, at least if identity is a logical constant, that the condition is true of the objects.

Even if identity is not a logical constant, a relevant difference between the two approaches will still emerge. For on the object-sensitive approach, a condition of the form '$Fx \supset Fy$' will be logically true of an identical pair of objects; while on the object-blind approach, it will not be, since '$\forall x\forall y(Fx \supset Fy)$' is not a logical truth.

A further question that arises is whether the mechanism, upon looking at the various objects, will be aware of which of them are distinct. Will it be sensitive to relative difference? It needs to be emphasized that this is a separate question from the one concerning relative identity. In the earlier case, we were asking whether the mechanism would provide *positive*

information concerning the identity of two of the objects a_i and a_j; in the present case, we are asking whether the mechanism will also provide *negative* information concerning identity. Of course, if it were known that the mechanism provided *all* of the positive information, then the negative information could be gathered from the absence of the corresponding positive information. But the mechanism could *in fact* provide all of the positive information without it being known that it did. It would then be left open which of the objects a_i and a_j about which no information was given were the same and which were distinct. They could all be distinct; they could all be the same, in which case there would only be one object; or some intermediate possibility could obtain.

But although the question is distinct, the answer should be the same; the mechanism should be sensitive to relative difference. For surely it should be sensitive to the occurrence of different symbols in an expression; and it seems equally plausible that it should be sensitive to the occurrence of different objects.

If this is so, then the non-identity condition '$\sim (x = y)$' will be logically true of distinct objects a_1 and a_2, at least if identity is a logical symbol; for the mechanism will report back to control that the objects are distinct, and then, on the basis of the logical form of the condition, it can be determined to be true of the objects. However, in contrast to the case of relative identity, the distinction between sensitivity and insensitivity to relative difference would not appear to show up when identity is not a logical constant or when identity or its cognates do not appear in the condition.[4]

It is a remarkable fact that, although we expect the mechanism to provide negative information concerning identity of occurrence, this information is not required to show that truth-functionally valid sentences are logically true. To take an example: suppose that χ is a sentence of the form $\phi \supset (\phi \vee \psi)$, where ϕ and ψ are distinct. Then to show that this sentence is true on the basis of its logical form, we need to know that the two occurrences of ϕ are the same, but we do not need to know that ψ is distinct from ϕ.

Whether the same point holds for predicate logic depends upon a subtle question concerning variables. The question might be stated in the paradoxical form: are variables constants? What I mean is: are variables to be given the same status as the logical constants or as the non-logical constants in the determination of logical form? Do we take such sentences as $\exists x(x = x)$ and $\exists y(y = y)$ to have a different logical form, because of the difference between

---

[4] In case $\phi(x, y)$ is an identity-free open sentence with two variables x and y, the point can be put in the following way: when $\forall x \forall y(x \neq y \supset \phi(x, y))$ is a logical truth then so is $\forall x \forall y \phi(x, y)$; but when $\forall x \forall y(x = y \supset \phi(x, y))$ is a logical truth, $\forall x \forall y \phi(x, y)$ may not be.

x and y, or do we take them to have the same logical form, in much the same way that '$o = o$' and '$1 = 1$' have the same logical form?

In the former case, there is no essential difficulty in extending our observation to predicate logic. But in the latter case, there is. For consider the sentence $\forall x \forall y (x = y) \supset \forall u \forall v (u = v)$. Then surely it is logically true; and yet in order to show that it is logically true we need to know that the variables x and y are, indeed, distinct.

If such negative information can be relevant, then the possibility is significant both for the characterization and for the representation of logical truth. According to the standard account, endorsed by Quine and many others, a logical truth is a sentence whose truth is preserved under arbitrary substitutions for its non-logical constants. But apply this account to the sentence $\forall x \forall y (x = y) \supset \forall u \forall v (u = v)$, and we find that it is not a logical truth; for upon substituting x for y, we obtain the sentence $\forall x \forall x (x = x) \supset \forall u \forall v (u = v)$, which is false in a domain of more than one individual.

It would be better, in the account of logical truth, to restrict the substitutions to those that preserve the exact logical form of the original. Indeed, so much should have already been clear from the intuitive conception of logical truth. For, according to this conception, a logical truth is a sentence that is true in virtue of its logical form. We can therefore expect sentences with *exactly* the same logical form as a logical truth to be true. But we can have no a priori guarantee that sentences with a different logical form, even when obtained by substitution from the original logical truth, should also be true.

These remarks are relevant to the schematic representation of logical form. Suppose we use 'α', 'β', 'γ',... as schematic letters for variables in much the same way that we use 'p', 'q', 'r',... as schematic letters for sentences. Then the logical form of $\forall x \forall y (x = y) \supset \forall u \forall v (u = v)$ can be represented by $\forall \alpha \forall \beta (\alpha = \beta) \supset \forall \gamma \forall \delta (\delta = \gamma)$. But we cannot take the validity of the scheme to consist in the logical truth of all its concrete instances; for upon substituting x, x, u, and v for α, β, γ, and δ, we obtain what we have already seen not to be a logical truth. The validity of a scheme should therefore be taken in general to consist in the logical truth of all its 'exact' instances; and the class of valid schemes will be closed, not under arbitrary substitutions, but only under those substitutions that preserve exact logical form.

Let me not be misunderstood. I am not actually endorsing a schematic approach to variables or arguing that negative information on logical form is ever relevant to the determination of logical truth. This is a matter for investigation. I am merely pointing out that the very definition of logical truth should not be taken to exclude the possibility. If the class of logical truths are closed under arbitrary substitutions, then this

should be treated as a theorem concerning the notion, not as part of its characterization.[5]

One final question remains concerning the concept of logical satisfaction. Do the objects $a_1, \ldots, a_n$ contribute anything else to logical form? When the mechanism looks at them, does it see anything else besides their relative identity and difference? The mechanism is only programmed, if I may put it that way, to detect logical form: and so the question turns, not on the external logical form of the sequence $a_1, \ldots, a_n$ itself, but on the internal logical form of the objects that go to make it up. As we have noted, if they are ordinary concrete objects, they will possess no logical structure and the mechanism will treat them in the same way that it treats the simple symbols of an expression. But the objects may themselves be expressions; and there then seems to be no reason why the mechanism should not be as sensitive to their structure as it is to the structure of the condition itself.

However, even in this latter case, the extra information will be of no use; if the mechanism reports it back to control, it might as well be discarded. For all that we have to go on is the logical form, not the meaning, of the condition; and there is nothing about the logical form of the objects that can help us decide whether a condition of given logical form is true of them. It is therefore, for all intents and purposes, as if no further information were available.

Putting together the three answers, we see that whether a condition is logically true of certain objects depends entirely on their identity type, i.e. on the relative identities and differences among them. Let $\pi(a_1, \ldots, a_n)$ be the statement that says which of $a_1, \ldots, a_n$ are the same and which are distinct. Then the condition $\phi(x_1, \ldots, x_n)$ will be logically true of $a_1, \ldots, a_n$ just in case $\forall x_1 \ldots \forall x_n(\pi(x_1, \ldots, x_n) \supset \phi(x_1, \ldots, x_n))$ is a logical truth.[6] The sensitivity to the objects shows up, if you like, in the presence of the restriction $\pi(x_1, \ldots, x_n)$.

This conclusion needs to be somewhat modified for the concept of analytic satisfaction. In the normal case in which the objects $a_1, \ldots, a_n$ are lacking in semantic structure, we may say, as before, that the condition $\phi(x_1, \ldots, x_n)$

---

[5] The issue of negative information is also relevant to the question of characterizing the metalogical terms. For is the negation of a metalogical predicate also metalogical? If there is unrestricted access to the negative information, then it will be. We may say, for example, that a sentence $\phi \lor \sim \psi$ is not a logical truth on the grounds that ϕ and ψ are distinct (and there is no other relevant logical form). However, without the unrestricted access, there is no guarantee that the negative predicate will also be metalogical. Perhaps there is some interesting way of mathematically exploiting the fact that, in many cases, the application of a metalogical predicate depends only upon positive information.

[6] Related principles are discussed in Fine [1978] in connection with the elimination of *de re* modality. Similar conceptions of modality have been propounded in Parsons (Linsky [1971], 85) and Kaplan ([1986], 250–1).

is analytically true of $a_1, \ldots, a_n$ just in case the sentence $\forall x_1, \ldots,$ $\forall x_n (\pi(x_{\pm 1}, \ldots, x_n) \supset \phi(x_1, \ldots, x_n))$ is an analytic truth. But in the case in which some or all of the objects $a_1, \ldots, a_n$ possess semantic structure, this result may not hold. For the 'analytic' control sees, not only the logical form, but also the meaning of the condition; and for this reason it is able, in contrast to the 'logic' control, to make use of information presented to it concerning the semantic structure of the objects. It is able to determine from the semantic structure of the sentence 'All bachelors are unmarried', for example, that the condition 'x is analytic' is true of it.

It is to be noted that, even in this case, it is not the meaning of the term used to denote the sentence that is relevant to the determination of analytic satisfaction, but the meaning of the sentence itself. It might plausibly be argued, however, that there is a canonical designation M of the meaning ψ of any sentence, identical or analogous to a Frege-Church structural description of its sense. Relative to such a system of canonical designators, it might then be maintained that a condition $\phi(x)$ is analytically true of ψ just in case the sentence '$\forall x(x$ means $M \supset \phi(x))$' is an analytic truth.

We therefore see that the notions of logical and analytic satisfaction do indeed make sense. Our understanding of logicality and analyticity combines with our understanding of satisfaction to produce intelligible hybrid notions. It might also be remarked that there appears in each case to be a natural reduction of the *de re* to the *de dicto*. This is in itself sufficient, though maybe not indispensable, for establishing intelligibility.

What might be conceded is that these notions are to some extent indeterminate (or perhaps one should say equivocal). There is nothing in our intuitive conception of logical or semantic structure, it might be argued, that will enable one to determine whether the resulting concepts of logical and analytic satisfaction should be object-blind or object-sensitive. But this concession itself is open to question, for it is not clear whether there is a difference of opinion over the application of the concepts or a genuine indeterminacy in the concepts themselves. There is perhaps a similar worry over the logical status of identity. Do conflicts indicate a difference of opinion or a difference in concepts?

But even if the concession is made, the indeterminacy should not be blamed on the notions of logical or analytic *satisfaction*, but on the notions of logical or semantic *structure*. Make it clear what is meant by 'logical' and 'analytic', i.e. by 'logical' and 'semantic' structure, and it then becomes clear what is meant by 'logical' and 'analytic' satisfaction. The difficulty is not over the intelligibility of these notions, but their identity.

In the light of these considerations, Quine's own remarks on the topic are curious. He claims that one can quantify into modal contexts (and here only the meaningfulness of the notion of necessary fulfilment is in question), but at a price. That price is the commitment to 'Aristotelian essentialism', by

which he means 'adopting an invidious attitude toward certain ways of uniquely specifying' an object, seeing these ways, rather than others, 'as somehow better revealing its "essence"' (Quine [1963], 155, and [1966], 173–4). Others have agreed with Quine on this, even if they have differed from him on the question of whether the price should be paid.

But it does not seem to be true that the price needs to be paid. Certainly, it is not as if the notions of logical or analytic satisfaction have to be *defined* or otherwise *explained* in terms of a class of preferred descriptions. Anyone looking for such a class in the case of our own account will look in vain. (Cf. Kaplan [1986], 252.) Even in the very special case in which a condition is analytically true of certain expressions and not others, such descriptions are not presupposed; the condition will simply be true of an expression in virtue of its semantic structure. There will indeed be a description of the expression that is especially revelatory of its semantic structure. But this no more constitutes a preferred description in this case than it does in the standard case of analytic truth.

Quine seems to assume that the explanation of *de re* necessity must take a certain form. With each object a is associated a class of 'preferred' conditions Γ in a single free variable, say x (and, similarly, with each n-tuple of objects $a_1, a_2, \ldots, a_n$ is associated a class of preferred conditions Γ in n variables, say $x_1, x_2, \ldots, x_n$). The object a is then said necessarily to satisfy the condition $\phi(x)$ (perhaps we should say $Ex \supset \phi(x)$) just in case the universal sentence $\forall x(\psi(x) \supset \phi(x))$, for $\psi(x)$ a conjunction of conditions from Γ, is necessarily true (and similarly for an n-tuple of objects $a_1, a_2, \ldots, a_n$).[7]

Such an analysis then invites the criticism that it calls for an invidious distinction to be made between those conditions in the relevant variables which belong to Γ and those which do not. Indeed, we seem to have a kind of proof that such a distinction among the descriptions of the objects is required. For not every condition in the relevant variable can belong to Γ on pain of the distinction between the necessary and the contingent collapsing; and some condition, at least in the case of two variables or more, must belong to Γ in order that $x = y$ should be necessarily true of a pair of identical objects a, a.

But from the fact that a class of preferred descriptions always exists, it does not follow that the analysis of *de re* necessity should be given in terms of such a class. Indeed, as we saw in the case of the notions of logical and analytic necessity, the most natural explanations take another form altogether (although whether this is also true of the notion of metaphysical

---

[7] Various subtle questions concerning (1) the presence of constants or modal operators in Γ, (2) the use of relational descriptions in Γ, and (3) the possibility of variant analyses have been ignored.

necessity is a much more difficult question). Even if we insist, somewhat perversely in my opinion, that the explanations take the required form, it will still not be true, at least for the object-blind notion of logical satisfaction, that an invidious distinction among descriptions must be made; for we can simply take the class of preferred conditions to be empty. There will remain a distinction of sorts; for we choose the class to be empty rather than 'universal' or something in between. Now this is a choice that can be motivated without any regard to the content of the conditions, for a universal class would collapse the modalities and an intermediate one would require an invidious discrimination *among* descriptions. Of course, the resulting notion of logical satisfaction will not make $x = y$ logically true of a pair of identical objects; but the desirability of such a requirement for the purposes of interpreting quantified modal logic is, as we shall see, more problematic than is commonly supposed.

A commitment to Aristotelian essentialism is therefore not necessary for making sense of the *de re* notions of logical and analytic necessity. But neither is it, in the intended sense, *sufficient* for making sense of these notions. For were we to characterize a notion of necessary fulfilment in terms of a non-trivial class of preferred descriptions, we would not thereby obtain a notion of *logical* or *analytic* fulfilment, but some other notion altogether. As Quine himself points out, 'essentialism is abruptly at variance with the idea...of explaining necessity by analyticity' (Quine [1963], 155). Yet our problem just was to make sense of a notion of analytic fulfilment.

But perhaps I am being unfair to Quine. It seems as if he is only interested in the notion of analytic truth, and not in other notions of necessary truth; and this suggests that he is only interested in the notion of analytic fulfilment, and not in other notions of necessary fulfilment. But it is conceivable that he only wishes the notion of necessity to be subject to the requirement that 'Necessarily ϕ' is true iff ϕ is analytic; and this is then compatible with interpreting the notion of necessary fulfilment in terms of privileged descriptions. But this way of viewing the problem is only really appropriate to the logical issue. The constraints that the requirement imposes on a solution to the metaphysical issue are so slight as to be almost worthless.

It appears that Quine and others have conflated two versions of the problem of essentialism: one for the analytic concept of necessity; and the other for the metaphysical concept. They have unwittingly assumed that any *de re* concept of necessity must be a concept of metaphysical necessity. What has perhaps made this conflation so easy is the confusion of subject with source; given that certain objects are the subjects of attributions of necessity, one naturally takes them to be their source. But, of course, a concept of necessity is none the less intelligible for having no interesting application to objects.

III

We proceed to an evaluation of the logical argument against quantifying in. This argument may be broken down into five steps. (1) It is argued that occurrences of singular terms within modal contexts are not open to substitution. (2) From this it is inferred that such occurrences are not purely referential. (3) From this it is inferred that the corresponding occurrences of variables are not purely referential. (4) From this it is inferred that the concept of objectual satisfaction is not meaningfully applicable to the conditions formed with the variables. (5) From this it is then inferred that quantification with respect to these variables is incoherent.

This breakdown of the argument may appear excessively minute. But I feel that each step raises significant and independent issues. My discussion of the argument, which will take us to the end of the chapter, will be in two major parts. In the first, I shall discuss each of the steps in turn, paying particular attention to the general theoretical issues that they raise. In the second, I shall be concerned to apply the results that will emerge from the first part to the specific question of whether quantification into modal and other problematic contexts is possible.

Step (1): The Failure of Substitutivity

Quine's standard example concerns the number of planets; it is necessary that 9 is greater than 7 and yet not necessary that the number of planets is greater than 7, even though 9 is in fact the number of planets. One small objection against this example is that it is needlessly controversial; for it presupposes the necessity of mathematical truths, and even their analyticity, given that necessity is explained in terms of analyticity. A better example, also used by Quine, concerns identity: it is necessary that $9 = 9$ and yet not necessary that the number of planets $= 9$. But even this example is open to the objection, in case the necessity is logical, that identity is not a logical constant and so '$9 = 9$' is not necessary. Perhaps the most satisfactory kind of example for avoiding needless controversy is something rather quaint. We may say: it is necessary that if 9 is odd then 9 is odd and yet not necessary that if 9 is odd then the number of planets is odd.

A more serious objection concerns not the status of the condition in the proposed example, but the status of the singular terms. It might be argued that no sentence containing singular terms can be an analytic, let alone a logical truth; for there is nothing about the meaning or logical form of a sentence that will guarantee that the terms have a reference; and it is only if the terms have a reference that a truth-value for the sentence itself is guaranteed. It is not my purpose to discuss this line of reasoning here; it raises large and difficult questions in the philosophy of logic and language.

Quine's own views, which would call for a very thorough discussion in themselves, are contained in his paper 'Meaning and Inference' [1963]. But we may note that if the reasoning is accepted, then no attempt to find a modal example of substitutivity failure can succeed.

It is hardly satisfactory, however, to fault Quine's argument on such a point, both because it is highly controversial and because the general issue of the transition from substitutivity failure remains. I shall therefore proceed on the assumption that there is indeed a failure of substitution in modal contexts.

Step (2): The Inference to Irreferentiality of Singular Terms

It might be thought that there is no difficulty here; for does not the irreferentiality of an occurrence of a singular term follow by definition from a failure of substitution? But matters are not so simple; and, in fact, our discussion of the step will raise some of the most significant issues that are involved in Quine's argument.

It has to be recognized that there are two concepts of referentiality; one informal and the other technical.

The informal concept of a purely referential occurrence of a term is explained by Quine ([1960], 177) in the following words: 'the term is used purely to specify its object, for the rest of the sentence to say something about'. So, within a sentence containing a purely referential occurrence of a singular term, there is a certain division of labour: the term picks out its object; and the rest of the sentence picks out what is said of the object. The sentence as a whole then says the one of the other.

Given this understanding of the concept, the occurrence of '9' in '9 > 7' would appear to be purely referential. For the sentence says of the number 9 that it is greater than 7, with '9' picking out its object, namely 9, and the rest of the sentence picking out what is said of the object, namely that it is greater than 7. On the other hand, the occurrence of 'Cicero' in '"Cicero" contains six letters' would appear not to be purely referential; for the sentence says nothing of Cicero at all. The occurrence of 'Giorgione' in 'Giorgione was so-called because of his size' has an intermediate status. The sentence does indeed say something of the object Giorgione, but the term 'Giorgione' is not used solely to pick out that object. We might say that the occurrence of the term is referential but not purely referential.

The explanation of the technical concept is rather different. An occurrence of a (referring) singular term in a sentence is purely referential in this sense if truth-value is preserved upon the substitution of coreferential singular terms. Let us be a little more exact. Suppose t is a singular term which occurs in the sentence $\phi(t)$; and let $\phi(s)$ be the sentence that results from $\phi(t)$ upon substituting s for the given occurrence of t. Then the occurrence of t is

said to be purely referential if $\phi(t)$ and $\phi(s)$ have the same truth-value whenever the identity sentence $t = s$ is true.

Quine is not careful to distinguish the informal and the technical concepts. We shall be. For the informal concept, we shall reserve the phrase 'purely referential'; though often we shall simply say 'referential' and use 'partly referential' in place of Quine's 'referential'. For the technical concept, we shall use our earlier phrase 'open to substitution'. Our question therefore is whether a failure in being open to substitution implies a failure of pure referentiality.

We could, of course, state the argument without resort to the informal concept of being purely referential; we could attempt to go directly from a breakdown in substitution to an incoherence in the application of the concept of objectual satisfaction. But such a version of the argument could only be as plausible as the one in which the intermediate step was supplied. For our most direct reason for holding objectual satisfaction to be incoherent is that the position of a variable can be occupied by a term that is not purely referential. If, through some freak, substitution could break down and yet the singular term still be purely referential, we would have no reason for supposing objectual satisfaction to be incoherent. There is therefore no advantage to be gained from attempting to avoid the use of the informal concept.

The argument only requires that openness to substitution be a necessary condition for pure referentiality. Quine believes that it is also a sufficient condition, and it is presumably for this reason that he is not careful to distinguish the two notions; one can simply be regarded as a definition of the other.

Considered as a definition, it is from the same mould as the definition of logical truth. In both cases, we have an informal concept that can be expressed in the form: x is F in virtue of x's being G. So the concept of logical truth is that of a sentence being true in virtue of its logical form; and the concept of pure referentiality is that of a sentence saying what it does in virtue of the reference of a singular term occurrence (and whatever else the rest of the sentence does). In both cases, the informal concept is given a definition of the form: feature F is preserved under all transformations that preserve G. So a sentence is said to be a logical truth if its truth is preserved under all transformations that preserve its logical form; and an occurrence of a term in a sentence is said to be purely referential if the truth-value of the sentence is preserved under all substitutions for the term that preserves its reference. (To preserve parity with the definition of logical truth, we should either have 'what the sentence says is preserved' or we should appropriately modify the informal concept. To avoid needless complication, we may leave the informal concept alone and we may take the definition to encapsulate the result that preservation of content is equivalent to preservation of truth-value.)

I am suspicious of all such attempted definitions. For a significant aspect of the informal concept is lost in the process. The informal concept is explanatory; it is required that the presence of the feature *G* actually explain the presence of the feature *F*. The defining concept is purely extensional; it is required that the feature *F* be preserved, but without regard for *why* the feature is preserved. It is plausible that if *G* explains *F* then *F* is preserved under transformations that preserve *G*; it is possible, though considerably less plausible, that if *F* is preserved under transformations that preserve *G*, then *G* explains *F*. But even if the implications hold, they should be regarded as consequences of a correct definition of the informal concept and not as constitutive of it.

Of course, general suspicion is no substitute for detailed investigation; and it may be worthwhile to consider further the sufficiency part of Quine's claim, even though it is not strictly relevant to the rest of his argument.

One difficulty in maintaining sufficiency is that there may be grammatical restrictions on which singular terms can be substituted for the given singular term. For all I know, there is no singular term of English which is distinct from and yet coreferential with 'nine' and which can be substituted for 'nine' in the sentence 'Fido is canine' without loss of grammaticality. Yet we would not want to say that, on this account, the given occurrence of 'nine' is purely referential.

A related difficulty is that there may be an impoverishment of terms in a given language. There is no single-letter word of English that is coreferential with 'nine'. Yet we would not want to say, on this account, that the occurrence of 'nine' in '"nine" consists of more than one letter of the alphabet' is referential. (A similar example is considered by Lewy [1976], 25.)

The first difficulty may be removed by requiring, of a term *t* that is to be open for substitution in a sentence $\phi(t)$, that not only should the sentence $\phi(s)$ have the same truth-value as $\phi(t)$ whenever *s* is coreferential with *t*, but also that $\phi(s)$ should *be* a sentence whenever *s* is coreferential with *t*. But this makes for difficulties with the necessity part of Quine's claim; and it does not, in any case, solve the second difficulty. It seems preferable to confine the claim of sufficiency to those languages in which there is no grammatical restriction on the substitution of singular terms and in which any expression whatever can be used to refer to a given object. In so far as a language is not of this sort, it must be suitably related to one that is.

There is a more serious difficulty. Take a sentence ϕ in which it is clear that an occurrence of a singular term is not purely referential. Perhaps 'Giorgione is so-called because of his size' will do; but if it will not, then another example may be chosen. Now form the disjunction of this sentence with '$2 + 2 = 4$'. In the Giorgione example, we obtain 'Giorgione is so-called because of his size or $2 + 2 = 4$'. Then in the resulting sentence, truth is preserved upon substitution of coreferential terms for the given

occurrence—indeed, upon substitution of arbitrary terms. Yet the given occurrence is still not purely referential.[8]

This should be intuitively clear in any particular case. It is intuitively clear, for example, that in the sentence 'Giorgione is so-called because of his size or $2 + 2 = 4$', the term 'Giorgione' is not being used solely to pick out an object—or, at least, this is as clear as it is in the original sentence 'Giorgione is so-called because of his size'. There is also an argument for this conclusion. For surely the referential status of a term is preserved under disjunction (and the other logical operations). Indeed, if we accepted openness to substitution as both a necessary and sufficient condition, we would be forced to say, in case ϕ contained an irreferential occurrence of a term, that that occurrence remained irreferential in the disjunction $\phi \vee \psi$ when ψ was false but became referential when ψ was true. But surely it is absurd that the referential status of the term should depend upon the non-linguistic facts in this way.

The general point is that truth-value may be preserved under substitution of coreferential terms for reasons having nothing to do with the referential status of the given occurrence of a term. The occurrence is, if you like, 'accidentally' open to substitution.

In this connection, Quine's own example (Linsky [1971], 141):

> 'Giorgione played chess' is true

is of interest. We cannot conclude, simply on the grounds that the occurrence of 'Giorgione' is open to substitution, that it is also referential. Whether it is will depend upon what account we give of the phrase 'is true'. If we regard it as a device of disquotation, analogous in its operation to the erasure or crossing out of the quotation marks, then the occurrence will be referential and, indeed, there will be no difficulty in quantifying into the quotation context, as with:

> $\exists x$ ('x played chess' is true)

If, on the other hand and as is much more plausible, we take 'is true' to be a predicate of sentences and take the quotation-mark expression to be referring, here as elsewhere, to the expression under the quotes, then the occurrence of 'Giorgione' will not be referential and quantification into the quotation context would appear to be impossible.

Presumably, what makes sufficiency so appealing is the thought that if the occurrences of terms t in $\phi(t)$ are generally open to substitution, then the context $\phi(-)$ can be construed as picking out a condition which the individual sentences $\phi(t)$ attribute to the object picked out by t. But although

---

[8] Linsky [1967], 100–4, has formulated somewhat similar counterexamples to the sufficiency of Quine's criterion.

it may be true that the context $\phi(-)$ can be so construed, the actual linguistic composition of the context may prevent it, as we have seen, from being so construed.

The question of sufficiency is not relevant to the *validity* of Quine's argument, but it is relevant to its *scope*. The argument is stated in terms of an example: he shows that substitutivity fails for a particular modal context; and then argues that quantification into that context is not possible. One naturally supposes that what goes for the one case goes for all cases and that there must be a general difficulty over quantifying into modal contexts.

But Quine's own principles rule out this extension of the argument. In the sentence '□ (9 is self-identical)', the occurrence of the term '9' is open to substitution. It therefore follows from sufficiency that the occurrence of the term is referential; and it then follows from his other principles that the context is accessible to quantification. So, by his own lights, Quine should be taken to have shown that quantification into modal contexts is sometimes impossible and that it is sometimes possible![9]

If one held the view, mentioned above, that no necessary truth can contain a singular term, then *all* of the singular terms within modal contexts would be open to substitution and so *all* the contexts would be accessible to quantification. Indeed, this result would hold good even if the necessity statements were written in explicitly quotational form. It would make sense to say '$\forall x$("$\phi(x)$)" is analytic)', with the quantifier actually binding the variable under the quotes.

All this is quite absurd and is merely another indication of the implausibility of the substitution test as a sufficient condition for referentiality. If sufficiency is rejected, then the proofs of possibility are blocked and one is thereby free to take a more sensible overall view of modal contexts. One can hold that all alike are resistant to quantification, even though there is only a partial breakdown, or no breakdown at all, in the substitution of singular terms.

Let us return to the question of necessity. But first, some terminology: in any alleged case of substitution failure, three sentences are used: $\phi(t)$, $t = s$, and $\phi(s)$. These involve two terms t and s and a context $\phi(\ \ )$. It is interesting to note that the terms and the context each occur twice in the sentences. We call t the *given* and s the *substituted* term. We talk similarly of the given and substituted *occurrences* of those terms. Sometimes we are sloppy, and talk of terms when we should mean occurrences. We call $\phi(t)$ the *given* and $\phi(s)$ the *resultant* or *final* sentence; $t = s$ is called the *identity sentence*. We shall sometimes think of the claim of necessity as providing a *test* for

---

[9] A similar point is made by Baker [1978]. He draws the conclusion 'so much the better for quantified modal logic'. I draw the conclusion 'so much the worse for Quine's account of referentiality'.

referentiality. If, in a particular case, the given and the identity sentences are true, then we may talk of a *positive* result from the test if the final sentence is true, and a *negative* result if it is false.

I shall now present a series of counterexamples against the substitution test. In the case of each counterexample, we shall show how the test might be appropriately modified. In this way, we hope eventually to arrive at a correct formulation.

I have tried to find my counterexamples from actually existing languages. But this is merely for dramatic effect. I could, with equal legitimacy, have drawn my examples from artificial languages constructed specifically for the purpose of faulting the test. For I take it that Quine would not want to restrict the scope of his criterion to the languages that happen to exist and would not consider it relevant, in assessing the correctness of the criterion, to engage in a detailed investigation of different languages.

The first counterexample turns on the syntactic status of the given and substituted terms. We note that the sentences 'Eve's elder son was Cain', and 'Eve is the mother of Cain' are true, while the sentence 'The mother of Cain's elder son was Cain' is false. So the result of the test is negative, and yet it is not to be doubted that the occurrence of 'Eve's elder son' in the first sentence is referential.

Such examples are not confined to natural language. An example from arithmetic (with the usual conventions governing the scope of '+' and '.') goes as follows: the sentences '2.2 = 4' and '2 = 1 + 1' are true; the sentence '1 + 1.2 = 4' is false; and yet the given occurrence of '2' in the first sentence is referential.

It seems clear that these counterexamples go against the intended application of the substitution test. What is wrong is that there has been a shift in syntactic function: an expression that originally performed the syntactic function of a singular term is replaced by an expression that no longer does so. A proper formulation of the test is therefore one in which the given and substituted singular terms are required to function *as* singular terms.

The difference in the two formulations of the thesis might be expressed as a difference in the sense of 'occurrence of a singular term'. A *typographic* occurrence has the same shape or appearance as the given singular term; a *syntactic* or *constituent* occurrence must also function as a singular term. On the original formulation, the occurrences of the given and substituted terms were merely taken to be typographic; on the present formulation, they are required to be syntactic.

The distinction between typographic and syntactic occurrence is related to Kaplan's distinction between accidental and vulgar occurrence ([1969], 112). Every accidental occurrence is typographic; 'nine' in 'canine', for example, is both. On the other hand, not every purely typographical

occurrence is accidental, at least if I correctly understand how Kaplan wishes to use his term. The occurrence of '1 + 1' in '1 + 1.2' is purely typographic. But it is not accidental; it is not, in the relevant sense, a mere accident that we have this combination of symbols. A change in the system of spelling is capable in principle of eliminating all accidental occurrences of singular terms or other meaningful expressions; but it is not in general capable of eliminating all purely typographic occurrences of such expressions.

Quine tends to ignore the distinction between the two kinds of occurrence. He seems to think simply in terms of typographical occurrence. But this has its dangers. As we have seen, it leaves his test open to counter-example. It also leads him to overlook a significant alternative test for irreferentiality; for any purely typographic occurrence of a singular term will, on that count alone, be irreferential. We know, from our discussion of sufficiency, that purely typographic occurrences of singular terms may still be open to substitution; and so we see that this test will genuinely extend the substitution test. A similar point holds in regard to the explanation, as opposed to the test, of irreferentiality. The fact that an occurrence is purely typographic is sufficient by itself to explain its irreferentiality. But in case the occurrence is syntactic, some other explanation must be sought.

There are other difficulties for the substitution test; for even when a substitution preserves the syntactic status of the singular term, it may still induce a syntactic shift in the rest of the sentence. We may suppose that the sentence 'The Smith family leap frogs' is true. The identity sentence 'The Smith family is the same as The Smiths' is presumably also true; and we may take it that the resultant sentence 'The Smiths leap frogs' is false. So the upshot of the test is negative; and yet no one would doubt that the original occurrence of 'The Smith family' was referential.

The shift in the syntax may sometimes be more subtle and may turn on the relationship of the syntax of the predicate-expression to the syntax of the subject-expression. For the purposes of the next example, we must imagine that three men are in a line, with Bill at the back and Fred at the front, and that Fred subsequently leaves. The sentence 'The man behind Fred saw him leave' is presumably true; the identity sentence 'The man behind Fred = The man before Bill' is also true; and the resultant sentence 'The man before Bill saw him leave' is false. Yet again, no one would doubt that the original occurrence of 'The man behind Fred' was referential.[10]

What has gone wrong in each of these examples is that the syntactic identity of the context has been altered. This is clear in the first two

---

[10] Linsky [1967], 104, has a similar example involving 'latter'. It should be clear from these examples that I go further than Kaplan in my criticisms of Quine's argument from substitution. For he concedes (Kaplan [1986], 235) that if substitution fails then one of the terms is irreferential.

examples. But also it is plausible in the last example. For in the sentence 'The man behind Fred saw him leave', it is part of the syntax of the context '—saw him leave' that 'him' stands in an anaphoric relationship to 'Fred'. Upon making the substitution to obtain the sentence 'The man before Bill saw him leave', 'him' comes to stand in an anaphoric relationship to 'Bill' and the syntactic identity of the context is thereby altered.

It should therefore be required, in a proper formulation of the substitution test, that the syntactic identity of the context remain the same in the given and resultant sentences. Given this requirement, the previous stipulation concerning singular terms then becomes redundant; for the syntactic identity of the context cannot stay the same without the syntactic status of the occurrences of the singular terms also staying the same.

The difference between the present and the previous formulations might be expressed as a difference in the sense of 'same context'. Again, the difference rests upon the distinction between the typographic and the syntactic. Typographic context is a matter of appearance; as long as the symbols remain the same, the context remains the same. Syntactic context is a matter of analysis; it is only when the syntactic analysis remains the same that the syntactic context can be said to remain the same. Before it was required that the substitution be made within a given typographic context; now it is required that the substitution be made within the same typographico-syntactic context.

Once the test is formulated in this way, it is seen to be entirely incidental that the typographic identity of the context remains the same; all that matters is its syntactic identity. A more general formulation of the test is therefore one in which the contexts are taken to be purely syntactic and in which it is allowed that different instances of the context may be given different typographic realizations. Indeed, so much is already apparent from the actual use we might make of the test; for it would be taken to be of no consequence, when performing a substitution, that the main verb was modified to agree with the subject or the initial letter of a singular term was capitalized in order to let it head a sentence.

It is hard not to think of these syntactic shifts in occurrence and context as somehow anomalous or irregular; the mechanism for discerning syntax from typography does not take the simplest possible form. We are therefore led to the idea of a syntactically uniform language. Within such a language, these anomalies do not arise: there is agreement both in syntactic and typographic context.

Let us be a little more precise. The requirement on occurrence is a global one. It says that, for any meaningful expression, all syntactic and typographic occurrences of that expression within a larger meaningful expression should coincide. The requirement on context is a local one and may be put in terms of substitution. Any occurrence of a meaningful expression

within a larger meaningful expression creates a context, which can be regarded either typographically or syntactically. In requiring these two contexts to agree, we are requiring that the substitutions within them should coincide: any result of making a typographic substitution of a meaningful expression within the typographic context should coincide with the result of making the corresponding syntactic substitution within the syntactic context.

Such a language is, in an obvious sense, completely perspicuous with regard to its syntax: the appearance of an expression is an infallible guide to its syntactic status; and the way it is put together is an infallible guide to its syntactic analysis. The syntax is discerned from the typography in the most straightforward possible way.

It seems reasonable to suppose that the different syntactic anomalies of a language can be ironed out and that it can be converted, upon a suitable adjustment in its notation, to a language that is syntactically uniform. In this way, the underlying syntactic regularities of the language are laid bare. Various familiar devices can be regarded as means to this end. Single-letter spelling, often introduced for purposes of notational economy, also serves to eliminate accidental occurrences of expressions; and bracketing, often introduced for purposes of disambiguation, also serves to eliminate irregular contexts.

Uniformity of language is of great relevance to the substitution test. The original test would appear to apply directly, without qualification, to syntactically uniform languages; for identity of typographic context automatically guarantees identity of syntactic context. But the test also applies *indirectly* to other languages; for we may first convert to a uniform language, and then apply the test to it. In this way, the question of whether the first occurrence of '2' is referential in the sentence '$2.2 = 4$' of an unbracketed irregular language is reduced to the question of whether the corresponding occurrence of '2' is referential in the corresponding sentence '$(2.2) = 4$' of a bracketed syntactically uniform language. Conversion also provides a test for purely typographic occurrence: for such occurrences disappear upon conversion; no occurrence in the converted expression corresponds to them.

It should be noted that these indirect tests require us to set up a suitable correspondence between the typographic occurrences of the given and the converted expression; for it is only in terms of such a correspondence that we can say 'what happens' to the given expression. What the correspondence in effect does is to isolate the syntactic (and subsequently, the semantic) contribution of a typographical item. It is not to be taken for granted that this can always be done. We can imagine a code for a language that does not work on compositional principles. Perhaps the sentences of English and of French are independently enumerated and then the nth French sentence is used as a code for the nth English sentence. Within the language of the code,

the sub-sentential components would not have isolable syntactic or semantic roles. It is against such an example as this that we can appreciate how remarkable it is that isolable roles can be attributed to individual components in the ordinary case.

The substitution test is beset not merely by syntactic but also by semantic ills; and again, both subject- and context-expression are equally prone to attack.

Let us deal first with the ills of the subject. Just as the syntactic status of the singular term may shift from the given to the final sentence, so may its reference, notwithstanding the truth of the identity sentence. One kind of case arises from ambiguity. The sentence 'Queen Elizabeth II weighs over a million tons' is true when the subject-phrase refers to the ship. The identity claim 'Queen Elizabeth II = The present queen of England' is true when the subject-phrase refers to the person. The first occurrence of 'Queen Elizabeth II' is clearly referential, and yet the sentence 'The present queen of England weighs over a million tons' is false when the reference of the subject-phrase is again to the person.

A similar kind of case arises from indexicality. At some time before a fateful hour in 1953 it would have been true to say 'The monarch lives' and 'The monarch is King George V'. But at any time after that fateful hour, it would not have been true to say 'King George V lives', even though the first occurrence of 'The monarch' is clearly referential.

I am not sure either of these counterexamples needs to be taken too seriously; I mention them more for the sake of completeness than for their intrinsic interest. They may be avoided by insisting that the test only apply to sentence-types, and not to sentence-tokens. This is not to say that the test has no application to the actual use of language; for instead of speaking of the truth-value of a sentence-type, we may speak of the truth-value of a sentence-type in a context that fixes the reference of an indexical or ambiguous term. It then suffices to require that the context remain the same from one sentence to the next.

However, there is a more devious kind of ambiguity, which does raise difficulties at the level of types. With our previous examples, a single occurrence of an expression within a larger expression-type had several 'meanings'. But it is also possible that different occurrences of the same expression should have different meanings, even though each individual occurrence has a single meaning. The first kind of ambiguity might be called *intracontextual*, and the second *extracontextual*. The second is more devious in being less apparent. The detection of intracontextual ambiguity is relatively straightforward and, in many cases, is simply a matter of consulting our intuitions; the detection of extracontextual ambiguity is much more problematic and often calls for a heavy exploitation of semantic theory.

The classic case of inter-contextual ambiguity is provided by Frege's account of oblique contexts. According to Frege, a term may either refer to its standard referent, or, in special contexts, it may refer to a non-standard referent. The consequence for the substitution test is that it may fail in its application to terms that occur within the special contexts, not because the terms are irreferential, but because the truth of the identity sentence fails to guarantee that the reference of the terms is the same in those contexts.

A somewhat less controversial case than Frege's is provided by the decimal system of notation. Perhaps the most plausible semantic account of this notation is this: a digit denotes according to its position; the dot helps to indicate position; and juxtaposition is used to signify addition.

Consider now the sentences '2.3 = 2.3', '3 = 03', and '2.03 = 2.3'. The first two are true and the third false; and yet it seems clear that the initial occurrence of '3' is referential. If the proposed semantic explanation of the notation is correct, this is because the given occurrence of '3' and the final occurrence of '03' are not coreferential within their respective contexts, notwithstanding the truth of the identity sentence.

These counterexamples raise a subtle issue concerning the intuitive concept of referentiality, one which we have not had to consider until now. Quine says that a referential occurrence of a singular term is used solely to specify its object. But what is 'its' object here? Is it the object that the term is used to refer to in the given context? Or is it the object that the term is standardly used to refer to?

It seems advisable to separate the question of whether a term is used to refer to some object or another for the rest of the sentence to say something about from the question of whether this object is the standard referent. Accordingly, we shall henceforth use 'referential' in the weaker non-committal sense and use 'standardly referential' for the stricter sense (cf. Kaplan [1969], 118). (Strictly speaking, it is not a question of whether the referent is standard but of whether the reference-relation is standard. It is possible to imagine that a term might non-standardly refer to its standard referent. It is perfectly conceivable, for example, that for Frege the referent and the sense of a term might coincide, as in 'the sense of this term'.)

If referentiality is standard, then the counterexamples do not work and no revision to the test needs to be made. Terms in Frege's special contexts and digits after a decimal point will fail to be referential, simply because they do not have their standard reference.

If referentiality is neutral, then the counterexamples stand and a revision to the test does indeed need to be made. A further restriction needs to be placed on the given and resultant occurrences of singular terms. It should not only be required that they have the same syntactic status as singular terms, but also that they should be coreferential.

This restriction differs in an essential way from our previous requirement that the identity sentence be true. For the requirement on the identity sentence concerned the reference of the terms in isolation from the context in which they occurred; while the present restriction concerns the reference of the terms *in situ*. Moreover, given this new condition, the old one should be dropped. For if the occurrences have their standard reference, then the condition on the identity sentence is idle; while if the occurrences fail to have their standard reference, the condition is irrelevant.

However, the new condition is still not enough to protect the substitution test against counterexample. For just as the substitution of a singular term can induce a shift in the syntactic identity of the context, it can also induce a shift in its semantic identity. In one kind of case, a lexical item within the context may change its meaning. An example is provided by languages that have a common notation for arithmetical and Boolean operations. (Some programming languages are like this.) We may suppose that the sign '−' may be used to signify either subtraction or complementation, according as to whether the expression following it is arithmetical or Boolean. We may also suppose that 't' and 'f' are Boolean expressions for the True and the False, that '1' and '0' are arithmetical expressions for 1 and 0, and that the True is identified with 1 and the False with 0. Then the sentence '$-t = f$' is true and the sentence '$-1 = f$' is false, even though the occurrences of 't' and '1' in their respective contexts are coreferential. In this case, we cannot even place the blame on the syntactic identity of the context, for it remains the same. What has gone wrong is that the substitution has induced a change in the meaning of the operation-symbol '−'.

Examples from natural language are hard to come by. An example from Hebrew (proposed by Ran Lahav) goes as follows. The word 'TSAFA' can either be the present, third-person, feminine form of a verb meaning to float or the past, third-person, masculine form of a verb meaning to observe. The word for the moon can either be 'YARE'ACH', which is in the masculine, or 'LEVANA', which is in the feminine. So we see that 'The moon (LEVANA) floats (TSAFA) in the sky (RAKEIA)' is true; and that 'The moon (YARE'ACH) observed (TSAFA) the sky (RAKEIA)' is false, even though the occurrences of 'LEVANA' and 'YARE'ACH' are coreferential. However, this is not a pure example, one in which there is only a semantic shift; for there is also a shift in the syntactic status of the verb 'TSAFA'.

In the case of this example, and others like it, there may be some doubt as to whether the given occurrence of the singular term is genuinely referential. For is the term 'YARE'ACH' being used solely to pick out an object? Is it not also being used to disambiguate the verb 'TSAFA'?

But referentiality is a matter of the direct role of the term in determining what is said. The given sentence says that the moon floats in the sky; and the immediate contribution of the term to this content is its object. This is not to

deny that the term may have an indirect role in determining what contribution other expressions make to the content. This is the case with our example, the gender of the subject-term determines the meaning of the verb; but this is no way impugns the referentiality of that term. The term, if you like, performs both off-stage and on-stage; and it is only the on-stage performance that is relevant to referentiality. (Later, we shall have further occasion to be highly discriminating over exactly what role of a term is relevant to its referentiality.)

Another kind of example arises when the substitution induces a shift, not in the meaning of any simple constituent, but in the semantic significance of the syntactic operations by means of which the simple constituents are combined. Again, ordinary arithmetical notation provides an example. The inequality '$02 > 1$' is true. But upon substituting '$(1 + 1)$' for '2', we obtain the false inequality '$0(1 + 1) > 1$'. Here the occurrences of '2' and '$(1 + 1)$' are coreferential, and even the syntactic identity of the context remains the same. What has happened is that the substitution has induced a change in the semantic significance of juxtaposition; from signifying addition (or perhaps something involving exponentiation), it has come to signify multiplication.

There may also be some doubt as to whether this example is pure; for does not juxtaposition correspond to two distinct syntactic operations in the context '02' and '$0(1 + 1)$'? More generally, it might be argued that a single syntactic operation should have a single semantic interpretation and that, where this requirement appears to be violated, the syntactic operation should be differentiated according to the different semantic interpretations which it can bear. This point of view raises difficult questions concerning the nature of syntax and its relationship to the semantics. Let me here record my own belief that the relationship between syntax and semantics should not be taken to be as tight as this point of view would have us suppose and that, just as we should be tolerant over a single lexical item bearing several different meanings, so we should be tolerant over a single syntactic construction bearing several different semantic interpretations.

To take care of these counterexamples, the requirement that the syntactic identity of the context remain the same must be strengthened. It must also be required that the semantic identity remain the same. In the simplest case (which, on a certain view, is the most general case), corresponding constituents should have the same meaning, and corresponding syntactic operations should have the same semantic significance.

Again, we may see the difference in the present and the previous formulations of the test as merely a difference in the sense of 'context'. Previously, we took a context to be syntactic; now we take it to be syntactico-semantic.

Considerations of uniformity in language may also be introduced in the same way as before. A uniform language is now one that is completely

perspicuous in regard to its syntax and semantics; there is complete coincidence in the application of the typographic, syntactic, or semantic concepts of occurrence and context. Syntax is determined from typography and semantics from syntax in the most straightforwardly possible way.

Two varieties of uniformity may again be discerned, one local and the other global. Local uniformity is typified by the occurrences of different expressions having the same syntactico-semantic function in a given context. Global uniformity is typified by the different occurrences of the same expression having the same syntactico-semantic function.

Again, it appears reasonable to suppose that the different semantic anomalies of a language can be ironed out and that it can be converted, upon a suitable adjustment in its notation and possibly its syntax as well, into a language that is both syntactically and semantically uniform. The original test will then apply, without qualification, to such languages; and it will apply indirectly to other languages by means of their conversion to such a language.

Philosophers have tended to be very suspicious of the idea of an 'ideal' or uniform language. But there is really no more to this idea than the attempt to be completely systematic about disambiguation. Instead of using one word with two different meanings, we use two words each with a single meaning. Apply this same procedure at the syntactic level and also to the means by which expressions are constructed, and we obtain the general conception of an ideal or uniform language.

A certain form of radical scepticism concerning the concept of uniformity and its application is indeed possible. Perhaps a radical scepticism over the purely syntactic aspect of uniformity could be maintained. But let us here give a partly semantic example (somewhat along the lines of Goodman's 'grue' and Kripke's 'quus'). We suppose that the word 'grun' is used ambiguously; it can mean either 'green' or 'not green'. In any context, the ambiguity is resolved by the number of letters in the subject-expression: if the number is even then the meaning is 'green'; and if the number is odd then the meaning is 'not green'. So sentences of the form '—is grun' will not be semantically uniform in the sense that I have previously tried to convey.

To this, the sceptic will object that one can, with equal justice, maintain that the ambiguity lies with 'green' rather than 'grun'. 'Grun' is unambiguous; but 'green' means 'grun' when the subject-expression is of even length, and it means 'not grun' when the subject-expression is of odd length. Of course, either hypothesis will require adjustments elsewhere in the semantic analysis of a sentence. But as long as this can be consistently executed, there is no reason to prefer one hypothesis to the other.

I no more have a telling objection to this form of scepticism than I do to any other. But I should like to point out that doubts over the existence of a uniform or ideal language would appear to belong more with these sceptical misgivings than with our natural good judgement.

Granted that there is a coherent concept of uniformity to be used, then our reformulation of the test strikes me as not being merely impervious to counterexample, but also susceptible to some kind of proof. We have that the syntactico-semantic contexts of $\phi(t)$ and $\phi(s)$ are the same. We know that the given occurrence of t is referential and that the substituted occurrence of s is coreferential with it; and we wish to deduce that the truth-values of $\phi(t)$ and $\phi(s)$ are the same.

What is lacking for a demonstration to go through is a precise account of what it is for an occurrence of a term to be referential. Quine's own account in terms of substitution was rejected long ago. But the problem remains; and a solution to it would not only be of great interest in itself, but would help us see what was correct in Quine's original account.

I would like to suggest that an occurrence of a term t in a sentence $\phi(t)$ is referential if there is a semantical analysis of the whole sentence in which the semantic value assigned to the occurrence of a term is its referent. We may imagine a roving semantic eye that picks out whatever is relevant to the given semantical analysis; when it comes to a referential occurrence, it picks out only the referent. It should be noted that this account actually supplied the explanatory factor that was found lacking in Quine's account. It shows the concept of referentiality to be theoretical in nature and to require, for its application, an implicit semantic analysis of the sentence under consideration.

If this is correct, we may assume that our given $\phi(t)$ has a semantical analysis in which the semantic value assigned to the given occurrence of t is its referent. This analysis determines an analysis of the context. This may be compared with the semantical analysis of the context which is meant to be preserved on substitution. If the two are in agreement, it would seem to follow, by some sort of compositionality principle, that the semantic values of $\phi(t)$ and $\phi(s)$ are the same; and from this and the principle that the semantic value of a sentence must always determine its truth-value, it follows that the truth-values of the sentences are the same. On the other hand, if the two analyses are not in agreement, it still seems reasonable to suppose that the analysis of the context will determine an analysis which *is* in agreement with the relevant part of the analysis of the sentence, either because this analysis exists at a lower level (at the level of reference, say, rather than sense) or because it is an alternative, but equivalent, analysis at the same level. The argument may then proceed in the same way as before.

I do not present this reasoning as an actual proof, but it does make clear how a proof might go.

There is a way in which Quine himself concedes the need for something like a uniformity requirement. In a very revealing passage from 'Reference and Modality' (Quine [1963], 150), he writes:

Nonsense is indeed mere absence of sense, and can always be remedied by arbitrarily assigning some sense. But the important point to observe is that granted an understanding of the modalities . . . , and given an understanding of quantification ordinarily so-called, we do not come out automatically with any meaning for such sentences as (30)–(31).

But what is lacking in Quine is any explicit acknowledgement of the significance of this requirement for his own argument.

In fact, the presence of the requirement makes an enormous difference to the epistemological status of the test. In unqualified form, the application of the test is completely unproblematic. That one sentence $\phi(s)$ is obtained by substitution from another $\phi(t)$ is a matter for 'inspection'; and that the respective sentences involved in the test—the given and final sentences $\phi(t)$ and $\phi(s)$, and the identity sentence $t = s$—have the appropriate truth-values is a matter for ordinary judgement. On the other hand, the application of the qualified test calls for what may be a highly theoretical judgement on the preservation of syntactico-semantic context. We are no longer making straightforward observations about the world, but problematic theoretical claims about the syntax and semantics of language.

It is as if Quine had attempted to state a theoretical truth within the confines of an observation language. In this respect, we might compare his formulation of the substitution test with the commonplace generalization 'Unsupported objects fall'. It may be granted that this generalization holds under conditions of *ceteris paribus*; and yet any proper account of these conditions will have to make use of theoretical terms.

Indeed, what is generally remarkable about Quine's discussion of referentiality is its atheoretical character. Although he was writing at a time when the views of Frege and Russell were already well known, there is no admission, at least in his earlier work, that they had any bearing on the issues at hand. Referentiality was to be decided by a simple 'observational' test; semantical analysis was, at worst, incoherent and, at best, irrelevant.

It might be wondered whether our own formulation of the substitution test not only makes it theoretical in nature, but also makes it useless as a way of testing for irreferentiality. For in any application it must be determined whether the syntactico-semantic context stays the same. And how can this be done unless it has already been determined whether or not the given occurrence of a singular term is referential?

However, the fact remains that we may be able to determine that two syntactico-semantic contexts are the same without knowing what that context is. Indeed, it suffices to determine that the contexts are the same under the hypothesis that the given occurrence of a singular term is referential. For if irreferentiality (something of the form $\neg p$) can be inferred from referentiality (something of the form p), it can be inferred without the benefit of that hypothesis. To take a concrete example, it may be unclear how the name

'Cicero' is functioning in the sentence ' "Cicero" has five letters'. But it is clear that the name 'Tully' in ' "Tully" has five letters' is functioning in the same way; and so, under the hypothesis that the original occurrence of 'Cicero' is referential, the two syntactico-semantic contexts will be the same.

If we are testing for standard referentiality, then identity of syntactico-semantic context is enough. We may therefore conclude from the failure of substitution in the example above that the name 'Cicero' in the original sentence is not standardly referential. However, if neutral referentiality is in question, then we also require that the given and resultant occurrences of the singular terms should be coreferential. But whether this is so, even under the hypothesis that the given occurrence is referential, will be highly questionable. For it will always be possible that the given and final terms refer, at their respective occurrences, to themselves; and so it will always be possible that they are not coreferential, except in the uninteresting case in which they are the same.

There would therefore appear to be a significant difference between the amended tests for standard and neutral referentiality: one is of some use; the other is not!

We may conclude this part of the discussion with the consideration of two topics with a more general bearing. As we have seen, the concept of uniformity is required for the proper formulation of the substitution test. This is no accident; it is also required, it seems to me, at many other places in the philosophy of language and logic. An interesting example for us, since it involves referentiality, comes from the attempt to characterize logical truth.

According to the standard account, a logical truth is a sentence whose truth is preserved upon arbitrary substitutions for its non-logical constants. But on this account (and, indeed, merely on the necessity half of it), the sentence

$$2 = (1 + 1) \supset (2.2 = 4 \supset (1 + 1).2 = 4)$$

is not a logical truth.

For upon substituting '1 + 1' for '(1 + 1)', we obtain the falsehood:

$$2 = 1 + 1 \supset (2.2 = 4 \supset 1 + 1.2 = 4).$$

What is one to make of this argument? One could take the brave step of denying that the original sentence was a logical truth. But although the sentence may not be a logical truth, it can hardly be because the given substitution results in a falsehood, as the proposed account of logical truth would seem to imply. Moreover, the bravery may require one to go too far. For the example generalizes to the extent that one would be forced, on this view, to deny that there were *any* logical truths that significantly involved singular terms.

Another response is to restrict the substitutions to those that preserve the referentiality of the singular terms. But this restriction appears to be completely unmotivated in terms of the original account of logical truth. Moreover, even if it could be motivated, it would still fail to deliver the right results; for referentiality will not guarantee that the syntactico-semantic context remains the same. So Quine ([1963], 146) is mistaken in thinking that the principle $\phi(t) \supset \exists x \phi(x)$ will hold when the term t refers and the given occurrences of t are referential (in either the intuitive or technical sense); for, in substituting x for t, the syntactico-semantic context may change.

In this regard, Quine's attitude to the substitutivity principle

$$t = s \supset (\phi(t) \supset \phi(s))$$

is of interest. He takes the failure of the principle merely to be evidence that the given occurrences of the term t are not referential ([1963], 140). But in so doing, he denies himself the possibility of making any significant generalization concerning the cases in which the principle holds; he can only say, the principle holds in the cases in which it holds. But it is hard not to believe that such a significant generalization exists, especially when the correct instances are regarded as logical truths.

Surely the proper response to these counterexamples is to require that the substitutions preserve the appropriate syntactico-semantic context. The counterexamples are then avoided, and in a properly motivated way. We may still abide by the usual schematic representation of logical truths, but we must take the substitution-instances, of which the schemes are representative, to be the result of a syntactico-semantic substitution, and not a purely typographic substitution.

The other general topic concerns the quest for rigour. Quine's discussion of referentiality is typical of much work in the philosophy of logic and language and even in linguistics; it involves very general considerations concerning the nature of language. How should such considerations be formalized? What is the proper framework of concepts and principles within which they should be set?

At the foundation for any such a framework will be a discipline of *universal abstract syntax*. This discipline attempts to formulate the general concepts and principles of syntax, the ones applicable to any possible language, and it attempts to formulate them in the most basic terms. The aim is to get at the idea of syntax as such.

My own conception of this discipline (which I hope to develop more fully elsewhere) differs in two fundamental ways from the orthodox conception, as found in the work of Montague [1970] and others. First the discipline is not regarded as an extension of the theory of concatenation. The basic entities of the discipline are not taken to be expressions or strings of symbols,

even in a suitably abstract sense. Instead, they are taken to be primitively given and are assumed only to have whatever structure their syntax endows upon them. It is not essential to the idea of syntax that the objects capable of possessing syntactic structure should be strings of symbols. Matrices, diagrams, bodily acts, propositions, even facts, may possess syntactic structure and yet not be, or relevantly taken to be, strings. Thus our discipline might equally well be regarded as the *general theory of constituent structure*.

Second, the basic syntactic structure of the entities is taken to be given, not by certain syntactic constructions, but by the operation of substitution. To be exact, there will be three basic syntactic notions: *occurrence of, occurrence in*, and *substitution. Occurrence of* is a two-place relation holding between an occurrence e and the entity of E of which it is an occurrence. *Occurrence in* is a two-place relation holding between an occurrence e and another occurrence f within which it occurs. (So we must therefore posit an ontology of occurrences of entities, in addition to the entities themselves.) *Substitution* is a three-place operation: if e is an occurrence and E' and F are entities, then substitution gives the result $F^{E'}/e$ (if any) of substituting E' for e in F. Our discipline must lay down the basic principles for these notions. One basic principle, for example, is that if F' is the result of substituting E' for the occurrence e of E within F, then there is an occurrence e' of E' within F' such that the result of substituting any expression E'' for e' within F' is identical to the result of substituting E'' directly for e in F.

I do not wish to dispute the existence of syntactic constructions. It is just that I do not take them as basic. Indeed, if c is any syntactic construction, of two places let us say, then c may be recovered, with the help of substitution, from any one of its instances. For let F be the result of applying c to E_1 and E_2, and let e_1 and e_2 be the corresponding occurrences of E_1 and E_2 in F. Then the result of applying c to any entities E_1' and E_2' is the same as substituting E_1' and E_2' for e_1 and e_2 in F. On the other hand, it is not always clear that the notions of occurrence and substitution are recoverable from the syntactic constructions. I therefore reverse the usual order of definition and take substitution as primitive, rather than the syntactic constructions themselves.

The syntactic structure of a language is given by a domain of entity-types, a domain of occurrences, the relations of occurrence in, and occurrence of, and the operation of substitution. Given such a structure, there is a natural notion of homomorphism. Roughly speaking, a homomorphism is a correspondence between the entities of the two domains that preserves the structure of substitution.

The various notions of interpretation, translation, and uniformity should all be explained in terms of the concept of homomorphism. Since we take an abstract view of syntax, the semantic values of an interpretation can themselves be regarded as constituting a syntax. An interpretation may then

be regarded as a homomorphism from the syntactic domain of expressions to the semantic domain of semantic values.

Suppose now that f_1 interprets the language L_1 in the semantic domain M_1, and f_2 interprets the language L_2 in the semantic domain M_2. Let g be a homomorphism from M_1 to M_2 which, intuitively speaking, maps each semantic value of M_1 into that aspect of it which the translation should preserve. (In the extreme case, g could be an identity function). Then a translation is a homomorphism h from L_1 to L_2 for which the diagram at Fig. 2.1 commutes. On this perspective, universal semantic becomes a part of universal syntax.

$$
\begin{array}{ccc}
M_1 & \xrightarrow{\ \ g\ \ } & M_2 \\[4pt]
f_1 \uparrow & & \uparrow f_2 \\[4pt]
L_1 & \xrightarrow{\ \ h\ \ } & L_2
\end{array}
$$

Finally, let us note that the expressions of a language will usually have a purely typographic structure, which can be specified in terms of suitable concepts of occurrence and substitution. The notions of syntactic and semantic perspicuity can then be defined in terms of the existence of appropriate homomorphisms between the typographic, syntactic, and semantic structures. Uniformity, or syntactico-semantic perspicuity, is usually regarded as an all or nothing matter. But it is in fact possible to define various intermediate concepts: the homomorphism may also be isomorphisms; they may or may not be onto; they may only have a local application; and so on.

Two interesting areas of investigation arise within this framework, although they also have an existence outside it. First, if one takes an abstract view of syntax or semantics then it becomes a definite question to what extent and in what manner a given syntax or semantics can be realized in a given concrete medium, such as the written or spoken word. We therefore have a study of what might be called *syntactic* or *semantic realizability*. Second, philosophers and logicians have concentrated their efforts on *uniform* languages. But most languages that are actually used are not uniform in nature and, indeed, many of their most desirable features would appear to derive from their lack of uniformity. We need to understand not only how a non-uniform language can be transformed into one that is uniform but also how it is that, in the process of transformation, certain desirable features of the original language may be lost.

Universal syntax also provides the natural basis for a completely general theory of abstraction and application. The operation of substitution can be regarded as the combined result of two other operations. Given the occurrence e of an entity E within F, first we remove e. This is the operation

of abstraction and it gives us an abstract or form λeF. (The notation should not confuse: e here is an occurrence, not a variable; it is abstracted from F and is no part of the resultant entity λeF.) We may then insert E' in the gap left by e. This is the operation of application and it gives us a new entity $\lambda eF(E')$. The fundamental law of abstraction then states that the combined effect $\lambda eF(E')$ of these two operations is the same as the result $F^{E'}/e$ of substituting E' for e in F.

Contexts are merely abstracts in the case in which the underlying entities are syntactic; properties (or relations) are merely abstracts in the case in which the underlying entities are semantic. Our previous considerations concerning syntactico-semantic context can therefore be formalized within the general theory of abstraction. It also seems to me, though this is not the place to develop the point, that any satisfactory account of the identity of propositions must be based upon the assumption that abstraction is one of the fundamental operations by means of which they are generated.

We turn to the remaining steps in Quine's argument. Fortunately, we shall not need to be so expansive.

Step (3): From the Irreferentiality of the Singular Term to the Irreferentiality of the Variable

Before the cogency of this step can be evaluated, it must be determined what the concept of referentiality for a variable is. Quine, so far as I know, provides no intuitive account of the concept. It is, however, fairly easy to construct one, by analogy with the account of referentiality for (closed) singular terms.

An occurrence of a singular term is referential if it is used solely to pick out its object. By analogy, we may say that an occurrence of a variable is referential if it is used solely to pick out its value. This account is all right as far as it goes, but we may make it a little more precise. First, we may note that 'value', unlike 'referent', is a relative term. We should therefore say that the role of a referential occurrence of a variable is relative to an assignment of values to the variables, and that, relative to such an assignment, the sole role of the occurrence is to pick out the value of the variable. Second, we may be more explicit about the role. A referential occurrence of a singular term is used to pick out an object for the rest of the sentence to say something about. In the same way, a referential occurrence of a variable is used to pick out a value for the rest of the open sentence to say something about. Finally, we may distinguish, as before, between neutral and standard referentiality. Just as a (closed) term has a standard referent, so a variable has a standardly assumed value, one that is given by an extracontextual specification of the domain. The variable will be standardly referential if it takes the standardly assumed value.

Our question therefore is whether the (standard) referentiality of an occurrence of a variable x in an open sentence $\phi(x)$ implies the (standard) referentiality of the corresponding occurrence of a term t in the sentence $\phi(t)$. We may concentrate on the neutral case of referentiality, since the standard case raises no special problems.

It should be apparent from our previous discussion of the substitution test that a battery of counterexamples might be marshalled against the inference from the referentiality of the variable to the referentiality of the term. Let me here mention just two. First, the occurrence of the variable 'x' in '$2.x > 0$' is referential and yet the corresponding occurrence of '$1 + 1$' in '$2.1 + 1 > 0$' is not. Indeed, it is not even a syntactic occurrence of a singular term. Second, using Quine's own notation for quasi-quotation, the metalinguistic variable 'α' has a referential occurrence in '$\ulcorner \alpha \urcorner$ is a term', but the corresponding occurrence of the term '0' in '$\ulcorner 0 \urcorner$ is a term' is not referential.

I would like to suggest that, for the inference to go through, it should again be required that the semantic or the syntactico-semantic context remain the same. Even if this requirement were not necessary for the validity of the inference, it might still be necessary for the purposes of the argument of which it is a part. For our aim is to show that, where there is a failure of substitution, there is a hindrance to quantification. But this conclusion could lose its interest if the context into which we could not quantify was not the same, semantically speaking, as the context for which substitution failed.

With the requirement, there appears to be some sort of possibility of proving that the inference goes through. For suppose that an occurrence of a variable x in $\phi(x)$ is referential, where we take this to mean that there is a semantical analysis of the open sentence $\phi(x)$ which is such that, relative to an assignment, the semantic value of the occurrence is the value assigned to the variable x. It would then seem to follow, if the semantic contexts remain the same, that the semantic value of t in the corresponding semantical analysis of $\phi(x)$ is its referent. In other words, uniformity would appear to dictate that the role of the value of a variable is the same, though relative to an assignment, as the role of the referent of a singular term, or, to put the matter in quasi-psychological terms, our understanding of the open sentence must be implicit, in the appropriate way, in our understanding of the corresponding closed sentence $\phi(t)$.[11]

Although it is not strictly relevant to the evaluation of the argument, it may be of interest to consider the other links that can hold among the

---

[11] We have here a response that might be made to Kaplan's alleged refutation of Quine's 'alleged theorem' (Kaplan [1986], section III). The refutation does not go through because it takes no account of the hidden premiss concerning uniformity. Kaplan himself responds to the gap he finds in Quine's argument by insisting on a requirement of coherence (p. 238). But it is not clear to me exactly what coherence is or whether it is intended, like uniformity, to constitute a repair to the argument.

referential status of terms. There is, first of all, the remaining cross-categorial link, going from the referentiality of the singular term to the referentiality of the variable. The ordinary symbolism of the predicate calculus suffices to yield a counterexample in this case. For '0' in '$\exists x(x > 0)$' is referential, while the last occurrence of 'x' in '$\exists x(x > x)$' is not referential.

There is then the remaining intracategorial link, going from the referentiality of one variable to the referentiality of another. Again, a counterexample from the predicate calculus may be given. The variable 'x' has a referential occurrence in '$\exists y(x > y)$', but the corresponding occurrence of 'y' in '$\exists y(y > y)$' is not referential.

It is usual to get around these difficulties by requiring that the occurrences of the substituted variables should be free, where the term 'free' is defined in purely typographic terms. But if what we are after is a general principle, then such a restriction is far too narrow and ad hoc. We might have different conventions for the position of the quantifiers; they might be placed at the end of an open sentence, as is common in mathematics, rather than at the front. There are also other ways in which a variable may become bound; the phrase 'the polynomial', for example, binds all the variables (though none of the parameters) within its scope. It is clear that there is no purely typographic concept that will cover all of these cases.

It again seems advisable to make the links dependent upon a requirement of a semantic or syntactico-semantic uniformity. A similar requirement, we may note, should be placed upon the substitution test that Quine provides for the referentiality of variables. If the given occurrence of 'x' in '$\phi(x)$' is referential, then the sentence '$\forall x \forall y(x = y \supset \phi(x) \supset \phi(y))$' should be true, but only on condition that the syntactico-semantic context is preserved.

Step (4): From the Irreferentiality of the Variable to the Non-Objectuality of Satisfaction

Like the second step, from the failure of substitution to irreferentiality, this step is unlikely to be challenged by the average reader of Quine. But it is again highly problematic, in my opinion, once the critical terms are made clear.

An occurrence of a variable is referential if it is used solely to specify its value. Whether an open sentence containing referential occurrences of variables is satisfied therefore depends only upon the values of those variables. On the other hand, the relation of satisfaction is objectual if it is a relation that holds between an open sentence and an assignment of values to variable.

It is presupposed, in the definition of referentiality, that the corresponding notion of satisfaction is objectual. But the notion of satisfaction may be objectual even though the occurrences of the free variables are not

referential; for there is nothing in the idea of satisfaction as a relation between an open sentence and an assignment of values to variables that makes it necessary that the identity of the variables should be of no relevance to whether the relation obtains.

One needs to distinguish between two ways in which an occurrence of a variable may fail to be referential. On the one hand, satisfaction may depend not merely on the value of the variable, but also upon how that value is specified. Call this *external dependency*. On the other hand, satisfaction may depend not merely on the value of variable, but also upon the identity of the variable itself. Call this *internal dependency*.

Referentiality excludes both forms of dependence. Objectuality of satisfaction excludes only the first. It is therefore possible, when only the first is excluded, that satisfaction should be objectual and the variables irreferential.

Quine seems to overlook this possibility in his argument against quantifying into modal contexts. He notes that '$\Box(9 > 7)$' is true but '$\Box$ (The number of planets > 7)' is false. He then supposes that whether an object satisfies '$\Box(x > 7)$' will depend upon the external means by which the object is given. But it is also possible that the satisfaction of the condition should depend upon the internal means by which the object is given, i.e. on the variable 'x' itself.

As far as I know, the possibility of distinguishing in this way between objectual and referential quantification has not previously been noticed. But the possibility should really have been already evident from the comparable case for singular terms. With what is taken to be a partly referential singular term, such as 'Giorgione' in 'Giorgione was so-called because of his size', the truth-value of the sentence may depend not only on the referent of the term but also on the term itself. In the same way, the satisfaction of an open sentence containing a free variable may depend not only on the value of the variable but on the variable itself. What goes for the singular term goes equally well for the variable.

The new style of variable, and the quantifier that goes with it, may be called *literalist*, since satisfaction will in general depend not only upon the object but also upon the letter. It is important to appreciate that the literalist quantifiers are not, like the substitutional quantifiers, let us say, an alternative to objectual quantification but are themselves a species of objectual quantification. The form of the satisfaction relation is standard; it is a relation that holds between an assignment of objects from the domain to the variables and an open sentence. The clause for the satisfaction of a quantified open sentence is also standard; an assignment satisfies an existential sentence $\exists x \phi(x)$, for example, just in case an appropriate variant of the assignment satisfies $\phi(x)$.

Indeed, there is something misleading about calling our variables or quantifiers literalist in the first place. We may take it that it is of the essence

of the use of variables that the satisfaction conditions for open sentences be given relative to an assignment of objects from the domain to the variables and that it is of the essence of the use of quantifiers that the satisfaction conditions be given by the standard clauses. So the distinction between literalist and referential quantification may be seen to lie not in the use of the variable or the quantifier but in the choice of the context: some contexts require us to look at the variable sign; others do not. The decision to use only referential variables or quantifiers amounts therefore to a restriction not on the use of the variables or quantifiers but on the contexts in which they may appear.

It seems to me that literalist variables are not a mere theoretical oddity, but may actually be closer to the use of variables outside logic than the more familiar referential variables. In most programming languages, for instance, we may make assignment statements in which variables occur both on the left and the right; a typical example is '$x := x + 1$'. The evaluation of the expression on the right depends only upon the 'value' of the variable; but the evaluation of the expression on the left depends upon the identity of the variable itself. Or, in ordinary mathematical discourse, we may say 'Let $y = x^2$. Then $dy/dx = 2x$'. Here again, it is the identity of the variables rather than their values that appears to be relevant to the proper evaluation of the term 'dy/dx'. (I do not intend these remarks as a full account of the use of variable signs in ordinary mathematics but merely as a step in the right direction. My own view, which I hope to develop elsewhere, is that variable signs are used in such examples as the above to signify variable objects.)

Literalist quantifiers may also be used to modify existing formal languages. We shall assume that the languages contain no individual constants, although our account could be tailored to their presence by allowing them, in the appropriate way, to be partly referential. One example, of immediate interest to us, concerns modality. Let us use the notion of necessity as object-blind in the sense previously explained, so an assignment of values to variables satisfies the open sentence $\Box \phi(x_1, \ldots, x_n)$ just in case $\Box \forall x_1 \ldots \forall x_n \phi(x_1, \ldots, x_n)$ is true. This gives us a perfectly acceptable relation of objectual satisfaction; but one that is literal, not referential; for $\Box x = x$ is satisfied by any assignment of an object to x, but $\Box x = y$ is not satisfied by the assignment of that object to both x and y. The use of literal variables therefore enables us to formulate an appropriate quantified modal logic for such a notion of necessity.

Another example concerns belief. We may wish to express that Ralph is in the situation typified by the case in which he believes that Cicero is Cicero but does not believe Cicero is Tully. It is tempting to suppose that the sentence $\exists x \exists y (x = y \& B_r x = x \& \sim B_r x = y)$ will do. But the logic of referential quantification will not permit it; for the sentence will have $\exists x (B_r x = x \& \sim B_r x = x)$ as a consequence. However, a literalist account

of the quantifiers allows us to succumb to temptation with logical impunity. For where $\phi(x_1, \ldots, x_n)$ is an open sentence in which $x_1, \ldots, x_n$ are all the free variables and are pairwise distinct, we may take $B_r\phi(x_1, \ldots, x_n)$ to be satisfied by the objects $a_1, \ldots, a_n$ just in case there are terms (perhaps appropriate terms) $t_1, \ldots, t_n$ for the objects such that $B_r\phi(t_1, \ldots, t_n)$ is true. On letting x and y both assume Cicero as values, we then see that our sentence is true. It is of course well known that the sentence can also be made true upon combining non-objectual quantification with a non-standard interpretation of identity. But our account is one in which the quantifiers are objectual and identity is standard. It should be noted, however, that the account renders $\exists x(B_rFx\&B_rGx)$ equivalent to $\exists x\exists y(x = y\&B_rFx\&B_rGy)$. It is the intra-belief, not the cross-belief, connections among terms that get expressed.

Literalist quantifiers may be used for the formalization of classical as well as intensional theories. One example, suggested to me by Allen Hazen, concerns a many-sorted theory in which the domains associated with the different sorts may overlap. Perhaps we have a sort for numbers and a sort for sets under a logicist construal of numbers as sets. Notwithstanding the overlap of domains, we may wish $x = n$ always to be false when the variables 'x' and 'n' are of different sorts. In such a way, we could permit within the metalanguage ('off-stage') a reduction of numbers to sets and yet not accept within the object-language ('on-stage') that any number was identical to a set. On a referential treatment of the variables such an account is impossible. But on a literalist treatment it offers no special difficulties; for we can take it to be part of the satisfaction conditions for the identity sentence $x = n$ that the variables 'x' and 'n' should be of the same sort.

Another example concerns the theory of truth. We may so use the truth-predicate T that the formula $T\alpha$ is satisfied by an assignment just in case (1) a formula ϕ of the object-language is assigned to the metalinguistic variable α, (2) individuals from the domain of the object-language are assigned to the free variables $x_1, \ldots, x_n$ of ϕ, and (3) the assignment of those individuals to $x_1, \ldots, x_n$ satisfies ϕ. On a referential treatment of variables, such an interpretation of the truth-predicate is incoherent since the satisfaction of the formula $T\alpha$ does not depend simply upon the formula assigned to α. But on a natural extension of the literalist treatment, under which *any* aspect of the assignment may be relevant to satisfaction, such an interpretation is unproblematic. With this use of the truth-predicate, the metalinguistic formula $T\alpha \supset \forall xT\beta$ is true when the object-language sentence $\forall xFx$, let us say, is assigned to α and the object-language formula Fx is assigned to β. For $T\alpha$ is true under this assignment just in case $\forall xFx$ is true; while $\forall xT\beta$ is true under the assignment just in case $T\beta$ is true whenever Fx is assigned to β and an arbitrary individual from the object-domain is assigned to x, and this holds, by the interpretation of T given above, just in case Fx is true under an assignment of an arbitrary individual to x. Indeed, if we so use metalinguistic

terms s that the sentence Ts is true under an assignment just in case $T\alpha$ is true under the variant assignment in which the denotation of s is assigned to α, then we may be so bold as to assert $T'\forall x Fx' \supset \forall x\, T\,'Fx'$, something which is usually regarded as a blatant case of use-mention confusion. By using such equivalences, we might even give a direct recursive definition of truth, one which made no appeal to an intermediary notion of satisfaction, although the approach would suffer from certain peculiarities of its own.

It is a characteristic feature of the use of literalist variables that the Leibnizian scheme $\forall x \forall y (x = y \supset (\phi(x) \supset \phi(y)))$ may fail. Our concept of necessity suffices to make the point. For let $\phi(x)$ be the formula $\Box(x = x)$, with the second 'x' designated. Then the resulting instance $\forall x \forall y (x = y \supset (\Box(x = x) \supset \Box(x = y))$ will be false, at least if there is more than one object in the domain; for, as we have seen, $\Box(x = x)$ will be true under the assignment of an object to x, while $\Box(x = y)$ will be false for the assignment of the same object to x and y. Again, it needs to be emphasized that this is not a failure of the sort familiar from the literature; for the quantifiers are objectual and identity is standard.

It is common to draw a distinction between the substitutivity principle of the above sort for objectual quantification and a substitutivity principle for terms: $s = t \supset (\phi(s) \supset \phi(t))$. It is thought that the second may fail, but that the first must hold. But we see that the one may fail for much the same reason as the other. Of course, this is not to deny that the quantificational principle is valid when the objectual quantifiers are required to be purely referential. But, in this respect, the two principles are on a par; for the principle for terms is equally well valid when the terms are required to be purely referential.

The deviation from the canons of pure referentiality is somewhat greater in the case of our truth theory than in the case of our modal or doxastic logics. For the two logics, it is only the relative identity of the variables that is of any account. As long as one formula is an alphabetic variant of another, its truth-value will remain the same. But for the truth theory, the absolute identity of the variable is also of account. Change a single variable, even when there are no others, and truth-value can change. So even though the statement $T\,'\forall x Fx' \supset \forall x T\,'Fx'$ is correct, its alphabetic variant $T\,'\forall x Fx' \supset \forall y T\,'Fx'$ is not. We therefore have a use for the distinction between relative and absolute identity at the level of the variables that is analogous to our previous use of the distinction at the level of the objects.

In addition to the particular object-theories, the metatheory of literalist quantification may be of some interest. We suppose that we are working with a standard first-order language. However, we do not follow the standard semantics in assigning an extension to each predicate and, on this basis, determining which assignments of values to the variables satisfy which atomic predicates. Instead, we simply stipulate which assignments are to

satisfy which atomic formulas, as with the truth-predicate *T*. On the standard semantics, it will follow from the clause for the satisfaction of atomic formulas that satisfaction in this case possesses certain normal properties. For example: it will be *local*—satisfaction will only depend upon the values of the variables occurring in the atomic formula; and it will be *referential*—change the variables, make a corresponding change in the values and satisfaction will be preserved. From the other clauses in the definition of satisfaction, it will then follow that these normal properties are preserved under the logical operations.

However, for the literalist semantics, there is no guarantee that even atomic satisfaction will conform to these properties and so there is no basis for an induction. For example, the satisfaction relation for atomic predications of truth will conform neither to locality nor to referentiality. Since these properties are required to validate the standard principles of quantificational logic, there is no guarantee that the literalist semantics will validate these principles.

Instead, we have a minimal logic that corresponds to the case in which no special assumptions are made about atomic satisfaction. We then have a hierarchy of stronger logics corresponding to the various special assumptions that might be made. However, the detailed investigation of these logics is not something which we shall pursue here.

Step (5): From the Non-Objectuality of Satisfaction to the Impossibility of Quantification

Various philosophers have attempted to evade Quine's conclusion by adopting a non-objectual account of quantification. They have supposed, for example, that the quantifiers are conceptual or substitutional. The question therefore arises as to whether the quantifiers must, in the intended sense, be objectual.

Given our own distinction between the objectual and the referential, it might be thought desirable to miss out the previous step of the argument altogether and to argue directly from the non-referentiality of the variables to the impossibility of quantification. The question would then be whether the quantifiers must be referential.

There does not appear to be any good reason for insisting upon the objectuality or referentiality of quantification. Quantifiers are not made in heaven, but on earth. Whether they are to be taken one way or another is a matter for stipulation, not discovery.

This is indeed how Quine sometimes regards the matter. Later, we shall give an objective meaning to the question of whether the quantifiers are to be taken to be referential for the purposes of Quine's argument. For the moment, we may note that there appears to be a sense in which the referential

quantifiers are basic and so fundamentally the only quantifiers that there are. Although the issue is not strictly relevant to the evaluation of Quine's argument, it may be of some interest to discuss it briefly here. I wish to claim not merely that the *truth-conditions* of the products of non-referential quantification can be explained in terms of statements involving only referential quantification. This is relatively unproblematic. I also wish to claim that our *understanding* of the one is to be explained in terms of our understanding of the other. If the statement ϕ is the product of a substitutional quantification, let us say, then there must be a statement ψ, formulated in terms of referential quantification alone, which is such that what we understand in understanding ϕ is to be given by what we understand in understanding ψ.

To this, it may be objected that the interplay between quantification and predication reveals a difference in our understanding. Let the product of the substitutional quantification be $\Sigma x Fx$. Let the corresponding statement in terms of referential quantification be $\exists x Fx$, where x is now the appropriate quantifier over terms and F is taken to be a predicate that is true of a term t just in case Ft (on our original understanding of F) is true. It may now be argued that, in our understanding of $\Sigma x Fx$, the meaning of F remains fixed; it has the same meaning in the quantified statement as it has in a simple unquantified statement Ft. However, in our understanding of $\exists x Fx$, the meaning of F has changed; it is now metalinguistic.

What is so hard to see, though, is that we have an understanding of $\Sigma x Fx$ that does not involve a shift in our understanding of F. I am clear in my own mind that I understand $\exists x Fx$, shift and all. But it is completely unclear to me that I have any independent understanding of $\Sigma x Fx$. There would appear to be no difference in our understanding of either sentence or perhaps one should say, more cautiously, that any difference would appear to be equally susceptible to explanation in terms of referential quantification.[12]

Similar remarks apply to the conceptual interpretation of the quantifiers, according to which the variables receive 'senses' rather than 'substituends' as values. There are other interpretations of the quantifier to be considered, which differ from the referential account not only in regard to the status of the values but also in regard to their singularity; the variables may be multivalued, and not merely mono-valued. Perhaps the most famous account of this sort is Carnap's, in which the variables receive both 'referents' and 'senses' as values. There are some treatments of quantification into belief contexts (Kaplan's [1971], 138, is an example) that call quite naturally,

---

[12] Van Inwagen [1981] has experienced similar difficulties in achieving a distinctive understanding of substitutional quantification. It should be noted, in contrast to what his paper suggests, that there is no need for the concept of truth to be involved in our understanding of substitutional quantification; it suffices if the meaning of each primitive predicate appropriately changes.

at the semantic level, for a distinction between two types of value. And even some cases of my own literalist quantifiers could be regarded as a kind of quantification in which both 'referents' and 'expressions' serve as values, but with the expression-value for each variable restricted to the variable itself.

But even quantification of this sort would appear to consist in the simultaneous or 'parallel' use of referential quantifiers. The point is, indeed, a general one. Any explanation of quantification will involve two elements. The first is a specification of the kinds of values that a variable can receive. The second is an account of the role that the different kinds of value play in interpreting an open sentence. By appropriately multiplying the referential quantifiers, and by appropriately tuning the interpretation of the predicates and other parts of speech to accord with the role of the different kinds of values, it would appear that any type of quantification could be seen to consist in the more or less devious application of referential quantification.

It may be wondered, given that there is a genuine indiscernibility in our understanding of the different kinds of quantification, why one rather than another should be taken as basic. I am inclined to think that the answer rests upon two desiderata of semantic explanation: one is the proper disassociation between the sign and what it signifies; the other is the analysis of any particular relationship of signification into its simplest elements. It seems plausible, once these desiderata are pursued to their limit, that only referential quantification should remain.

IV

Although the preceding critique has been largely negative, two positive conclusions do emerge. The first is that an occurrence of a singular term is irreferential if substitution fails under conditions of uniformity. The second is that referential quantification into a position occupied by an irreferential occurrence of a singular term is impossible, again under conditions of uniformity.

This means that an argument for the impossibility of quantifying in can be recovered from the ruins of our critique. For we may proceed in two steps. First, we may go from the truth of $s = t$, the difference in truth-value between $\phi(t)$ and $\phi(s)$, and the identity of the syntactico-semantic context across $\phi(t)$ and $\phi(s)$, to the irreferentiality of the given occurrence of t in $\phi(t)$ (where referentiality is here taken to be standard). Then we may go from the irreferentiality of t in $\phi(t)$ and the identity of the syntactico-semantic context across $\phi(t)$ and $\phi(x)$ to the unintelligibility of $\forall x \phi(x)$, for $\forall x$ a referential quantifier.

We are therefore not in the position of many of Quine's critics in denying altogether the validity of the considerations he puts forward. If we are to

dispute the conclusion of the reconstructed argument, we must dispute one of its premises—either by finding fault with the truth-values attributed to the statements $t = s$, $\phi(t)$ and $\phi(s)$ or, what is more likely, by finding fault with one of the assumptions of uniformity.

I want now to consider whether this reconstructed argument can be used to show that referential quantification into modal and other such contexts is impossible, dealing with each step of the argument in turn. My main concern is dialectical; I want to see to what extent Quine's argument can be used to establish conclusions concerning irreferentiality or the unintelligibility of quantification. But I also have an interest in the truth; I want to know whether given occurrences of term are irreferential and whether given attempts at quantification are unintelligible.

In considering various alleged examples of irreferential terms or unintelligible quantification, it will sometimes be helpful to be explicit about the languages from which they come. Three main sources may be distinguished. The first, which we call Modalese, is the language of quantified modal logic. It is a first-order language enriched with an operator whose interpretation is given in terms of necessity or belief or some other such intensional notion. The second is an ordinary language such as English, whose quantificational apparatus is given by the appropriate use of determiners and pronouns. The third is a cross between the first two. Ordinary language provides the basic sentences; and logic provides the quantificational apparatus. So in case the ordinary language is English, a typical sentence of the resulting hybrid 'Loglish' would be the sentence '$\exists x(x$ is a spy$)$'. The interpretation and, indeed, the specification of a language such as Loglish is not completely clear; and, to a large extent, the problems raised by Quine can be seen to be problems about just how the specification or interpretation should go.

The critical difference between these languages, for our purposes, lies in the degree of control we have over the interpretation of the referential and quantificational apparatus. For English we have no control; the interpretation is simply given to us. For Loglish we have control over the quantificational part, but not the referential part. And for Modalese we have control over both parts.

In the work of Quine, we may note a transition in interest from the more formal languages to the less formal. His initial preoccupation was with first-order modal logic; the question was whether appropriate sense could be made of its symbolism (see especially Quine [1947]). But his considerations soon became more general, and took in examples from both semi-formal and informal languages. In some of his work, he switches from one kind of example to another; and often it is not clear whether an example stated in Loglish is merely meant to be a proxy, for the symbol-wary reader, of a purely formal example. Probably, for Quine, the transition is of no great significance; but we, at any rate, will attempt to bear it in mind.

It will be illuminating in the first place to look at the occurrences of terms within unproblematic contexts, what Kaplan has called 'vulgar' occurrences (Kaplan (Linsky [1971]), 112). If we were going to be completely thorough, we would distinguish cases according as to whether the language was natural or artificial and according to whether the terms were names or description. But it will be sufficient, for our purposes, to concentrate on the case of descriptions from natural language. The main difference between natural and artificial languages here is that hypotheses concerning the use of terms in a natural language may become stipulations concerning their use in an artificial language; and the important differences between descriptions and names may be noted as we go along.

A typical example of a vulgar occurrence of a description is provided by the sentence 'The number of planets is greater than seven'. Is the given occurrence of the description 'The number of planets' referential? Quine would have no hesitation in answering 'Yes'. For truth is preserved under substitution and so it follows, from the very definition of referentiality, that the given occurrence of the description is referential.

But we have already had occasion to observe that the technical and informal concepts of referentiality are to be distinguished and referentiality in the technical sense, i.e. openness to substitution, is not a sufficient condition for referentiality in the informal sense. We therefore do not have this compelling reason for supposing that the occurrence of the description in question is referential.

We may consult the informal explanation of the intuitive concept. Is the description 'The number of planets' used solely to pick out an object? It might appear clear that it is. But here there is a difficulty. For what is the relevant sense of 'pick out'? There is a sense in which 'A man' in 'A man came to my room last night' may be said to pick out a man. But we would not want to say that, on that account, the given occurrence of 'A man' was referential, or even partly referential.

I have suggested that the relevant sense of 'pick out' is one that relates the role of the term to an account of what a sentence containing it says. The sole contribution of the referential occurrence of a term to the content of the sentence will be its object; the appropriate semantical analysis of the sentence will simply assign the object to the occurrence.

If this is correct, then whether the ordinary occurrences of descriptions are referential will be a theoretical question and will turn on what semantical account is provided of the sentences containing them. If one holds a Russellian view of these sentences, then definite descriptions will function in the same way as indefinite descriptions and it will be equally inappropriate, in either case, to say that these descriptions 'pick out' or 'contribute' an object to the content of a sentence. These terms will therefore not be referential. Although their occurrence will be open to substitution, this will be an

'accident' of a kind that was previously considered in connection with counterexamples to the sufficiency of the substitution test. Russell would hold a similar view of names if they were treated as disguised descriptions; but they would be referential if taken to be 'genuine' or 'logically proper'.

On a Fregean view, the matter is more problematic. Frege advocates, not a duality in the category of singular terms, but a duality in the conception of semantic analysis; for it may be conceived to proceed either in terms of sense or in terms of reference. At the level of sense, a description (or name) will not contribute 'its object', though at the level of a reference it will. However, it is the analysis at the level of sense that is most naturally taken to disclose what the sentence says. Indeed, it is somewhat odd to think of the rest of the sentence as saying something about the referent of a term that it contains. For unless the levels of analysis are to be crossed, the rest of the sentence will pick out something extensional like a set, rather than something intensional. So on the most plausible view of what a sentence says, its terms will fail to be referential, though on a less plausible view, they will be.

It is curious that the two leading views in the philosophy of language produce results that are not straightforwardly in conformity with our ordinary intuitions on the matter. For one is tempted to follow Quine and to suppose that a sentence containing a vulgar occurrence of a description is used to say something about its referent. Of course, it could be that our intuitions are in error. One would then have the familiar situation in which examples used to illustrate a theoretical concept turn out, upon proper consideration, not to illustrate it at all. But I am inclined to believe that the fault lies with the views and that a more satisfactory account of descriptions could be developed which would indeed provide for a level of analysis at which a description was capable of picking out an object for the rest of the sentence to say something intensional about.

Let us now look at terms that do occur within the problematic contexts. Quine presents a variety of examples, which differ widely in the degree to which they carry conviction. Perhaps the most convincing kind of example is provided by the occurrence of 'nine' in 'Fido is canine'. We have here no hesitation in declaring the occurrence of 'nine' to be irreferential. But the substitution test is of dubious relevance in reaching this conclusion. For we have a much more straightforward reason; the occurrence is not even a meaningful constituent of the sentence.

Another example is provided by quotation. The occurrence of 'Cicero' in '"Cicero" contains six letters' is meant to be irreferential. Here we do not have the previous compelling reason for supposing the occurrence to be irreferential; for it is entirely problematic whether the occurrence of 'Cicero' is a meaningful constituent of the sentence. But, as I have already indicated, the substitution test is of some use; for the appropriate assumption of

uniformity is relatively unproblematic. However, it needs to be emphasized that the test only entitles us to draw the conclusion that the occurrence of the term is not standardly referential. It is possible, and indeed plausible, that a term under quotes refers to itself (cf. Kaplan (Linsky [1971]), 120).

A less convincing example concerns the term 'Giorgione'. In the sentence 'Giorgione is so-called because of his size', this term is not open to substitution; for upon substitution of the coreferential term 'Barbarelli', a truth is converted into a falsehood. But does uniformity hold? Is the term referential?

Quine suggests that the term is partly referential; it is used to pick out its object, but it is not solely so used. He seems to assume that the term makes a double contribution to what the sentence says: it picks out its object; and it also picks out, or 'presents', itself. The rest of the sentence then says of the one that it is called the other because of its size. (We might note that Quine would here appear to be appealing to his intuitions on the matter. For what would a test for partial referentiality, analogous to the substitution test for pure referentiality, look like?)

But there is in fact a much more plausible account of how the sentence functions, one that would make the term referential and consequently lead one to reject the uniformity assumption upon which the application of the substitution test depends. On this alternative account, the term 'Giorgione' is used solely to pick out its referent. However, the expression 'so' in 'so-called' is used to refer to that term. So what the rest of the sentence says of the referent Giorgione is that it is called 'Giorgione' because of its size. The sentence attributes a property to the single thing picked out by the subject-term, not a relation to the two things picked out by the subject-term.

The way 'so' gets to refer to the term is analogous to the way that a demonstrative gets to refer. Indeed, the word 'so' can itself be used just like a demonstrative. We may say 'Giorgione is so-called because of his size', while pointing, at the moment of uttering 'so', to the expression that is to give it its reference. Imagine now that we point to the given occurrence of 'Giorgione' in the sentence and that the pointing becomes implicit. Then we get to something very close to the present use of 'so'.

There is nothing in the given example to favour the one analysis over the other. But consider the following conversation between two people *A* and *B*:

A: Giorgione is Italian.
B: Yes, and the man is so-called because of his size.

It is hard to make sense of the conversation on the analysis which requires that reference to 'Giorgione' must already have been made. For where does the reference come from? Surely not from the original occurrence of 'Giorgione', which appears to be straightforwardly referential, and surely not from the subsequent occurrence of 'the man'. On the other hand, the

conversation presents no special difficulties for the analysis which requires no antecedent reference to the term.

Of course there is a way, even on the second analysis, in which the term 'Giorgione' performs two roles: for it refers to Giorgione; and it secures a reference for 'so'. But we are not here interested in any role; that a term was used to shock would not make it irreferential. We are only interested in a certain kind of linguistic role. The question is: what does the occurrence of a term inject into the content of a sentence of which it is a part. It is therefore entirely irrelevant to the referential status of a term that is used as the referent of another expression; it plays this role not as an item of language but as an object of the world.

Quine's most contentious examples concern belief and modality. Since the two raise somewhat different issues, let us deal with each in turn. Consider the sentence 'Ralph believes that the man in the brown hat is a spy.' Is the occurrence of the description 'the man in a brown hat' referential or not?

This case is complicated by the fact that there are two readings of the sentence, what we may call the *de re* and the *de dicto* readings. Under the first, it is not necessary that Ralph think of the subject of his belief under the description of being the man in a brown hat; under the second, it is necessary.

In adverting to this difference, I am not attempting to provide any explanation of what it consists in. The suggestion that emerges from the work of Quine strikes me as implausible. He appears to attribute the difference to a lexical ambiguity in the term 'believes' (Quine [1960], 146–7, Linsky [1971], 103). But the possibility of a dual reading in this case is an instance of a much more general phenomenon, one in which any sentence containing an appropriately embedded description may be given a *de re* or a *de dicto* reading. Since there is no explanation of the general phenomenon in terms of lexical ambiguity, there is no reason to suppose that this is the explanation in the given case.[13] Russell's account in terms of scope avoids this particular difficulty. It also has the advantage of providing readings of intermediate scope. But it is subject to a difficulty of its own; for it can provide no plausible account of the dual reading of such sentences as 'The commissioner is looking for the chairman of the board.'

Given the existence of two readings, the question of referentiality should be considered separately for each. Under the *de dicto* reading, substitutivity may fail. It may be true that Ralph believes that the man in a brown hat is a spy and yet false that he believes that the mayor is a spy, even though the

---

[13] Other objections have been levelled against Quine's postulation of a lexical ambiguity; but I find none of them convincing. It should be noted that the present objection still holds if the lexical ambiguity is replaced with a structural ambiguity in the application of a single lexical item of belief.

man in a brown hat is the mayor. There is no reason here to suspect a shift in the syntactico-semantic context. We may therefore conclude that the given occurrence of the description is not (standardly) referential.

Under the *de re* reading, the given occurrence of the description will be open to substitution. But we cannot thereby conclude that it is referential. However, it does seem plausible that the present case should be assimilated to what we previously called the vulgar use of descriptions. So whatever goes for that case goes for this one too. According to our intuitions on the matter, the descriptions will be referential; on a Russellian view, they will not; and on a Fregean view, they will, though not straightforwardly so.

Similar considerations apply to ordinary names. In particular, there is the possibility of both a *de re* and a *de dicto* reading when names are used in place of descriptions. So on a *de re* reading of the sentences, the truth of 'Ralph believes that Tully is an orator' will follow from the truth of 'Ralph believes that Cicero is an orator'; while on a *de dicto* reading, the implication will not hold. (It would be good to have an account of the different readings that covered the cases of both names and descriptions. This is another reason for not liking the Russellian account in terms of scope, at least when it is combined with an account that makes the names of ordinary language 'genuine' or 'logically proper'.)

The final examples concern modality. A typical case is 'Necessarily, the number of planets is greater than seven.' Is the given occurrence of 'the number of planets' referential? Here it would be tempting to appeal to a difference between the *de re* and *de dicto* reading, just as in the example concerning belief. But Quine would presumably dispute the *de re* reading in this case.[14]

It is important, however, to be clear on the grounds upon which a *de re* reading might be disputed. There is no doubt that the words are capable of bearing a *de re* reading, as much in this case as in the case of belief. The only question is whether this reading is intelligible. The words are capable of carrying the thought, if only the thought is there to be carried. We are therefore back to our old metaphysical worry concerning *de re* modality, which at least in the case of the strict modalities may be allayed by the considerations in the earlier part of the chapter.

There might be thought to be a special reason why a *de re* reading is not available when necessity is interpreted as a strict modality. For on such an interpretation, the necessity operator should be subject to the principle that, for any sentence *S*, 'necessarily, *S*' is true iff *S* is necessarily true. But then the behaviour of the necessity predicate would seem to preclude a *de re* reading of the sentences containing the necessity operator.

---

[14] In his later work (Quine [1977]), he also disputes the intelligibility of a *de re* reading for the corresponding belief sentences. The two cases therefore become alike.

But once the possibility of a *de re* and *de dicto* reading has been conceded, the principle linking the operator and predicate for necessity should be formulated with more care. For, in case the sentence *S* contains descriptions (or perhaps names too), we need to ask whether the compound sentence 'Necessarily, *S*' is to be given a *de re* or a *de dicto* reading. It is then clear that the principle is only plausible for the *de dicto* reading; and so that still leaves open the possibility of a *de re* reading.

Indeed, it would be a mistake to think that necessity was any different from belief in this respect. For the concept of belief is subject to a link principle of its own, one roughly to the effect that, for any sentence *S*, the sentence '*P* believes *S*' is true iff *P* is prepared to assent to *S*. But the existence of a link in this case does not preclude the possibility of a *de re* reading of belief sentences; it only requires that the belief-sentences of the link itself be given a *de dicto* reading.

This completes our survey of the examples. Quine's own substitution test has been replaced by a more sophisticated test, one in which it has been required that there be syntactico-semantic uniformity of context. Using this alternative test, we have been able to find with Quine on some of his examples, and not on others.

In case a term is referential, no further question arises as to how it is being used; it is being used solely to pick out an object. But in case the term is irreferential, this further question does arise. We know the term is *not* being used solely to pick out an object. So how exactly *is* it being used?

Kaplan has noted a tendency on Quine's part to treat all such cases alike: the irreferential (in the technical sense) is assimilated to the accidental (Linsky [1971], 113). He himself has a tendency to assimilate the not standardly referential to the non-standardly referential (p. 119). But our brief survey of some examples would seem to indicate that there should be no general presumption in favour of one kind of explanation rather than another. The phenomenon of substitution failure is too diverse to admit of general explanation. Sometimes the occurrence of the term is not even a constituent; sometimes the context shifts; sometimes the reference of the term shifts. What is required in this area is not presumption but detailed investigation.

We turn to the question of whether quantification into modal and other such contexts is possible. This question divides up in two separate and independent ways. There is, first of all, the division according to whether the quantifiers are referential or not; and there is, secondly, the division according to whether the use of the quantifiable variables is or is not uniform with the use of singular terms. In this section, we take up the question for referential and non-referential quantifiers without assuming uniformity; and in the next section, we take up the same question, but under the assumption of uniformity.

The difference between assuming and not assuming uniformity might be put in the following way. Without uniformity, the sense of quantified statements is independent of the sense of their instances; we are therefore free to decide what sense to attach to them. With uniformity, the sense of the quantified statements is already determined by the sense of their instances; we therefore have no freedom to decide what sense they are to have. In the one case, the quantified statements are born with sense; and in the other case, they have it thrust upon them.

Quine is no enemy to the autonomous use of quantifiers; for his own account of quasi-quotation would appear to commit him to it. For in case no variables fall under the quasi-quotes, they are read in the same way as ordinary quotes. But quantification into ordinary quotation contexts is impossible; and so it is only by means of a special convention that any sense can be made of quantification into quasi-quotation contexts. There is also a way in which Quine allows for the autonomous use of quantifiers in connection with the problem of opacity. For he contemplates the possibility of quantifying into modal contexts in the case in which all singular terms have been excised from the primitive notation. The interpretation of the quantified statements is then unconstrained. But the reason is that there exist no instances by which their interpretation might be constrained. Our own perspective is somewhat different. There may be singular terms in the language; substitution may fail; the terms may even be irreferential. It is just that the behaviour of the singular terms is not considered relevant to the interpretation of quantification.

Quine is of the opinion that objectual quantification into modal contexts, even of an autonomous sort, is impossible without a commitment to Aristotelian essentialism; certain ways of describing an individual must be taken as preferable to others. It is here that our previous considerations on logical (or analytic) satisfaction may be brought into play; for we may transform our non-committal account of necessary satisfaction into a non-committal account of objectual quantification into necessity contexts.

It will be recalled that two concepts of logical satisfaction were distinguished. One was object-blind: the condition $\phi(x, y)$, say, was logically true of individuals a and b just in case $\forall x \forall y \phi(x, y)$ was logically true. The other was object-sensitive: the condition $\phi(x, y)$ was logically true of a and b just in case $\forall x \forall y (x \pi y \supset \phi(x, y))$ was logically true, where π was '=' or '$\neq$' according as to whether $a = b$ or $a \neq b$. Both concepts were seen, in their own way, to be free of essentialist presuppositions.

Let us now interpret the satisfaction of $\Box \phi(x_1, \ldots, x_n)$ in the following way: $\Box \phi(x_1, \ldots, x_n)$ is true of $a_1, \ldots, a_n$ iff $\phi(x_1, \ldots, x_n)$ is logically true of $a_{\pm 1}, \ldots, a_n$. If the concept of logical truth is taken to be object-sensitive, then we obtain a referential reading of the quantifiers; satisfaction depends only upon the values. If the concept of logical truth is taken to be object-blind,

then we obtain a 'literal' reading of the quantifiers; satisfaction depends only upon the *assignment* of values. In both cases, the interpretation of the quantifiers is objectual and yet innocent of commitment to essentialism.

The possibility of the first interpretation is not new; it is already implicit in the work of T. Parsons (see Linsky [1971], 73–87) and has been underscored by Kaplan and by myself in recent work (Kaplan [1986], Fine [1984]). But the possibility of the second interpretation would appear to be new. It depends critically on the distinction between objectual and referential quantification; for it is only in terms of that distinction that the usual constraints on objectual quantification can be avoided.

The difference between the two interpretations of the quantifier, though both are objectual, may still appear to be great. It is therefore of interest that there is an account of quantification in which the difference disappears. For we may so interpret the variables that distinct variables assume distinct values (the assignment of values to variables is one–one).

This convention was first proposed by Wittgenstein in the *Tractatus*. It has not turned out to be very useful for logical purposes, though it has been applied both by myself ([1978], 297–8) and by Hintikka [1956]. It is sometimes to be found in the writings of mathematicians who use '$\{x, y\}$', for example, to refer to a set whose members are distinct.

Under this convention, the literalist and referential readings of the quantifier become indiscernible. For the difference only shows up in the case in which the same individual is assigned to two distinct variables; and, under the convention, this can never happen. If, therefore, the convention had been adopted from the start, the distinctive problems involved in securing a referential reading would never have arisen!

Autonomous interpretations of quantification into belief and other such contexts can also be given. Indeed, Kaplan's 'trick', as propounded in 'Quantifying In', n. 3 is one such interpretation, of a very general sort. He treats the belief operator as a predicate that holds between the objects picked out by the free variables within its scope and the condition that they help define. Thus the sense of the operator varies with the disposition of the variables.

The possibility of autonomous interpretations of referential quantification is of great relevance to the application of Quine's argument to sentences containing *ordinary language* quantifier phrases. One kind of example is provided by the sentence 'Ralph believes someone is a spy' on a *de re* reading; another kind of example, to which similar considerations apply, is provided by the sentence 'There is someone who is such that Ralph believes that he is a spy.'

The question is: what can the argument from substitution tell us? Given a failure of substitution (using the terms 'the mayor' and 'the man in the

brown hat', let us say), does it follow that the quantifier sentences are unintelligible? The issue is complicated by the fact that the sentences to which the substitution test apply could be given a *de re* reading. The argument would not then get off the ground since the sentences would not even change their truth-value. So let us assume, in order to see the argument in its most favourable light, that the sentences are given a *de dicto* reading.

It might then be objected (indeed, often is) that the possibility of the *de re* reading for the test sentences shows that the quantified sentences must be intelligible. This may be so. But it still does not show what, if anything, is wrong with Quine's argument. We still need to know why, if at all, the unintelligibility of the quantified sentences does not follow from the failure of substitution for the test sentences under a *de dicto* reading. To put the point in a particularly graphic form, let us suppose that English never permitted a *de re* reading for sentences containing descriptions. There would then be no possibility of a counterargument to Quine. But the force of Quine's own argument would remain the same.

The weakness of the argument lies, of course, in its implicit assumption of uniformity. For it is always possible that the ordinary language quantifiers should be given an autonomous interpretation. The failure of substitution and, indeed, the general behaviour of singular terms would then be irrelevant to the issue of intelligibility.

If it could somehow be assumed that the quantifier constructions were uniform in their interpretation with the use of the terms under a *de dicto* reading, then the argument could get some grip. But it is hard to see how such an assumption might be justified. There is perhaps a general presumption in favour of uniformity. But uniformity with what? It is perfectly conceivable that the quantifier constructions might be uniform in their interpretation with the use of terms on a *de re* rather than a *de dicto* reading. Even if our language was one which, as envisaged above, permitted no *de re* reading, it would still seem more plausible, in the face of a conflict between the direct intuition of the intelligibility of the quantified sentences and a general presumption of uniformity, to give up the presumption, with its highly theoretical character, rather than the intuition. Thus, Quine's argument would still lack force against an opponent for whom there was intuitive evidence of intelligibility.

We have seen that there are no essential difficulties involved in giving an autonomous interpretation of quantification into modal contexts. We must now look at the question of whether a uniform interpretation is possible. In talking of uniformity here, I mean to indicate that the interpretation of a quantified statement such as $\exists x\phi(x)$ is to be uniform with that of its instances, $\phi(t)$; the syntactico-semantic context is to remain the same. The

interpretation of the quantified statement must, in this sense, already be implicit in that of its instances.

The second leg of our reconstruction of Quine's argument depends upon the assumption of uniformity. But what reason is there for supposing it to hold? The language of first-order modal logic, it should be appreciated, is one whose interpretation is up to use; its uniformity, or lack of it, is therefore a matter for stipulation. This is equally true of the formal parts of a semi-formal language such as Loglish, though perhaps less likely to be appreciated. It may of course be part of the conventions governing the interpretation of either language that $\forall x$ is to be a universal quantifier and $\exists x$ an existential quantifier. It may even be agreed that the quantifiers are to be referential. But that still leaves open what condition $\phi(x)$ in $\forall x\phi(x)$ or $\exists x\phi(x)$ is to express. The question of uniformity remains undecided.

Although we are always free to adopt an autonomous interpretation of the quantifier, the uniform interpretation is much more natural. It is the one which the quantifiers are naturally assumed to have in the absence of any special explanation. It is the only one to provide a general account of quantification that is of any interest. Moreover, part of the point in setting up an artificial language may be to expose uniformities; the expressions of the language are to bear their interpretation on their face. This purpose is hardly served by tolerating autonomous interpretations. Of course, Loglish will not, in any case, be fully uniform; but it may be seen as a step in the direction of a language that is.

All the same, there may be good reasons for tolerating autonomy. There are clear advantages, for example, in adopting the autonomous interpretation of the quantifier that is required by Quine's account of quasi-quotation. Another example will be given later. But even here, uniformity exerts its pull. For to a large extent, the advantages of these autonomous interpretations derive from the the fact that they simulate a uniform interpretation. Quine's account of quasi-quotation works *as if* Frege were right; it works as if terms under quotation marks did denote themselves and the quantifiers were accordingly interpreted. We may operate such languages and their logics under the illusion of a uniform understanding.

Whatever the reasons for adopting uniformity, its consequence, once adopted, would appear to be straightforward. Referential quantification, as in $\exists x\phi(x)$, is possible only if the occurrences of the singular term t in its instances $\phi(t)$ are referential. We may therefore refer to the results of our earlier section. Given that descriptions within belief or modal contexts are not referential under a *de dicto* reading, it would appear that Quine is vindicated; referential quantification into such contexts is not possible.

But matters are not so simple. For the interpretation of the substituted terms in $\phi(t)$ may not be uniform; and, as a consequence, the term t may be

referential in some of the instances $\phi(t)$ and not in others. This can happen in a variety of ways. A sentence containing a given term may be ambiguous, as with the *de re* and the *de dicto* reading of sentences containing embedded descriptions. There may be accidental variations within a given category of expressions, as with the occurrences of '3' and '1 + 1' in '3.2 = 6' and '1 + 1.2 = 6'. Or they may be systematic differences across categories. The most notable example is provided by Russell's theory of descriptions. On this view, there is a radical difference in the interpretation of descriptions and genuine names; and, as a consequence, the description will occur irreferentially, while the names will occur referentially.

In such cases, it seems unreasonable to require that the quantified statement should be interpreted uniformly with all of its instances. To take an extreme case, we would not want to deny a uniform interpretation to the quantifier in '$\exists x(x.2 = 4)$' on the grounds that the substituent term '1 + 1' occurred irreferentially in '$(1 + 1.2 = 4)$'. The general aim is that the interpretation of a quantified statement should be derivable from its instances. But then it suffices if there is a single instance from which the interpretation might be derived. (In the unlikely event that the interpretation of the quantified statement can be derived from instances which determine different conditions, the quantified statement will be ambiguous.)

The point might be put in terms of a distinction in the meaning of 'instance'. We may take a *substitution* instance of a quantified statement, say $\exists x\phi(x)$, to be the result of making an arbitrary meaningful substitution of a singular term t for x in $\phi(x)$: and we may take a *proper* instance to be a substitution instance which is uniform with the original quantified statement. The point then is that we require, not that a quantified statement be uniform with its substitution instances, but that it have a proper instance.

In the particular case in which the quantifier is to be referential, a uniform interpretation of a quantified statement will be possible just in case there is an instance in which the substituent term is referential. In demonstrating that referential quantification into a modal or other problematic context is impossible, therefore, it does not suffice to show, as Quine would appear to presuppose, that there is an instance in which the substituent term is irreferential. It must be shown that all instances are of this sort.

The idea that the Quinean difficulties might be evaded by discriminating among singular terms is hardly new; it is to be found in the earliest discussion of the problem. But it is important to be clear on what the proposal amounts to and what it will produce. Here the discrimination is on grounds of linguistic function. What we require are instances of the quantified sentence in which the given term is referential. In fact, one such instance will do. But if we have one, we are likely to have many; and if uniformity reigns among the resulting instances, then so will substitutivity. However, this is entirely incidental. Given one proper instance, what we get is a

uniform interpretation of quantification; for the one instance will determine a condition into which we can quantify.

Contrast this with the proposal that we should select a class of standard names, where being a standard name is a matter, not of linguistic function, but of having the right 'core content' for the type of intensional context in question. It is now essential that substitutivity should hold when restricted to the occurrences of the standard names in the relevant contexts. It is also essential that there should be several standard names, indeed one for each object of the domain. However, there is no need for the terms to be referential in the chosen contexts; and given that they are selected on the basis of content, they are unlikely to be so. What we then obtain with such a class of terms is an *autonomous* interpretation of quantification into the chosen contexts: satisfaction is given in terms of the truth of the instances formed with terms from the class; and quantification is explained in terms of satisfaction.

What has perhaps made it so hard to keep the two proposals apart is that the standard terms under the autonomous interpretation behave, in regard to their substitutivity properties, as if they were referential terms in a uniform context. But whatever the reasons, there has certainly been a tendency to confuse the two forms of the proposal. They are lumped together without regard for either the difference in their requirements or the difference in their results. One finds, for example, this statement in a relatively late work of Quine ([1976], 862): 'Instead of bandying a uniquely fulfilled predicate "G", one may forge a corresponding singular term "g". [This presumably would be the description $\iota x G x$.] Here, then, is what Føllesdal called a genuine name, and what Kripke has lately called, more quotably, a "rigid designator"'.

Our own concern at present is with the logical issue. Whether there will indeed be proper instances of the required sort will depend upon what account is given of the referential status of singular terms in the corresponding contexts. If one adopts a Fregean view, then all the terms in the contexts will fail to be (standardly) referential; and (standardly) referential quantification into such contexts will therefore not be possible. Quite apart from a specifically Fregean view, it follows that if one insists upon the uniform use of all terms, constant or variable, within a modal or similar context and if, also, one permits the use within that context of terms, characteristically descriptions with a *de dicto* reading, for which substitution fails, then standardly referential quantification into the context will not be possible. It is this, it seems to me, that is the main negative conclusion to emerge from the work of Quine. The consequences for quantified modal logic are not indeed damning, but a significant restriction is imposed on the behaviour of terms in languages for which uniformity and standard referentiality are desiderata.

It is possible to conceive of a Fregean view, though not of an orthodox sort, that provides for both the *de re* and the *de dicto* readings of descriptions within the problematic contexts (intermediate readings would be somewhat harder to obtain). Descriptions under a *de dicto* reading would refer to their sense, as on the orthodox view. Descriptions under a *de re* reading would refer to their ordinary referent. They would therefore be standardly referential in character. This is, on its own account, a plausible view and, if adopted, would then provide the basis for a uniform interpretation of the standardly referential quantifiers.

On a Russellian view, even the descriptions on a *de re* reading will not be referential and so will not secure a uniform interpretation of the quantifier. This result may appear surprising; for surely the interpretation of the quantified statement is already implicit in what for the Russellian is the wide scope reading of the description. It is, but not in the required way. We have here the uniformity of quantifier to quantifier, not of quantifier to term.

To obtain the requisite uniformity, it must be supposed that the language contains genuine names. The proper instances of the quantified statement will then be obtained, if they exist at all, by substituting the genuine names for the variables. If the language contains no genuine names, then they must be added. But the question then arises as to how the resulting sentences are to be interpreted.

It seems reasonable, especially if uniformity is a consideration, that they should be interpreted in conformity with the *de re* reading of the corresponding sentences with descriptions. We have something like the uniformity requirement in reverse: the interpretation of a quantified sentence, though of a non-orthodox sort, determines the interpretation of its proper instances. These instances then determine the interpretation of the quantified sentences of a more orthodox sort; and so again, the possibility of quantification turns ultimately on the intelligibility of the *de re* readings of the sentences with descriptions.

For the case of the strict modalities, the Russellian will have a natural understanding of the sentences which result from applying the necessity operator to sentences that contain genuine names. For the component sentence will express a singular proposition; and so the compound sentence will say of the singular proposition that it is logically or analytically necessary in the sense that has already been explained. There will not be the usual problems over referentiality, since the objects in the proposition will contribute as much to its logical form as its more intensional constituents.

This understanding of the necessity sentences then automatically yields an interpretation of quantification into necessity contexts. So we have an interpretation of the language of quantified modal logic that is completely uniform, that makes the quantifiers referential, and that construes the

modalities as strict. Moreover, this interpretation is one that the Russellian naturally arrives at; it is not one that he need contrive.

But a warning is in order. There is nothing in a language or the set of its truths that enables one to determine whether it is uniform. Uniformity is not a surface phenomenon, but depends upon the underlying semantical analysis. We could imagine a Fregean appropriating the language of quantified modal logic for his own use. He would take closed terms within modal contexts to refer to their standard sense; and he might dispose of free variables within modal contexts by means of Kaplan's 'trick'. In this way, he could simulate the effect of the uniform semantics of the Russellian. But his own semantics would be non-uniform.[15]

The Russellian way out is the best known of the responses to Quine's argument. But it is important to appreciate on what its efficacy rests. It does not simply rest, as we have seen, on the *selection* of terms with desirable substitutivity properties. Nor does it clearly rest, as Quine and others seem to assume, on the *elimination* of terms with undesirable substitutivity properties. For how does that help? How can a sentence be rendered intelligible by the removal of certain terms from the language to which it belongs? Of course, the terms are eliminated in favour of a paraphrase. But the redundancy of the removed terms can hardly make the problematic sentence any the more intelligible. Of course, Quine's argument can no longer be stated once the descriptions are removed from the language. But one does not defeat an argument by refusing to allow one's opponent to use the terms with which it must be stated.

The real efficacy of the Russellian response lies in its differential stand on genuine names and descriptions. Because of the disparity in their semantics, a genuine name can be referential where a description would not be; uniformity can thereby be saved. The relationship of elimination to the differential stand is somewhat problematic. Certainly, elimination is not necessary for the differential stand. The essence of the Russellian position is that descriptions are a kind of quantifier phrase. Whether they can be eliminated in favour of other quantifier phrases is a separate and somewhat incidental matter.[16] Elimination may be sufficient, however, for a differential stand; it all depends upon exactly what it is meant to do. If it is intended as a semantical analysis, then it could indeed provide the basis for a differential stand. But if it is intended as something else, as seems to be the case with

---

[15] Kaplan [1975] has argued that the Russellian can accommodate the semantical ideas of the Fregean. But the Fregean can also accommodate the semantical ideas of the Russellian. For example, he may treat a singular proposition, along the lines of the 'trick', as an ordered pair of an object and a property. Given the possibility of interpreting either theory within the other, there would appear to be a deep sceptical problem as to whether there was a genuine difference between the two theories.

[16] Kaplan ([1972], 214) has made a similar point, though in a somewhat different connection.

Quine's own elimination of singular terms, then it is hard to see in what its relevance could consist.

We have so far discussed the question of referential quantification under conditions of uniformity. If the terms within the problematic contexts are referential, then uniformly interpreted quantification into those contexts of a referential sort is possible. But what if the terms are irreferential? Is uniformly interpreted quantification, though of a non-referential sort, still possible?

The most straightforward case is when the terms within the problematic contexts are referential, but not standardly so. Uniformity then indeed provides a meaning for the quantifiers; for they can be taken to be similarly referential—ranging over the non-standard, not the standard, referents of the terms. The argument from the failure of substitution fails to get a grip, even under the assumption of uniformity; for what the truth of the identity sentence yields is the identity of the standard referents of the given and substituted terms; yet what is required is the identity of their non-standard referents.

A specific example of this sort is provided by the Fregean account of modal contexts. The term '9', let us say, in $\Box(9 > 7)$, is taken to refer to its customary sense. So a uniform interpretation of quantification would have us quantifying over the appropriate senses; $\exists x \Box(x > 7)$ would be true simply because $\Box$(the number of planets > 7) is true (under a *de dicto* reading of the description).

This is the interpretation proposed by Church [1943] in an early review of Quine's 'Notes on Existence and Necessity'. Quine originally found it acceptable, scruples over sense aside, but subsequently found reasons to reject it. In evaluating those reasons, it will be important, once again, to distinguish between logical and metaphysical considerations. One can imagine an invidious form of essentialism that allows essential properties to senses or other intensional entities but not to ordinary individuals. To this the radical anti-essentialist (and also the radical essentialist) may object that intensional entities and ordinary individuals should be treated on a par. I take it that this is part of the point of Quine's attack on Church: essential attribution to intensions and individuals alike calls for favouritism among descriptions (Linsky [1971], 153).

This is an interesting question. But our present concern is with the logical issue. On this, Quine's position is that the modal contexts are equally susceptible to failure of substitution regardless of whether the reference is to intensions or individuals. His example (ibid.) is:

$$A = (\iota x)[\,p \,\&\, (x = A)]$$

where A is an attribute and 'p' stands for a contingent truth. Substitute one for the other in a modal context and truth-value may change.

But the failure of substitution is no more relevant to the coherence of the quantification in this case than it was in the original case with reference to individuals. For the terms of the identity sentence will refer to the attribute A; but those same terms will refer in a modal context to appropriate 'higher-order' intensions. The referents of the terms inside and outside the context will again pass one another by.

How can Quine have failed to see this? I think he must have been assuming not only the local uniformity of term to variable but also the global uniformity among different occurrences of a term. This entails that the occurrences of a term inside and outside of a modal context must have the same reference, and so an escape from the consequences of a failure of substitution is no longer possible. One is instead led to the kind of systematically disambiguated language that was later proposed by Church [1951].

Another case of non-referential uniformity is when the terms have a double or multiple role. An example, for Quine, is provided by the occurrence of 'Giorgione' in 'Giorgione is so-called because of his size', for the term both picks out the man and draws attention to itself. Perhaps all linguistic roles should be assimilated to the referential role. If this is done, then these will be the only two possibilities: either the term refers to a single thing, be it a standard or a non-standard referent; or it refers to several things.

In the present case, it is not completely clear what the requirement of uniformity amounts to. There is the uniformity of variable to term and of term to term. If a term has several roles, is there to be the uniformity of variable to term with respect to all of those roles or only with respect to some of them? We can imagine that the term has several syntactic occurrences, one for each role, though only one typographic occurrence; the several syntactic occurrences coincide in the single typographic occurrence. The variable expression will likewise have several syntactic occurrences. But will it occur *as* a variable in all of these occurrences or in only some of them?

At one extreme is the view that the variable is to serve as a variable in each of its roles or occurrences. The uniform interpretation will then be one in which the variable is multivalued, with as many values as there are roles. For recall, a variable, when uniformly interpreted, will behave just like a term, but relative to an assignment; so each of the roles enjoyed endogenously by the term will be enjoyed exogenously by the variable.

In the case of the Giorgione example, the corresponding uniformly interpreted quantifier will simultaneously range over a pair of values: one an individual, and the other a term for the individual. Under this interpretation, the existential sentence '$\exists x$ (x is so-called because of his size)' will not only be meaningful but true; for when x takes as its values the man Giorgione and his name 'Giorgione', the open sentence '(x is so-called because of his size)'

will be satisfied. So even on Quine's construal of the example, uniform quantification into the context will be possible!

A less bizarre example is provided by Carnap's method of extension and intension, as propounded in *Meaning and Necessity* [1947]. Carnap takes a singular term to have a double linguistic role: one given by its extension, and the other by its intension. The corresponding interpretation of quantification therefore requires the assignment of both an extension and an intension as a value to a variable. Under this interpretation, the intelligibility of quantification into modal contexts is then guaranteed.

Quine makes the same objection to Carnap's account as to Church's. But the two are not really on a par. We saw that it was essential, if Quine's objection against Church was to be sustained, that the interpretation of different occurrences of the same term should be uniform; the possibility of non-standardly referential occurrences of terms was then ruled out. But uniformity of this sort creates no difficulties for Carnap. Terms will have the same double linguistic function both inside and outside modal contexts; it is just that only the extensional aspect of that function is relevant to the outside occurrences.

What makes Quine blind to the resilience of Carnap's account is that he does not take seriously the intended interpretation of Carnap's language. He says (Linsky [1971], 153) of the 'curious double interpretation of the variables' that 'this complicating device has no essential bearing and is better put aside'. He treats the variables as referential over intensions instead and then has no difficulty in restating his objection. But if I am right, the 'curious double interpretation' is of the essence of the matter; it is this which renders a fully uniform account of the language possible. Of course, there may be independent objections to the double interpretation, either as an account of ordinary language or as an approach to logical symbolism. But that is another matter.

The other extreme view is that the variable will serve as a variable in only one of its roles, presumably the one that for the term is most unproblematically referential; with respect to the other roles, the variable will function exactly like a term. The variable will therefore be single-valued, though the satisfaction of an open sentence may depend not only on the value assigned to the variable but also on the variable itself. We will have, in fact, the literalist use of the quantifiers.

Now in the case of the Giorgione example, the variable 'x' of 'x is so-called because of his size' will take a single object as value, but will depend for its satisfaction on the identity of 'x' in the same way that an instance of the open sentence depends for its truth upon the identity of the corresponding term. The existential sentence '$\exists x$ (x is so-called because of his size)' is therefore presumably false; for even if there is someone called 'x', it is not

likely that he is called 'x' because of his size. Again, a uniform interpretation of quantification into the context is possible!

Similar considerations apply to certain construals of modal discourse. Suppose we so understand the necessity operator that terms within its scope are taken not only to have their standard referents but also to pick out themselves or some feature of themselves. Uniformity will then deliver an appropriate variety of literalist satisfaction and quantification.

It may well be that the concept of necessity is such that 'Necessarily, Cicero = Ciero' is true and 'Necessarily, Cicero = Tully' is false. The concept, in its operation, is object-blind; it is not capable of looking past the terms to the objects which they denote. It is then plausible that the corresponding concept of satisfaction is such that 'Necessarily, $x = x$' is always satisfied and 'Necessarily, $x = y$' is never satisfied. If the identity of the closed terms in 'Cicero = Cicero' can guarantee its necessary truth, then, by parity of reasoning, it would appear that the identity of the variables in '$x = x$' can guarantee its necessary satisfaction; and similarly for the case in which the terms or variables are distinct. In some such way as this, therefore, it should be possible to provide a uniform account of literalist quantification into contexts governed by an object-blind operator for necessity.

This completes our discussion of the possibilities of quantification under conditions of uniformity. I have talked as if it were a matter of decision whether or not the quantifiers were to be referential or not. It is for this reason that I have separately considered the two cases. But the requirement of uniformity actually gives an objective meaning to the question; for the nature of the quantifier will be implicit in the use of the term from which it derives. If it is asked, for this neutral notion of quantification, whether quantification into a given context is possible, then the answer may well be 'Yes', even though no term is capable of occurring referentially within that context.

Quine would be unhappy with this conclusion; for there is a presumption in his work that the quantifiers are referential. This would be understandable if there were an equal presumption that terms are to occur referentially. But there is not. A double standard operates; terms are allowed to be irreferential, but variables are not. It is hard to see what can possibly justify this bias. The proper conclusion to draw from the irreferentiality of a given term is not that the corresponding quantification is impossible but that it is, if possible, similarly irreferential.

3
Quine on Quantifying In

Quine's argument against quantifying into modal contexts can be presented in the following way. Suppose that a statement such as $\exists x \Box (x > 7)$ makes sense. Then one should intelligibly be able to ask: for which objects x is the condition $\Box (x > 7)$ rendered true? But one cannot so ask. Specify the object as '9' and the condition is apparently rendered true; specify the object as 'the number of planets' and the condition is apparently rendered false. Fail to specify the object one way or another and there is no saying whether the condition is rendered true or false.

Central to the argument is a fact about substitution. Upon substituting '9' for 'x', the condition $\Box (x > 7)$ becomes a true statement; upon substituting 'the number of planets' for 'x', the condition becomes a false statement. Call an instance of such a fact a *failure of substitution*. Call the notion of an object rendering a condition true independently of how it is specified *objectual satisfaction*. Then it is because of the failure of substitution that there is thought to be no coherent notion of objectual satisfaction.

However, there are two quite separate accounts of why the one is a reason for the other. According to the first, the failure of substitution provides a general reason for rejecting the notion of objectual satisfaction. Given such a failure, the occurrence of the term upon which a substitution is made will not be purely referential, it will not be used solely to pick out its object. But given that this is so, we can have no understanding of what it is for an object to satisfy the corresponding condition.

On the other account, the failure of substitution provides a special reason for rejecting the notion of objectual satisfaction in its application to necessity conditions. For there is no understanding of *de re* necessity except in terms of *de dicto* necessity, no understanding of what it is for an object to satisfy a necessity condition $\Box \psi(x)$ except in terms of the truth of the instances $\Box \psi(t)$ (or perhaps of comparable closed sentences in which

terms give way to predicates). So given a failure of substitution, the satis-faction of such a condition will arbitrarily depend upon the term (or predi-cate) by which the object is specified.

The two accounts give rise to very different considerations. The first belongs to the philosophy of language. It is part of the general question of what, if anything, accounts for our understanding of quantification and satisfaction. The second belongs to metaphysics, loosely construed. It is part of the general question of what, if anything, accounts for the existence of necessary *de re* connections.[1]

It is my aim in this chapter to evaluate Quine's argument against quanti-fying into modal contexts, dealing first with the peculiarly modal consider-ations and then with the more general logical considerations.

The Special Argument

I follow Quine in distinguishing between the strict and the non-strict mo-dalities. The strict modalities include logical and analytic necessity, while the others include natural and metaphysical necessity. I also follow Quine in confining my attention to the strict modalities. Indeed, I shall concentrate on the case of logical necessity, though most of what I want to say will apply without essential modification to the case of analytic necessity.

There is a general difficulty in defending or challenging the intelligibility of any notion. For if the dispute is to have any point, it must be possible to identify the notion in such a way as to not presuppose its intelligibility. One cannot therefore identify the notion in the usual way as 'the concept so and so'; for the objector to the notion will fail to attach any reference to the phrase and so will fail to say, in the intended sense, that anything is unintelligible.

The obvious way out of this difficulty is to make the dispute about words. The proponent of the notion will have his chosen way of expressing it. He will therefore take the chosen form of words to be intelligible, while his opponent will not.

But there is a danger in this way out. For it mixes up two distinct questions: one is the intelligibility of an idea without regard for how it might be expressed; the other is the intelligibility of a form of words without regard for what it might express. It is perfectly conceivable that a philoso-pher might find the chosen form of words intelligible, not because he found the notion under dispute intelligible, but because he took the words to express some other notion altogether. Likewise, it is perfectly conceivable

---

[1] It was Kaplan [1969] (Linsky [1971], 41), who first made explicit the different consider-ations to which the two kinds of reason give rise. See Barcan Marcus [1990] for an account of the transition from one reason to the other in Quine's work.

that a philosopher might find the chosen form of words unintelligible, not because he found the notion unintelligible, but because he thought the words incapable of expressing the notion or, indeed, any notion at all.

These possibilities need to be taken seriously in the case of *de re* necessity. For suppose our chosen form of words is '$\Box\psi(x)$'. Then someone who accepts Quine's general argument from substitutivity may find these words unintelligible and yet still believe in necessary fulfilment. Indeed, Quine [1960] was in this position in *Word and Object*, though in regard to belief and not modality. Or again, someone who rejects Quine's general argument from substitutivity may find the words intelligible and yet disbelieve in necessary fulfilment. Indeed, Carnap [1947] was in this position in *Meaning and Necessity*, since the satisfaction of the condition $\Box\psi(x)$ would depend upon the intension associated with the variable x.

If a particular form of words gives rise to a possible misunderstanding, then another form of words may be chosen in its place. So in the present case, instead of saying 'x satisfies $\Box\psi(x)$', we may use a relational idiom and say 'x necessarily fulfils $\psi(x)$'. Indeed, it was just by means of this alternative wording that I attempted to draw the contrast between the acceptance of the original form of words and the acceptance of the notion.

Now it is true that the peculiar difficulties concerning substitutivity are avoided upon the adoption of this new form of words. But the general difficulties in identifying the notion under dispute are not so readily avoided. For either the chosen form of words is simple or it is complex. Suppose it is complex. Then there are again two cases; either the construction from simpler elements is problematic, or it is unproblematic. If the construction is problematic, then there is the danger of interference from the different ways in which the expression might be understood. If the construction is unproblematic, then the question reduces to the intelligibility of the simpler elements and might as well have been stated in those terms in the first place. Suppose, on the other hand, that the expression is simple. It is then hard to see what the dispute is about. For what common understanding of the simple expression is it that falls short of a grasp of its meaning and yet enables the parties to the dispute to sensibly debate its intelligibility?

Suppose, for example, that we take the predicate for necessary fulfilment to be primitive. Then what is to stop someone saying that the predicate is intelligible on the grounds that an object necessarily fulfils a condition just in case it actually fulfils the condition, or even just in case it fails to fulfil the condition? What is it in our common understanding of the predicate which serves to rule out such interpretations?

It is clear, once this line of reasoning is pursued, that the proper formulation of disputes over intelligibility is in terms of an identifying set of desiderata. The question is ultimately not about a form of words, but about the existence of an intelligible notion which conforms to certain desiderata.

Traditional disputes over intelligibility are best seen in this way. Infinitesimals are what conform to the demands of proof in which they are used; the self is something which, in an appropriate sense, is 'beyond' experience; and so on.

In the present case, it seems possible to give a rather sharp formulation of the desiderata by which the *de re* notion of logical necessity is to be identified. There are notions of truth and of satisfaction. There is also a notion of logical truth, which stands in a certain relationship to the notion of truth. The question then is whether there is anything that stands in the same relationship to the notion of satisfaction. Is there anything that is to satisfaction what logical truth is to truth?[2]

Other issues of intelligibility also take this special form. We may ask: is there an intelligible notion of subtraction on the real numbers, or on the cardinal numbers (transfinite or finite)? This question may then be put in the form: is there a notion that is to addition on the reals (or on the cardinals) as subtraction on integers is to addition on integers?

Of course, we have here only a scheme. To get a definite question, we must say what relationship it is that holds between the original pair of notions, and in such a way as not arbitrarily to limit its application. This can be done for the case of numbers. For the result of subtracting b from a should be the solution x of the equation $b + x = a$. We see then, on this construal of the question, that there is an intelligible notion of subtraction for the reals but no intelligible notion of subtraction for the cardinals (even with the restriction that $a \geq b$).

Something similar can be done for the logical notions. Logical truth is truth that can be determined on the the basis of logical form. Logical satisfaction is therefore satisfaction that can be determined on the basis of logical form. Our question is therefore whether we can meaningfully single out a special subclass of the relationships of satisfaction, those that can be determined on the basis of logical form.

Now it might be thought evident, once the question is put this way, that there is no such intelligible notion of logical satisfaction. For the notion of logical form only properly applies to sentences and to other parts of speech (or perhaps to what they convey). So how can we even talk of logical form for a relationship of satisfaction that is capable of holding between an ordinary physical object, let us say, and a condition?

But matters are not so simple. Suppose that we have a relationship of satisfaction, one that holds between a single object a and a condition $\phi(x)$. Then the question is: when we attempt to determine whether the relationship holds on the basis of logical form, what is it that is meant to have the

---

[2] We shall not make anything of a distinction, which might be drawn, between logical truth and logical necessity.

logical form, what should we look at? Now one might wrap up the object and the condition into a single entity, perhaps the ordered pair $\langle a, \phi(x) \rangle$, and scrutinize that for logical form. And one might then decide, given the non-linguistic nature of the object a, that the single entity had no logical form and that there was therefore no logical form on the basis of which the relationship of satisfaction could be determined to hold.

But such a response cannot be right. For there is a notion of logical implication which stands to material implication as logical truth stands to truth. Now suppose we follow the same strategy in determining whether the sentence ϕ logically implies the sentence ψ. We will then wrap up the two sentences into a single entity, say $\langle \phi, \psi \rangle$, and go by the logical form of that single entity. But if logical form only properly applies to entities of a linguistic nature, then the non-linguistic nature of the ordered pair, or some comparable entity, is as much an impediment to crediting it with logical form as is the non-linguistic nature of one of its constituents. We will therefore incorrectly conclude that there is no coherent notion of logical implication.

It is clearly arbitrarily restrictive to suppose that the logical form on the basis of which a logical relationship is determined to hold should be that of a single entity. We should instead be allowed to look separately at the logical form of the relata. Thus if the sequence of relata is $a_1, \ldots, a_n$, then the information we should extract is a corresponding sequence $\alpha_1, \ldots, \alpha_n$ of logical forms.[3]

There still remains a difference, under such an approach, between the relations of logical satisfaction and of logical implication. For with the case of implication, both relata possess logical form; while in the case of satisfaction, only one of the relata will.

However, it is not clear that this should count as a reason for ignoring the logical form of *both* of the relata; it might simply be regarded as a reason for taking account of the logical form of only one of the relata. We thus arrive at what I call the *object-blind* account of logical satisfaction. A condition will be satisfied by an object and, in general, by a sequence of objects just in case

---

[3] This raises another problem, for the sequence of forms needs to be properly coordinated. If, for example, '¬(2 is even)' is to logically imply '¬¬¬ (2 is even)', then we need to know that the constituents in the logical forms that correspond to the two occurrences of '2 is even' are appropriately correlated. It is usual to think of the (most fine-grained) logical form of an expression as unique. But this is not a conception of logical form that will permit the proper co-ordination. For we must then say that '2 is even' and '2 is odd' have the same logical form and so will be incapable of distinguishing between '2 is even' logically implying '2 is even' and '2 is even' logically implying '2 is odd'. What then distinguishes the two cases is that in the first we have the same duplicate forms, while in the second we have appropriately different duplicate forms. However, this is a general problem about the concept of form, and has no special bearing on the particular problem at hand.

it can be determined to be so satisfied on the basis of the logical form of the condition alone.

On such a conception, many conditions will be logically satisfied; for example, a condition of the form '$Fx \vee \neg Fx$' will be logically satisfied by any object. However, there will be no distinction among satisfiers; if one sequence logically satisfies a condition, then so does any sequence. Indeed, a condition $\phi(x_1, \ldots, x_n)$ will be logically satisfied by a sequence just in case the universal sentence $\forall x_1 \ldots \forall x_n \phi(x_1, \ldots, x_n)$ is logically true.

So we see that even if we restrict logical form to linguistic items, it is still possible to make out an intelligible notion of logical satisfaction.

However, I am inclined to think that the principal assumption of the object-blind account, that the objects make no contribution to logical form, is itself in error and that the account is therefore to be rejected. The issue turns on what should be taken to be the more basic bearers of logical form—physical tokens or abstract types (or meanings). All can agree that there is a sense in which sentence tokens have logical form and also a sense in which sentence types have logical form. But do we attribute a logical form to a sentence token because it is a token of a type with that form or do we attribute a logical form to a sentence type because it is a type of a token with that form? Should we think of the types as codifying the logical form of the tokens or, alternatively, should we think of the tokens as representing the logical form of the types?

If we adopt the former view, that the tokens come first, then it does indeed seem reasonable that the objects should be ignored in any account of logical satisfaction. For take satisfaction, logical or otherwise, to be a relation between a sequence of objects and a condition. (It would not matter, for my purposes, if we took the first relatum to be something like an assignment of objects to variables). In a sequence $a, a, \ldots$ beginning with an identical pair of objects, can we now recognize the reoccurrence of the object a as part of the logical form of the sequence? Should logical form be primarily attributable to tokens, then reoccurrence, in the sense that is relevant to the determination of logical form, would appear to consist in the occurrence of appropriately similar physical tokens: in an identity sentence '$t = t$', for example, it is the occurrence of two appropriately similar tokens of 't' that helps define its logical form. But it is not in this way that we have a reoccurrence of objects in the sequence; and so it is not clear that the reoccurrence of objects in the sequence can, with equal legitimacy, be taken to help define its logical form.

Suppose, on the other hand, that we adopt the much more plausible view that it is the types (or the meanings) that come first. Then there would appear to be no good grounds for distinguishing between the reoccurrence of an object in a sequence and the reoccurrence of a constituent in a sentence-type (or proposition). Indeed, if we took a sentence-type simply to be a sequence

of symbols, then the two notions of reoccurrence would be exactly the same. But even without such a conception of sentences, it would still appear to be in the same abstract sense that we may talk of reoccurrence in the two cases.

We are therefore led, on this alternative view, to what I call the *object-sensitive* account of logical satisfaction. A condition will be satisfied by a sequence of objects just in case it can be determined to be satisfied on the basis of the logical form both of the condition and of the objects. The logical form of the sequence will be given by the pattern of identities and difference, i.e. by which members of the sequence are the same and which distinct. So the condition '$x = y$' will be logically satisfied by the sequence a, a, since it is part of the logical form of the sequence that its two members (which are respectively assigned to x and to y) are the same; and similarly the condition '$\neg(x = y)$' is logically satisfied by the sequence a, b, for a distinct from b. In general, let $\pi(a_1, \ldots, a_n)$ be the statement that says which of the objects $a_1, \ldots, a_n$ are the same and which distinct. Then the condition $\phi(x_1, \ldots, x_n)$ will be logically satisfied by the sequence $a_1, \ldots, a_n$ just in case the sentence $\forall x_1 \ldots \forall x_n(\pi(x_1, \ldots, x_n) \rightarrow \phi(x_1, \ldots, x_n))$ is a logical truth.

We therefore see that there is a perfectly intelligible notion of logical satisfaction, even if there is some doubt as to whether it should be object-sensitive or object-blind.

What then of Quine's objections? He has argued that such a notion can only be purchased at a certain cost, the commitment to Aristotelian essentialism. One must adopt 'an invidious attitude toward certain ways of uniquely specifying' an object, seeing these ways, rather than others, 'as somehow better revealing its "essence"' ([1963], 155; [1966], 173–4).

But the force of the objection in the present case is far from clear. First, we may note that if it is the logical satisfaction of a condition by a *single* object that is in question, then no discriminatory, let alone invidious, attitude toward ways of specifying the object is required. An object will logically satisfy a condition $\phi(x)$, on either the object-blind or the object-sensitive account, just in case the corresponding closed sentence $\forall x \phi(x)$ is logically true, and hence regardless of how the object might be specified.

When we turn to logical satisfaction by a sequence, the situation is somewhat different. On the object-blind account, again no discrimination is required. But on the object-sensitive account, it is. For the specification of a pair of objects as the same, should they be the same, or as distinct, should they be distinct, will possess a special status; such specifications can indeed be seen to somehow better reveal the essence of the objects.

So such an attitude is discriminatory. But is it invidious, without good reason, and therefore bad? If we were simply to stipulate that certain specifications were to be given a special status, then that would be objectionable; or if we were arbitrarily to presuppose that certain specifications were special, then that again would be objectionable. But we do neither.

We explain the notion of logical satisfaction in terms of logical form; no special specifications of objects are either stipulated or presupposed. Rather, it is a natural consequence of the account, not an arbitrary feature of its formulation, that identity and distinctness turn out to have a special status.

Quine's misgivings do indeed have some force in regard to the *de re* application of *metaphysical* necessity. We may want to say that 9 is necessarily greater than 7, but not necessarily the number of planets. And how can this be so unless a special status of the one specification of the object over the other is somehow presupposed? But in regard to the *de re* application of *logical* necessity, the very kind of modality that is Quine's primary concern, it would seem that similar misgivings are without foundation.

The General Argument

I turn now to the logical version of Quine's argument against quantifying into modal contexts. The argument may be broken down into five steps: (1) It is claimed that occurrences of singular terms within modal contexts are not open to substitution—replacement with coreferential terms will not in general preserve truth-value. (2) From this it is inferred that such occurrences of terms are not purely referential—they are not used solely to pick out their object. (3) From this it is inferred that the corresponding occurrences of variables are not purely referential—they are not used solely to pick out their values. (4) From this it is inferred that the concept of objectual satisfaction is not meaningfully applicable to the condition formed with the help of the variables. (5) From this it is then inferred that quantification with respect to these variables is incoherent.

I should first like to discuss some of the general issues raised by this argument, considering each of the steps in turn. The main upshot of the discussion will be that the argument only goes through once a certain requirement of uniformity is presupposed. With this and other qualifications in mind, I would then like to discuss the application of the argument to the question of quantifying into modal and other problematic contexts.

Step (1): The Failure of Substitutivity

Quine's standard example concerns the number of planets: it is necessary that 9 is greater than 7 and yet not necessary that the number of planets is greater than 7, even though 9 is in fact the number of planets.

To such examples it is often objected that they depend upon adopting a narrow scope reading of the critical terms. Employ a wide scope reading instead and the failure of substitution disappears. It will be necessary that 9 is greater than 7; and it will also be necessary that the number of planets is greater than 7.

Since this point has been the source of so much misunderstanding, it may be worthwhile to consider it with some care. It may be agreed that the intelligibility of the wide scope reading and of the quantification into the modal context stand or fall together: if the one is intelligible then so is the other. If therefore Quine's argument has depended upon excluding the wide scope reading, i.e. if it had been one of its assumptions that the wide scope reading was not available, then the argument would clearly have been circular. But it depends upon no such assumption. For the purposes of his argument, Quine can afford to be agnostic about the existence of the wide scope reading; he need merely insist on the legitimacy of the narrow scope reading and on the consequent failure of substitution for it.

Of course, if the argument succeeds, if it is shown that quantification into modal contexts is unintelligible, then the wide scope reading is thereby also excluded—but as a consequence of the argument, not as an assumption. Someone who holds to the intelligibility of the wide scope reading is therefore objecting to the conclusion of the argument. If he would maintain his position, then it remains incumbent upon him to show what is wrong with the argument itself.

Step (2): The Inference to the Irreferentiality of Singular Terms

The reader could be forgiven for thinking this step innocuous. For does not the irreferentiality of the occurrence of the given term simply follow by definition from the failure of substitution?

But it has to be recognized that there are two concepts of referentiality, one informal and the other technical. According to the technical concept, an occurrence of a singular term in a sentence is purely referential if the truth-value of the sentence is preserved upon the substitution of coreferential terms. According to the informal concept, an occurrence of a singular term in a sentence is purely referential if it 'is used purely to specify its object, for the rest of the sentence to say something about' (Quine [1960], 177). These two concepts are different in character: one concerns the behaviour of the sentence under substitution; the other concerns the role of the singular term.

If referentiality is taken in the technical sense, then irreferentiality of the occurrence of the term does indeed follow by definition. However, it is irreferentiality of the occurrence in the informal sense that is relevant to the cogency of the argument; it is this that makes the unintelligibility of quantifying in plausible. And it is therefore for this sense of referentiality that the validity of the present step needs to be considered.

But once it is, we see that there are clear counterexamples. I shall present four cases in all, each illustrating a somewhat different way in which a term may occur referentially in the face of a failure of substitution.

(1) The first example comes from arithmetic (with the usual conventions governing the scope of '+' and '×'). The sentence '$2 \times 2 = 4$' is true; the terms '2' and '$1 + 1$' are coreferential; and yet the result '$1 + 1 \times 2 = 4$' of substituting one term for the other is false. So there is a failure of substitution, even though the initial occurrence of '2' is referential.

(2) For the purposes of the next example, we must imagine that three men are in a line, with Bill at the back and Fred at the front. Suppose now that Fred leaves. Then the sentence 'The man behind Fred saw him leave' is presumably true; the terms 'The man behind Fred' and 'The man before Bill' are coreferential; and yet the resultant sentence 'The man before Bill saw him leave' is false. So again there is a failure of substitution, even though no one would doubt that the initial occurrence of 'The man behind Fred' was referential.

(3) The third example is taken from the old notation for pounds, shillings, and pence (which some of us still remember). Under this notation, '1d' would be used to denote one penny, and '1/1d' would be used to denote one shilling and one penny, or thirteen pence in all. Bearing this in mind, we have the following: the sentence '1d was Kit Fine's pocket money in 1952' is true; the terms '1' and '1/1' are coreferential, at least in the sense that the identity sentence '$1 = 1/1$' is unequivocally true; and yet the sentence '1/1d was Kit Fine's pocket money in 1952' is, alas, not true. So substitution fails with the initial '1' referential.

(4) The final example comes from Hebrew (and was produced to my specifications by Ran Lahav). In Hebrew the word 'TSAFA' can either be the present, third-person, feminine form of a verb meaning to float, or it can be the past, third-person, masculine form of a verb meaning to observe. The word for moon can either be 'YARE'ACH', which is in the masculine, or 'LEVANA', which is in the feminine. So we see that: the sentence 'The moon (LEVANA) floats (TSAFA) in the sky (RAKEIA)' is true, with the occurrence of 'LEVANA' referential; the terms 'LEVANA' and 'YARE'ACH' are coreferential; and yet the sentence 'The moon (YARE'ACH) observed (TSAFA) the sky (RAKEIA)' is false.

It is tempting not to take these examples seriously, since they are clearly not the kind of case that Quine had in mind in formulating his test for irreferentiality. But in fact the examples pose a serious challenge to the intended applications of the test. For how do we know that they are not similarly aberrant, though in a less blatant way?

So what has gone wrong? When we examine each example in turn, we see that the substitution for the given singular term has unintended consequences. In the first example, '$2 \times 2 = 4$', there is a shift in the syntactic status of the subject-expression. Upon substitution, it changes from occurring as a constituent singular term to not even occurring as a constituent at all.

In the second example, 'The man behind Fred saw him leave,' there is a shift in the syntactic status, not of the the subject-expression, but of the remaining predicate-expression. Upon substitution, the pronoun 'him' in 'saw him leave' gets to stand in an anaphoric relationship to 'Bill' rather than to 'Fred'.

In the third example, '1d was Kit Fine's pocket money in 1952', there is no external shift in the syntactic status of the subject-expression and no internal shift in the syntactic status of the predicate-expression. Instead, the referent of the subject-expression changes from being the number 1 to being the vector of 1 and 1.

In the final example, 'LEVANA TSAFA RAKEIA', there is a semantic shift in the predicate-expression; 'TSAFA' goes from meaning *float* to meaning *observe*.

In each of the counterexamples, there is an unintended shift in either the syntactic or in the semantic analysis of the given sentence. So if the counter-examples are to be avoided, a certain requirement of uniformity should be imposed; it should be insisted that the given and the resultant sentences 'work in the same way'. If the given sentence is $\phi(s)$ and the resultant sentence is $\phi(t)$, then the syntactico-semantic analysis of the context $\phi(\ )$ of $\phi(s)$ should be the same as the syntactico-semantic analysis of the context $\phi(\ )$ of $\phi(t)$, with the one context 'feeding off' its term s in exactly the same way the other context 'feeds off' its term t.

Step (3): From the Irreferentiality of the Term to the Irreferentiality of the Variable

We may take the occurrence of a variable in a formula to be referential if it is used solely to pick out its value (for the rest of the formula, under an assignment of values, to say something about). This is the analogue of the informal concept of referentiality for closed terms. There is also an analogue of the technical concept. The occurrence of a variable x in $\phi(x)$ is referential in this sense iff the sentences $\forall x \forall y (x = y \rightarrow (\phi(x) \leftrightarrow \phi(y)))$, for y any variable distinct from x, are true (cf. Quine [1960], 167).

Just as with the term-to-term case, several examples of anomalous contexts can be given to show that irreferentiality of a closed term does not imply irreferentiality of the variable. One example, amusingly enough, derives from Quine's notation for quasi-quotation. Under this notation, the metalinguistic variable 'α' has a referential occurrence in '$\ulcorner \alpha \urcorner$ is a term', but the corresponding term 'o' in '$\ulcorner o \urcorner$ is a term' has an irreferential occurrence.

Another example, discussed as such by Kaplan in 'Opacity' [1986], arises from the 'suggestion for notational efficiency' of his earlier paper 'Quanti-fying In' [1969], n. 3. According to this suggestion, 'the number of planets'

will have an irreferential occurrence in 'Necessarily, the number of planets is greater than 7', but 'x' will have a referential occurrence in 'Necessarily, x is greater than 7', since it is treated as shorthand for 'x is such as to be necessarily greater than 7'.

A third example comes from arithmetic. The variable 'x' has a referential occurrence in '$2 \times x > 0$', but the term '$1 + 1$' does not even have a constituent occurrence in '$2 \times 1 + 1 > 0$'.

Again, I would propose getting around such counterexamples by imposing a requirement of uniformity. The syntactico-semantic analysis of the context must remain the same under substitution and must operate on the variable in the same way that it operates on the term.

Step (4): From the Irreferentiality of the Variable to the Non-Objectuality of Satisfaction

The terms 'referential' and 'objectual' are often used interchangeably in connection with the variables of quantification. But given the natural explanation of these terms, it is possible to draw a subtle but significant distinction between them.

A variable is referential, it will be recalled, if it is used solely to pick out its value. On the other hand, for satisfaction to be objectual it must be a relation that holds between a condition and an assignment of objects from the domain of quantification to the variables. Such a constraint on the form of the relation excludes two main possibilities. The first is that satisfaction should depend upon the manner in which the value of the variable is specified. The second is that the value of a variable should be anything other than an object from the domain; it should not, for example, be a term or a concept for such an object.

However, such a constraint does not guarantee that the variables should be referential. True, satisfaction cannot depend upon the manner in which the value of a variable is specified nor upon a non-objectual value. But it may depend upon the variable itself; and so the variable, through its very identity, can make a contribution to the conditions of satisfaction that goes beyond its value.

Variables which are objectual without being referential I call *literalist*. It is important to appreciate that the use of literalist variables gives rise to a genuinely new form of quantification. Philosophers are familiar with various non-objectual alternatives to referential quantification, such as the substitutional interpretation or Carnap's method of extension and intension. But the literalist quantifiers provide an objectual alternative to referential quantification. The values of the variables are standard. The clause for the quantifier is standard; an existential formula $\exists x \phi(x)$, for example, is satisfied by an assignment just in case the condition $\phi(x)$ is satisfied by an appropriate

variant of the assignment. It is just that the variable, through its own identity, can make a contribution to the satisfaction conditions.

We shall later give some modal examples of literalist quantification. A simple non-modal example, suggested to me by Allen Hazen, arises from the use of a many-sorted language in which the domains can overlap. Perhaps we have a sort for numbers and a sort for sets under a logicist construal of numbers as sets. Notwithstanding the overlap of the domains, we may wish '$x = n$' always to be false when the 'x' and 'n' are of different sorts. In such a way, we could permit within the metalanguage ('off-stage') a reduction of numbers to sets and yet not accept within the object-language ('on-stage') that any number was identical to a set. On a referential treatment of the variables, no such account is possible. But on a literalist treatment, it offers no special difficulties; for we can take it to be part of the satisfaction conditions for the identity formula '$x = n$' that the variables 'x' and 'n' should be of the same sort.

Another non-modal example, though of a more sophisticated kind, arises from the theory of truth. We may so use the truth-predicate T that the formula $T\alpha$ is satisfied by an assignment θ just in case (1) a formula ϕ of the object-language is assigned to the metalinguistic variable α, (2) individuals from the domain of the object-language are assigned to the free variables $x_1, \ldots, x_n$ of ϕ, and (3) the assignment of those individuals to $x_1, \ldots, x_n$ satisfies ϕ. (Succinctly put, $\theta \models T\alpha$ iff $\theta \models \theta(\alpha)$). On a referential treatment of variables, such an interpretation of the truth-predicate is incoherent, since the satisfaction of the formula $T\alpha$ does not simply depend upon the formula assigned to α. But on a literalist treatment, in which α serves not only to contribute the formula which is its value but also the assignment of values to the free variables of that formula, the interpretation is unproblematic.

It is not necessarily possible to tell from the content of the satisfaction conditions alone whether or not the variables are referential. The point may be illustrated by means of our previous example concerning a many-sorted language. Where α and β are any two variables, an identity formula $\alpha = \beta$ will be said to be satisfied by an assignment just in case (1) the objects assigned to both α and β are the same, and (2) the sorts of α and β are the same. But how does this result come about? What is the division of linguistic labour? One view is that the variables have a double linguistic role: to pick out their value and to pick out their sort. The identity predicate then picks out a relation that holds of object, sort and object, sort just when the objects are the same and the sorts are the same. The other view is that the variables have a single linguistic role: to pick out their value. However, it is part of the linguistic role of the identity predicate to 'look at' the sorts of the variables. It is not that the variables supply the sorts; the identity-predicate takes its own initiative in seeking them out.

On the first of these views, the variable is literalist; but on the second, it remains referential. The notion of a literalist variable stands in contrast, if you like, to the technical and the informal notions of referentiality. So a variable may be literalist in contrast to the technical notion of referentiality and yet still fail to be literalist in contrast to the informal notion.

Step (5): From the Non-Objectuality of Satisfaction to the Impossibility of Quantification

Several philosophers have attempted to disarm Quine's argument by appeal to a non-objectual account of quantification. They have supposed, for example, that the quantifiers should be substitutional or that the variables should be assigned both a sense and a reference as values. The question therefore arises as to whether it is just a matter for decision how the quantifiers are to be interpreted or whether it is in some sense an objective matter.

It is hard to extract from Quine's writings any compelling reason for taking the quantifiers to be objectual; and so his arguments would only appear to show that objectual quantification into modal contexts is to be rejected, not any kind of quantification whatever. However, there seems to be a way of giving objective meaning to the question of how the quantifiers are to be interpreted. For we may require that the interpretation of the quantified statement, $\exists x \phi(x)$ let us say, should be uniform with that of its instances. Our understanding of the general use of variables should conspire, along with our understanding of a particular closed sentence $\phi(t)$, to produce the required understanding of the quantified sentence $\exists x \phi(x)$.

But in what does that uniform interpretation consist? The closed term t will make a certain contribution to our understanding of the sentence $\phi(t)$. We may now recognize a certain range of variation, not in how the contribution is made but in what it is. The variable will then make an identical contribution within the given range of variation in a way that is fixed, not by the language itself, but by a hypothetical specification or assignment.

However, this construal of the issue is not one that especially favours the objectual interpretation of the quantifier. In so far as it is recognized that a closed term can occur irreferentially, it should equally well be recognized that a quantifiable variable can occur irreferentially; for the variable may work, relative to an appropriate kind of assignment, in the same way as the term. It is therefore perfectly conceivable that the resulting interpretation of the variable should be non-objectual. Suppose, for example that we follow Carnap in taking a term to pick out both an intension and an extension. Then on the corresponding uniform interpretation of the quantifier, the variable will be assigned both an intension and an extension as values, and so is clearly non-objectual.

Cases

We come finally to the application of the considerations of the previous section. We take up three kinds of question in all. First, are terms within the given problematic contexts referential? Second, is formal quantification into the context possible? Third, is ordinary language quantification into the context possible? In the case of each kind of question, we consider three kinds of context: those generated by quotation; those generated by Quine's Giorgione-example, 'Giorgione is so-called because of his size'; and those generated by modality. Each kind of context has a certain intrinsic interest and serves to illustrate somewhat different points.

Two novel aspects of our approach should be noted. The first is that the application of Quine's test for irreferentiality will be made to depend upon the appropriate uniformity assumption, either from term-to-term or from term-to-variable. Although this qualification may appear slight it makes an enormous difference to the epistemological status of the test. For its application now depends not only on the 'empirical' matter of the truth-values of certain sentences, but also on the 'theoretical' matter of how they function.

Secondly, we shall be careful to distinguish between three sources from which the problematic cases can come. The first is ordinary language, in our case English; the second is Loglish, the result of adjoining variables and formal quantifiers to English; and the third is Modalese, the 'pure' language of quantified modal logic. These source languages differ in the extent to which their interpretation is 'up to us' and hence differ in regard to how the appropriate uniformity assumption might plausibly be maintained.

Referentiality of Terms

It will be recalled that the inference from the failure of substitution to irreferentiality only goes through under the assumption of uniformity. This may make one wonder whether the substitution test for irreferentiality is ever of any use. For how can one tell that $\phi(s)$ and $\phi(t)$ are uniform unless one already knows how s and t are functioning in their respective sentences?

But the fact remains that one may know (or have good reason to believe) that the terms s and t function in the same way without knowing what that way is. A good illustration is provided by the case of quotation. The sentence '"Tully" has five letters' is true; the corresponding sentence '"Cicero" had five letters' is false. But it is plausible to maintain that the two sentences are uniform, especially under the assumption that one of the terms is referential. It may therefore be argued that neither term is referential, that neither term is being used solely to refer to the man.

But note that such an argument only makes plausible the conclusion that the terms are not *standardly* referential, that they are not being used within

the quotation context to refer to what they standardly are used to refer to. The argument still leaves room for the Fregean hypothesis that the terms within quotation contexts are non-standardly referential, that they are being used to refer, not to their standard referents, but to the expressions themselves.

A case where the assumption of uniformity is much less plausible is provided by the Giorgione-contexts. Quine ([1953], 140) asks us to consider the sentence 'Giorgione is so-called because of his size'. This is a truth. But upon substitution of the coreferential term 'Barbarelli', it converts to a falsehood. Should we therefore conclude that the given occurrence of 'Giorgione' is not purely referential?

Quine thinks we should. Indeed, he suggests that the term is only partly referential: it is used to pick out its object; but it is not solely so used. It seems that the term makes a double contribution to what the sentence says: it picks out its object; and it picks out, or presents, itself. The rest of the sentence then says of the two things picked out by the subject-term that the first is called the second because of its size.

But there is a far more plausible account of how the sentence functions, one that would make the term referential and would consequently lead us to reject the assumption of uniformity upon which the inference to irreferentiality depends. On this alternative account, the term 'Giorgione' is used solely to pick out its referent. However, the expression 'so' in 'so-called' is used to refer to that term. So what the rest of the sentence says of the referent Giorgione is that it is called 'Giorgione' because of its size. The sentence attributes a property to a single thing picked out by the subject-term, not a relation to a pair of things picked out by the subject-term.[4]

The case of terms within modal contexts raises somewhat different considerations. When the terms are descriptions, as in 'Necessarily, the number of planets is greater than seven', there is the possibility of both a wide and a narrow scope reading. Should the descriptions be given a wide scope reading, there will be no failure of substitution and hence no way of establishing irreferentiality on the basis of the substitution test. However, one cannot conclude that the descriptions occur referentially, though this is a view with some independent plausibility.

Should the descriptions be given a narrow scope reading, then there may be a failure of substitution. For example, the substitution of 'the sum of seven and two' for 'the number of planets', in the above example, converts a falsehood into a truth. It seems plausible, in such cases, to suppose that the syntactico-semantic context remains the same. We may therefore conclude that the descriptions in these contexts do not occur as referential

---

[4] A similar suggestion has been made by R. H. Thomason in 'Home is Where the Heart Is'[1979].

terms (though it is still possible, of course, that they occur as referential quantifiers).

When the term within the modal context is a proper name, it is not clear whether both a wide and a narrow scope reading are available. Is there a reading (wide scope) on which 'Necessarily, Cicero = Tully' is true, and also a reading (narrow scope) on which it is false? If one thinks there is, then there exist the same reasons as in the case of descriptions for denying referentiality to the names on a narrow scope reading. If one thinks otherwise, then the argument to irreferentiality from the failure of substitution will not go through.

It might be thought that, in such a case, one could argue for irreferentiality on the basis of the substitution of a description for a proper name. 'Necessarily, Cicero = Cicero' is true; 'Necessarily, the most famous Roman orator = Cicero' is false; and therefore the initial occurrence of 'Cicero' is irreferential. Indeed, Quine's original argument was of this sort, with the description 'the number of planets' being substituted for the name '9', although in fairness to Quine it should be pointed out that he took the sentence '9 < 7' to be analytic and so presumably regarded the numeral as an appropriately disguised description.

However, such an argument would be extremely weak. For it rests upon the assumption that names work in the same way as descriptions with narrow scope. And it is hard to see how this can be the case, even with all other considerations put aside, given that substitution holds up for names but not for descriptions.

Autonomous Quantification

We wish to know what Quine's argument can tell us about the intelligibility of such statements as $\exists x \phi(x)$ where $\phi(x)$ is one of the problematic contexts. The question divides according to whether the interpretation of quantification is autonomous or uniform. Under a uniform interpretation, our understanding of the quantified statement $\exists x \phi(x)$ is implicit in the understanding of its instances; the condition to which the quantifier '$\exists x$' applies is determined in conformity with the meaning of its instances $\phi(t)$. Under an autonomous interpretation, on the other hand, our understanding of the quantified statement $\exists x \phi(x)$ is unconstrained by the understanding of its instances; we are free to let the condition to which the quantifier applies be anything we like.

Let us consider first the possibility of autonomous quantification. In this case, it is quite clear that Quine's argument from the failure of substitution or from the irreferentiality of the singular term is powerless to demonstrate the unintelligibility of the quantification; for, as we have seen, it is critical to the argument that the quantified statement should be uniform with that of its

instances. Indeed, Quine himself concedes as much. For he writes ([19], 150): 'Nonsense is indeed mere absence of sense, and can always be remedied by arbitrarily assigning some sense.' Here, we may suppose, the absent sense is that which should have accrued from a uniform interpretation, and the assigned sense is that which could accrue from an autonomous interpretation.

But arbitrary as they are, autonomous interpretations are not without their interest; and we would do well to consider some of the different ways in which they might be provided. Under the substitutional account of the quantifier, there is a single natural way of providing an autonomous interpretation for all contexts. For we may so understand $\phi(x)$ that $\exists x \phi(x)$ is true iff some instance $\phi(t)$ is true. Of course, there is uniformity here of sorts. But it applies across contexts and not, in the required way, from quantified statement to instance.

For other conceptions of the quantifier, such sweeping generality is not attainable; so let us consider the case of the autonomous referential account as standing in greatest contrast, in this respect, to the substitutional account.

Since the quantifier $\exists x$ is to be referential, the satisfaction of the condition $\phi(x)$ depends only upon the values assigned to its free variables. But the question is: what are the satisfaction conditions for $\phi(x)$, given that they are not implicit in our understanding of the instances $\phi(t)$? The problem can be represented as one of translation: how can the condition $\phi(x)$ be 'reorganized' so that the satisfaction conditions are now apparent from the instances?

It may well be possible to deal with this question in a highly systematic and rigorous manner; but let me limit myself here to making some general remarks and considering some suggestive cases. It should be noted that our concern is to provide interpretations of quantification *without regard* for considerations of uniformity. It is therefore possible, in a case that was susceptible of a uniform interpretation, that the autonomous and the uniform interpretations should coincide.

If our concern were only to obtain *some* account of the satisfaction conditions, then this could readily be done. For we could stipulate that no assignment was to satisfy the condition $\phi(x)$. However, such an interpretation would not respect the meaning of the truth-functional connectives or of the other unproblematic constructions. If this is also to be arranged, then we can stipulate instead that no assignment is to satisfy any of the immediately problematic contexts and then let 'recursion' do the rest. But even so, it is doubtful whether such an interpretation would be of much interest.

There are perhaps two main factors which make for the interest of an autonomous interpretation. They usually go together, but may sometimes come apart. The first is that when $\phi(x)$ is the immediate context for which there is a failure of substitution, then the concept of truth for $\phi(t)$ should cohere or 'be of a piece' with the concept of satisfaction for $\phi(x)$. It is hard to

say more exactly what this coherence amounts to, but we may take it to be typified by the connection that was previously argued to exist between the concepts of logical truth and satisfaction.[5]

The second factor is that the logic of the quantifier should approximate as closely as possible to the classical logic of referential quantification. Indeed, often what makes an autonomous interpretation of the quantifiers 'workable' is that, but for some readily understood restriction, we can reason with them as if they were referential.

However, it is never possible to get the whole of classical logic, given that there is an underlying failure of substitution. For suppose that there are terms t_1 and t_2 and a condition $\phi(x)$ for which $t_1 = t_2$ is true but for which $\phi(t_1) \leftrightarrow \phi(t_2)$ is false. Then we cannot accept both Specification, $\forall x \Phi(x) \rightarrow \Phi(t)$, and Substitutivity, $\forall x \forall y (x = y \rightarrow (\Phi(x) \leftrightarrow \Phi(y)))$, for arbitrary terms t and contexts $\Phi(x)$. For let $\Phi(x)$ be the particular condition $\phi(x)$. Applying Specification to the resulting instance of Substitutivity as antecedent yields $\forall y (t_1 = y \rightarrow (\phi(t_1) \leftrightarrow \phi(y)))$; and applying Specification once again then yields the falsehood $t_1 = t_2 \rightarrow (\phi(t_1) \leftrightarrow \phi(t_2))$.

We are therefore forced to make a choice between the principles of Specification and Substitutivity; and, indeed, one useful way of classifying quantification theories which have been 'imposed' over a failure of substitution is in terms of how they make this choice. Very roughly, we may say that referential interpretations, of the type we are considering, favour the retention of Substitutivity, whereas substitutional-type interpretations favour the retention of Specification.

Quine's argument itself may be understood in these simple formalistic terms: quantified modal logic is to be rejected since it is incompatible with the standard principles of quantification theory. This is not how Quine would want to be understood or how he should be understood; for in case the instancing term is irreferential, the principle of Specification itself stands in need of justification. But the formalistic argument does serve to highlight the purely logical problems which must be overcome if a workable form of quantification is to be sustained.

After these general remarks, let us turn to some examples. Perhaps the best known case of autonomous quantification is provided by Quine himself. Quine has enjoined against quantifying into quotes; but his device of quasi-quotation, as expounded in *Mathematical Logic* ([1951], sect. 1.6) and elsewhere, enables us to do just that. It may be objected that this device only allows us to quantify into quasi-quotes, not into quotes. But quasi-quotes behave just like quotes in the absence of free variables; and, indeed,

---

[5] I assume that it is coherence in this sense, rather than the earlier concept of uniformity, that corresponds to Kaplan's notion of coherence in 'Opacity' [1986].

the reasons Quine has, from the failure of substitution, against quantifying into quotes apply equally well to quasi-quotes.

It might be thought that the method of quasi-quotation is geared to the special case of quotation contexts, but in fact it is generalizable to any context whatever. For let $\phi(x)$ be an immediately problematic context. Then we may reinterpret $\phi(x)$ as a condition on expressions: 'the result of substituting x for "x" in $\phi(x)$ is true'. And similarly when the context is a term $t(x)$, as in the case of quotation.

Substitutional quantification may also be regarded as referential quantification over expressions. We therefore see a close connection between quasi-quotation, once it is suitably generalized, and substitutional quantification. The main difference is that, in the case of substitutional quantification, the metalinguistic device is applied across the board to all of the contexts, whereas in the case of quasi-quotation, the metalinguistic device is only applied to the immediately problematic contexts. The method of quasi-quotation is, if you like, a half-hearted version of substitutional quantification.

In one respect, though, the two methods are not on a par. The variables of substitutional quantification apply straightforwardly to the ordinary predicates of the language. The variables under quasi-quotation do not; for their values are expressions rather than objects from the given domain. As a consequence, atomic combinations of ordinary predicates with the quasi-quotational variables must either be excluded or some special interpretation for them must be proposed.

These difficulties are avoided under the method of arc-quotation of Kaplan's 'Opacity' [1986]. Under this method, it is the objects, and not the expressions, that are assigned to the variables. There is therefore no difficulty in interpreting the application of ordinary predicates to those variables. The innovative step comes in the interpretation of the open quotational terms. The term '⟨x is a philosopher⟩', for example, will denote the result of substituting Quine (the man) for 'x' in the expression 'x is a philosopher' under an assignment of Quine to 'x'. Thus the referents of arc-quotation expressions may be hybrid objects, consisting of both linguistic and non-linguistic material.[6]

Natural as this method is, it is not generalizable to arbitrary contexts. Suppose that we have a problematic open term $t(x)$. Then we might try saying that the denotation of $t(x)$ under an assignment of a to 'x' is the denotation of the result of substituting a for 'x' in $t(x)$. But it is not in general clear what the denotation of the resulting 'expression' should be. It is a peculiarity in the case of quotation-mark contexts that there is something

---

[6] A related procedure was adopted, at my suggestion, by B. Richards in 'A Point of Reference' [1974], 431–41.

which we can naturally take the denotation to be, namely, the 'expression' that falls under the quotes.

An interesting feature of arc-quotes is that they may be used to legitimate what would otherwise be illegitimate clauses in the truth-definition. Suppose we so use the truth-predicate T that '$T\langle x$ is a philosopher$\rangle$', for example, is true under an assignment of Quine to 'x'.[7] Then we may correctly assert:

$$T'\exists x(x \text{ is a philosopher})' \leftrightarrow \exists x T'\langle x \text{ is a philosopher}\rangle'.$$

Note the essential use of arc-quotes to the right. It may even be possible, under suitable conventions of disambiguation, to make reasonable sense of such general assertions as $\forall\alpha\forall\phi(T'\exists\alpha\phi(\alpha)' \leftrightarrow \exists x T'\phi(x)')$ and thereby to vindicate what are normally regarded as sloppy habits of use-mention.

Let us pass over the question of autonomous quantification for Giorgione-contexts as being of no independent interest and move directly to the case of modal contexts. Perhaps the best known method of autonomous quantification in this case consists in associating with each object of the domain a unique standard term or name. A modal condition $\Box\phi(x)$ is then taken to be true of an object just in case the result $\Box\phi(s)$ of substituting the standard name of the object for the variable is true.

In a way, such a method combines the advantages of purely substitutional methods, like quasi-quotation, and of purely objectual methods, like arc-quotation. In contrast to quasi-quotation, it avoids embarrassment over the interpretation of variables in ordinary contexts; and in contrast to arc-quotation, it avoids the embarrassment over the interpretation of the variables in problematic contexts. But the element of artificiality is not avoided altogether, since the association of objects with names must somehow be given.

The method is capable of considerable variation. (1) Instead of associating a single name with each object in the domain, several names may be associated. Of course, many pairs of names will be intersubstitutable *salva veritate*; but there is no need for the several names associated with a given object all to be equipollent in this regard. (2) Sequences of names may be associated with sequences of objects in a way that is not generable from the simple association of names with objects. (3) Predicates (or 'incomplete' names) may be associated with the objects or with the sequences of objects. (4) Given the association, either of several names or of predicates, there are two different ways in which satisfaction can be explained. We may say, for example, that $\phi(x)$ is true of an object iff $\phi(s)$ is true of *all* the associated names or, alternatively, of *some* of the associated names s.

---

[7] As Kaplan has observed in Appendix C of 'Opacity' [1986], this requires, in case the values of the variables include expressions, that some method be used to distinguish between the substitution of a value *qua* object and *qua* expression.

The previous object-sensitive account of modality can be regarded as a special case of this method. Suppose that with each sequence $a_1, \ldots, a_n$ of objects there is associated the identity-type predicate $\pi(x_1, \ldots, x_n) =_{df} \bigwedge_{i,j} (\pm x_i = x_j)$, where the $\pm$ is blank if $a_i = a_j$ and $\neg$ if $a_i \neq a_j$. We may then say that the modal condition $\Box \phi(x_1, \ldots, x_n)$ is true of $a_1, \ldots, a_n$ just in case $\Box \forall x_1 \ldots \forall x_n (\pi(x_1, \ldots, x_n) \rightarrow \phi(x_1, \ldots, x_n))$ is true.

The object-blind account may be obtained in a similar way, simply by letting the predicate associated with an n-tuple be a trivial predicate, true of all n-tuples whatever.

The method is generalizable to all contexts and is in no way peculiar to the modal case; although, of course, which names or predicates are appropriately associated with the objects may well vary from case to case. The method is, in particular, applicable to the earlier example of quotation. It might be thought odd that quotation contexts should be equally susceptible of an 'essentialist' interpretation as modal contexts. But many domains of objects are conceived in terms of a canonical system of notation which may then be exploited to this end. With each natural number, for example, may be associated its designation in the arabic notation. We may then take '*n* is even', say, to designate '12 is even' under the assignment of 12 to '*n*'. In such a way, we obtain a highly natural account of quantification into quotation contexts, one that is actually of use in the interpretation of some programming languages.

Uniform Quantification

The autonomous interpretations of the previous section call for an innovative act of understanding. Our question now is: what understanding of the problematic quantified statements is already implicit in the understanding of their instances? Is Quine right when he states in 'Reference and Modality' ([1963], 150): 'But the important point to observe is that granted an understanding of the modalities . . . and given an understanding of quantification ordinarily so called, we do not come out automatically with any meaning for quantified modal sentences . . .'

Granted this requirement of uniformity, it appears that Quine's argument can get a grip. For given that the instancing terms are irreferential, the corresponding variables cannot be referential and so quantification into the context is unintelligible.

But our earlier discussion prepares us for several ways in which this line of reasoning might be challenged. In the first place, the irreferentiality of the instancing term does not simply follow from the failure of substitution. The case for irreferentiality will vary from context to context. For the Giorgione-example, 'Giorgione is so-called because of his size', it seemed most plausible to maintain that the term 'Giorgione' was referential. What explained

the shift in truth-value under substitution was the shift in the reference of 'so'. Uniform quantification, of a standardly referential sort, would then also be possible, if only this referent could somehow be kept fixed.

For the modal and quotational examples, it did not seem plausible to maintain that the terms were standardly referential. Uniform quantification, of a standardly referential sort, would therefore appear to be barred. However, it was still left open whether the terms were non-standardly referential. If they were, then the variables would likewise be non-standardly referential. Thus quantification over a non-standard domain of entities, far from being a subterfuge, would actually be dictated by the requirement of uniformity. It is in some such way as this that one can understand Church's Fregean proposal in his early review [1943] of Quine: quantification into modal contexts is permissible, but with respect to senses not objects.

In the case of quotation, the non-standard entities are expressions and the quantification is similar in effect to the quasi-quotational method of Quine. We therefore have an example of quantification which seen from one point of view is uniform and which seen from another point of view is autonomous. Indeed, part of the appeal of the quasi-quotational method is that it enables us to engage in the fiction that the quantification is uniform.

The second challenge to the Quinean line of reasoning depends upon being more sensitive to the exact content of the requirement of uniformity. May a quantified statement be uniform with only one of its instances or must it be uniform with all of them? If the reason for insisting upon uniformity is that our understanding of a quantified statement should issue from our general understanding of quantifiers and variables, then one instance should suffice; for our general understanding will apply in this one case, regardless of how it may apply or fail to apply in other cases.[8] So in demonstrating the unintelligibility of a quantified statement, it is not sufficient to establish irreferentiality for one of the instances, as Quine's argument would lead us to suppose. Irreferentiality must be established for them all.

For the quotation and Giorgione contexts, the point is of no real significance. For the instances themselves are uniform; if one of the instancing terms is irreferential then so are the others. But for the modal and other such contexts, the matter is different. For them it is plausible to draw a distinction between names and descriptions with wide scope, on the one hand, and descriptions (and possibly names) with narrow scope on the other. Even it is conceded that descriptions with narrow scope are irreferential in the given contexts, it is still possible that either the names or the descriptions with wide scope are referential; and if this is so, then they may be chosen as a basis for a uniform understanding of quantification.

---

[8] Even in the bizarre case in which different instances with a referential term produce different conditions, the resulting quantified statement will be ambiguous rather than incoherent.

I have so far assumed that if we find anomalies in the linguistic function of the singular terms of ordinary language, then we should accept them. But our aim may be to construct a logical symbolism that is free from such anomalies, that is syntactically perspicuous in regard to the semantic function of its different terms. For such a symbolism, a sound form of the Quinean argument goes through. Its upshot will be that it is impossible to have a language in which: (1) there are terms subject to failures of substitution in modal or other problematic contexts; (2) there is uniformity from term-to-term and from term-to-variable; and (3) the variables of quantification are (standardly) referential.

This, it seems to me, is the only negative conclusion of any real value to emerge from the logical argument. The conclusion is indeed not damning; but it does impose a substantial limitation on the form that the symbolism for quantified intensional logics can properly take.

The third challenge to Quine's argument resists the restriction to referential quantification. It is conceded, if only for the sake of argument, that there are instances of the problematic contexts with irreferential terms; and it is also conceded that referential quantification, at least when grounded in those instances, is therefore impossible. But it is denied that quantification *simpliciter* is impossible; for the quantification may be uniform and yet not referential.

One way for this to happen, with which we are already familiar, is for the instancing terms to be non-standardly referential. The corresponding uniform quantifiers will then range over the appropriate domain of non-standard entities.

But another way is for the terms to have double or multiple roles. An example, *for Quine*, is provided by the occurrence of 'Giorgione' in 'Giorgione is so-called because of his size'. For there the term has two roles: it picks out the man; and it picks out the name itself.

In such a case, it may not always be clear what the requirement of uniformity amounts to. For there is the uniformity of variable-to-term and of term-to-term. So if a term has several roles, then is the variable to occur *as* a variable with respect to all of these roles or only with respect to some of them?

At one extreme is the view that the variable is to serve as a variable with respect to all of the roles. The uniform interpretation is therefore one in which the variable simultaneously takes several values, one for each of the roles. In the Giorgione-case, this means that the variable that supplants 'Giorgione' will take two values, one a term and the other an individual. As a consequence, the existential sentence '$\exists x(x$ is so-called because of his size)' will not only be meaningful but true; for under the simultaneous assignment of Giorgione and his name 'Giorgione' to 'x', the open sentence 'x is so-called because of 'his size' will be satisfied.

A more significant example is provided by Carnap's method of extension and intension, as propounded in *Meaning and Necessity* [1947]. Carnap takes a singular term to have a double linguistic role: one given by its extension; and the other by its intension. The corresponding interpretation of variables therefore requires the simultaneous assignment of an extension and an intension; and under such an interpretation, quantification into the appropriately intensional contexts will be unproblematic.

At the other extreme is the view that the variable is to serve as a variable with respect to only one of its roles, presumably the one that for the term is most unproblematically referential; with respect to the other roles, the variable will function exactly like an ordinary term. As a consequence, the variable will be single-valued, but the satisfaction of an open sentence may depend not only on the value assigned to the variable but also on the relevant aspects of the variable itself. We will have what amounts to a literalist use of variables and quantifiers.

In the Giorgione case (again under Quine's construal of it), the variable 'x' in 'x is so-called because of his size' will take a single object as its value, but the open sentence will depend for its satisfaction upon the identity of 'x' in the same way that an instance depends for its truth upon the identity of the corresponding term. So the existential sentence '$\exists x(x$ is so-called because of his size)' is still meaningful, but presumably false: for even if there is someone called by the *name* 'x', he is not likely to be called by the *variable* 'x'; and even if he is called by the *variable* 'x', he is not likely to be called by the variable 'x' because of his size.

A more significant example is provided by modality. Let us suppose that the terms within modal contexts serve as contributors to logical form in such a way that 'Necessarily, Cicero = Cicero' comes out true while 'Necessarily, Cicero = Tully' comes out false. Under a corresponding interpretation of the variables, in which the role of contributor to logical form is kept fixed, the open sentence 'Necessarily, $x = x$' will be satisfied by any assignment of objects, while the open sentence 'Necessarily, $x = y$' will be satisfied by none. We thereby obtain a uniform and literalist account of quantification into contexts governed by an object-blind operator for necessity.

Ordinary Language Quantification

Finally, let us consider the question of whether *ordinary language* quantification into the problematic contexts is intelligible.

It is important to distinguish the question of intelligibility for ordinary language quantification from the corresponding question for formal quantification. If it could somehow be assumed that the ordinary language constructions were to be rendered in terms of the formal ones, then the

two questions would be the same. But this is not something that can simply be taken for granted.

What can Quine's argument show us, then, about the intelligibility of such sentences as 'Something is such that necessarily it is greater than seven'? It is evident that the argument cannot even get off the ground unless it is assumed that the quantified sentences are appropriately uniform with their offending instances. So what reason is there for making such an assumption?

There is some plausibility to the view that we understand quantified statements and their various instances on the basis of certain general principles concerning the meaning of the relevant operators, quantifier constructions, and types of terms. There is indeed a preference for general principles over special cases. But it is far from clear that such an amorphous requirement of uniformity would lead to the specific requirement of uniformity from quantified statement to offending instance. On a Russellian conception of language, for example, it would be part of our understanding of the problematic operators that we understand how they apply to singular 'objectual' propositions; and so there would be no difficulty in seeing how the quantified sentences could be understood, no matter now aberrant the behaviour of the unquantified substructure.

But even if we decided, at the end of the day, that no general explanation of the meaning of the quantified statements and their various instances was forthcoming, this still would not provide us with a strong argument for the unintelligibility of the quantified statements. For given a conflict between a direct intuition of intelligibility and a theoretical presumption of uniformity, it seems clear that it is the theoretical presumption that should be given up.

II

Issues in Ontology

This page intentionally left blank

4
Prior on the Construction of
Possible Worlds and Instants

Fundamental to Prior's conception of modality were two theses:

> The ordinary modal idioms (necessarily, possibly) are primitive[1]
> Only actual objects exist[2]

The first thesis might be called Modalism or Priority, in view of its nature and founder. The second thesis is sometimes called Actualism, and the two theses together I call Modal Actualism. Prior held corresponding theses about time:

> The tenses (it will be, it was the case) are primitive;[3] and
> Only present objects exist.[4]

His acceptance of the last thesis was tentative: he was unsure about the existence of past individuals[5] and he was, in any case, attracted towards a more basic ontology of 'stuff'.[6]

In contrast to these doctrines about time and modality, one might hold that ordinary modal discourse is to be explained in terms of possible worlds and possible individuals and that tense-logical discourse is to be explained in terms of instants and individuals, past, present or future. Thus 'Possibly someone is over ten feet tall' would be analysed as 'Some possible individual is over ten feet tall in some possible world', and 'It will be that someone flies to Venus' would receive an analysis of the form 'Some person (past, present, or future) flies to Venus at some time later than t'. These analyses are often associated with the possible worlds semantics for modal and tense-logic, although there is no need to regard the semantics as providing an *analysis* of the notions that appear in the object-language.

---

[1] Many references might be given. See e.g. 'Modal Logic and the Logic of Applicability' (Prior [1968c]).

[2] See Prior [1968a], 143.

[3] Ibid., esp. ch. 12.

[4] Ibid. 143.

[5] See the discussion in Prior [1967], 171–2.

[6] Ibid. 174, and Prior [1968a], 80.

The possible worlds analysis is incompatible with both modalism and actualism: for it analyses the ordinary modal idioms (indeed, it uses no intensional connectives at all) and it admits possibles. One could, in theory, deny either tenet without denying the other. The modalist could admit possibles and the actualist could attempt to analyse the modal idioms. However, neither of these middle positions is particularly plausible. For if the modalist admits possibles then why does he not accept the possible worlds analysis, and if the actualist does not accept the modal idioms then how, exactly, are they to be analysed?

The most plausible positions are the extreme ones and, indeed, they stand in natural opposition to one another. Both find a distinctive role for modality, but they locate it in a different aspect of language. For the modal actualist, possible objects do not exist; rather, the possible exists as a *manner* in which things happen. It exists as a mode, not an object. In the proper language for expressing modal truths, the modal primitives will be adverbial (sentential connectives) and the quantifiers will range over actual objects alone. For the classical possibilist, there is no special manner in which things happen; the possible is located in ontology. It exists as an object not mode. In the proper language for expressing modal truths, the quantifiers range over possible objects and the connectives are all truth-functional.

This ontological difference can be brought out by considering general possibilities. Take as an example, 'Possibly some individual is not actual'. For the possibilist, this is an existential claim to the effect that some possible individual is not actual. Therefore there must be some specific individual who is not actual. But for the actualist, this singularity is spurious; there can be no instance in virtue of which the sentence is true. The sentence states an irreducible general possibility, and no matter how well the individual is described, he can have no specific identity. The same considerations apply, *mutatis mutandis*, to time.

Which of these positions is correct is a large and difficult question. I am inclined to think that Prior's views on both time and modality are correct, but that the arguments for the two cases are rather different. My aim is not to discuss whether modal actualization, or its tense-logical counterpart, is correct. Rather, it is to discuss a problem that arises once modal actualism is accepted. The problem is this: the possibilist can analyse ordinary modal discourse in terms of possible worlds and possible individuals; but what is the modal actualist to make of the possibilist's discourse?

One answer to this problem is that all such discourse is illegitimate. But this is hard to accept, for a great deal that we want to say about time and modality is put in these terms. Some of this talk may not make sense. The domain of possible worlds involves the notion of totality twice over, for each possible world is a totality (of what is the case) and then the domain is the totality of such totalities. Now it may be that certain references to these or

related totalities lead to the type of paradox that is familiar in set theory. But all the same, a great deal of possible worlds (or instants) talk will still be in order.

A more acceptable answer is that this talk is legitimate, but not basic; it stands in need of analysis. The modal actualist will eliminate talk of possible worlds and possible objects in favour of the ordinary modal idioms and quantification over actuals. His tense-logical counterpart will do likewise. Thus the previous analysis will be turned on its head, with the connectives and restricted quantifiers—be they modal or tense-logical—coming first and the worlds or instants and unrestricted quantifiers emerging as a construct from them.

My aim in this chapter is to carry out this programme of reconstruction, at least in outline. I have often followed the lead of Prior, much of whose later work[7] arose from this programme. However, I cannot be sure that he would have approved of all of the steps I take.

1. The Simplest Case

The simplest case of reconstruction concerns the language of S5, i.e. the language of truth-functions and necessity. Let us begin with the possible worlds analysis of this language. The main idea behind the analysis is that the relation '*A* is true in possible world *w*' can be expressed as a classical (indeed, first-order) condition $A'(w)$. Thus '*P* is true at *w*', for *P* a sentence-letter is expressed as '$P'w$', where P' is a possible worlds predicate corresponding to *P*; the truth-functional connectives are represented by themselves; and '$\Box B$' is true at *w* is expressed as '$B'(w)$ for all possible worlds *w*'. This last clause states that $\Box B$ is true at a world if *B* is true at every possible world. Given the condition $A'(w)$, the modal formula A can then be given the translation '$A'(\omega)$', where ω is a constant designating the actual world. Thus a formula is said to be true if it is true in the actual world.[8]

The language into which we translate is that for the monadic predicate calculus. It contains a monadic predicate P' for each sentence-letter *P*, a quantifier $\forall w$ over possible worlds, and a constant ω for the actual world. In order to translate back into a modal language, Prior suggested that each possible world be treated as a world-proposition. *p* is a *world-proposition*—in symbols Qp—if it is true in one world alone or, to put it in modal terms, if it is possible that *p* is true and necessarily implies all truths, i.e. $\Diamond(p \land \forall q(q \supset \Box(p \supset q)))$. Prior's suggestion, then, was to replace $\forall w$ with a quantifier $\forall q$ over all world-propositions, i.e. to translate $\forall wA$ as $\forall p(Qp \supset A^*)$ where A^* translates A. The predication $P'w$ would then be

replaced by the strict implication $\Box(q \supset P)$, where q is the world-propos-ition corresponding to w, and the constant ω by the description 'the true world-proposition'.

The whole procedure involves three languages and two translations. There is the original or *primary* modal language for S_5, the *classical* language of possible worlds, and the *secondary* modal language for S_5 with propos-itional quantifiers. There is the *standard* translation (A to A') from the first modal language into the classical language and the *reverse* translation (A to A^*) from the classical language into the second modal language; putting these two translations together results in a third translation (A to A'^*) from the one modal language into the other. For example, suppose one starts with $\Diamond P$ in the original modal language. The standard translation takes this formula into $\exists w P' w$ (= 'P is true in some possible world') and then the reverse translation takes that formula into $\exists p (Qp \wedge \Box(p \supset P))$, i.e. '$P$ is strictly implied by some world-proposition'.[9] Thus the possible worlds rendering of $\Diamond P$ is seen to rest on the fact that each possibility can be embedded in a maximal possibility.

Two aspects of the reverse translation are worth noting. The first is that the possible worlds predicate $A'(w)$ (= 'A is true in w') is translated into a rigid predicate of propositions, i.e. one that necessarily holds or necessarily fails of any given proposition. (I have used the term 'rigid' in analogy to Kripke's use of 'rigid designation'; a rigid predicate has the same extension in each possible world.) For example, each atomic formula $P'w$ is translated into a strict implication. In the reverse translation, the only source of contingency is in the reference to the world proposition. Thus $P'w$ becomes 'the true world-proposition enjoys the (necessary) property of strictly im-plying P'. The reference is contingent and the predication necessary, as in the well-known sentence 'the number of planets is greater than 7'. This aspect of the translation will later be exploited in eliminating quantification over possible individuals.

The second aspect is that the correctness of the reverse translations depends upon two assumptions about the existence of propositions. The first is that world-propositions (generally, propositions) necessarily exist or, to be more accurate, that *necessarily* world-propositions necessarily exist. This assumption will be questioned later. The second is that for each possible world there is a proposition true in that possible world alone or, to put it in modal terms, that necessarily there is a true world-proposition.

Is this assumption justified?[10] Probably not, if one makes the nominalist requirement that each proposition be expressible by a single sentence. But it is not clear that such a nominalist should accept possible worlds talk in the

---

[9] See Prior and Fine [1977], ch. 6, pp. 87–8.
[10] See ibid. ch. 2, p. 33, for a brief discussion of this question.

first place. For if the existence of propositions is tied to sentences, the existence of possible worlds should be tied to their names and then there is a similar difficulty.

If one is more platonic about the existence of propositions, then justification is possible. For let X be the set of true propositions, and let q be the proposition that all the members of X are true. Then q is a true world-proposition. For q is true since all members of X are true, and q strictly implies all truths since all truths belong to X. To establish that *necessarily* there is a true world-proposition, note that the premises of the above argument (that X and q exist) are necessary. Therefore the conclusion (that there is a true world-proposition) is also necessary.

Simple as this argument is, it resorts to quantification over sets of propositions and yet has a conclusion whose formulation requires quantification only over single propositions. I suspect that a language with quantifiers over sets of propositions is the natural setting for a great deal of significant modal discourse.

2.   Introducing Predicates

Let us now consider the complications that arise from adding predicates and quantifiers to the previous language, thereby yielding the language of first-order logic and necessity. I adopt the type of semantics I have adopted before[11] for this and most subsequent quantificational languages. I shall not justify this choice, but I will consider, in §7 below, the consequences of adopting the Q-type semantics of Prior. For convenience, we suppose that the original modal language contains a predicate E for existence; and to avoid notational confusion, we use $x, y, \dots$ for the actualist variables of a modal language and $\mathbf{x}, \mathbf{y}, \dots$ for the possibilist variables of the classical language. As before, the relation 'A is true in the possible world w' is expressed as a first-order condition $A'(w)$. Thus 'Rxy is true at w', R a binary predicate, is expressed as $R'\mathbf{xy}w$, where R' is a possible worlds predicate (in the last of of its argument) corresponding to R; the truth-functional connectives and necessity operator are dealt with as before; and '$\forall xB$ is true at w' is expressed as '$B'(w)$ for every possible individual such that $E'xw$'. If we let the domain of a world be the set of individuals that exist in the world, then the last clause says that $\forall xB$ is true at a world if B is true of every possible individual in the domain of the world. Again, A is given the translation $A'(\omega)$.

The language into which we translate is still first-order, but two-sorted and polyadic. It contains variables $\mathbf{x}, \mathbf{y}, \dots$ for possible individuals as well as variables $w, v, \dots$ for possible worlds, and for each predicate in the modal

---

[11] Fine [1970].

language, there is a predicate with an extra argument for possible worlds in the classical language. Unfortunately, there is no automatic extension of the previous reverse translation to the new classical language. The difficulty is that possible worlds should satisfy the condition that if $\exists x B(x)$ is true in a world then an instance of $B(x)$ is true in that world.[12] So if world-propositions are to do duty for possible worlds, they should satisfy the condition that if $\exists x B(x)$ is strictly implied by a world-proposition q then an instance of $B(x)$ is strictly implied by q. But given an actualist interpretation of the quantifiers, there is no reason to suppose that this is true. Take any counter-example to the Barcan Formula, say 'Possibly there is a divine being but there is no possibly divine being'. Then there is a world-proposition that strictly implies the existence of a divine being, but there is no (actual) individual whose divinity the world-proposition implies.

One solution to this difficulty is to replace each possible individual with a property that is possibly instantiated; the property could then exist even if the individual did not. To be more specific, say that a property φ is an *individual essence*[13] *of* the (actual) individual x if φ is true of x and of x alone in each possible world or, to put it in modal terms, if necessarily for all y (φ holds of y iff y is x); and call φ an *individual essence* if it is, or possibly is, the individual essence of some individual. Then possible individuals can be replaced with their individual essences. In the reverse translation, the individual quantifier $\forall \mathbf{x}$ will be replaced with the quantifier $\forall \varphi$ over individual essences, and the predication $R'\mathbf{x}\mathbf{y}w$ will be replaced with the formula 'Possibly there is an individual x and possibly there is an individual y such that p strictly implies φ of x, ψ of y and Rxy' where p is the world-proposition corresponding to w, and φ and ψ are the individual essences corresponding to $\mathbf{x}$ and $\mathbf{y}$ respectively. The above difficulty is now avoided, for if p strictly implies $\exists x B(x)$ then there is an individual essence φ such that possibly for some x, p strictly implies φ of x and $B(x)$.

This solution is natural. It is also consonant with the treatment of possible worlds; for in both cases, an object is identified with the sum of its necessary properties. The use of this sum again requires certain existence assumptions. The first is that individual essences (generally, properties) necessarily exist. This assumption, like the corresponding one for propositions, will be criticized later. The second assumption is that necessarily each individual has an individual essence. It can be defended in the following way. Let x be any (actual) individual and let φ be the property of being identical to x (I use identity without existential presupposition, so that each individual is necessarily self-identical). Then necessarily for all y (φ holds of y if y is identical to x); and so φ is an individual essence of x. Now necessarily for each x there

---

[12] Prior mentions this difficulty in Prior and Fine [1977], ch. 6, p. 95.
[13] Cf. Chisholm [1967], 1–8.

is the corresponding property φ, and so the conclusion holds necessarily. This argument is, in a sense, purer than the corresponding one for world-propositions, for it requires neither a platonic stance on properties nor a resort to higher order entities.

3. Being Extensional

In the previous reverse translations, there is apparently a gain in ontological simplicity. For the possible worlds and individuals are eliminated in favour of propositions and properties. But there is also a loss in logical simplicity. For the quantifiers over propositions and properties are both second-order and intensional. It is not always appreciated that in modal logic the classical hierarchy of orders ramifies with respect to the extensional/intensional distinction. For example, at the second order there are extensional quantifiers over sets and intensional quantifiers over properties. Thus in the second modal language there is a double jump in logical complexity, one in order and the other from extension to intension.

This increase in complexity is not altogether necessary. For in the reverse translation, the intensional entities can be replaced by corresponding extensional entities, world-propositions by extensions and individual essences by singleton sets. Thus the intensional jump is avoided, though not the jump in logical order.

Let us consider the extensional substitute for individual essences in more detail. The basic idea is to replace the individual essence of a possible individual with the set whose sole member is that individual. Let us say, then, that X is a *singleton of x* if x is the sole member of X; and call X a *singleton* if it is possibly the singleton of some individual. In the reverse translation, the quantifier $\forall x$ will be replaced with a quantifier $\forall X$ over singletons and the predication $R'\mathbf{xy}w$ will be replaced with the formula 'Possibly there is an x and possibly there is a y such that p implies $x \, \varepsilon \, X$, $y \, \varepsilon \, Y$ and Rxy', where p is the world-proposition corresponding to w, and X and Y are the singletons corresponding to x and y respectively.

The extensional method for possible worlds is a little more complicated. In this case, the basic idea is to identify each possible world with the extension it confers on each of the predicates $R'_1, \ldots, R'_n$ in the classical language. (The method does not work for an infinity of predicates, although it is applicable to any particular formula.) Say that X is the *extension of* the binary predicate R if the true world-proposition p is such that necessarily for all x and necessarily for all y, p implies $\langle x, y \rangle \varepsilon \, X$ iff Rxy. There is a similar definition for predicates of arbitrary degree. (These definitions give the *outer* extension of a predicate, the set of possibles of which the predicate is true. It is more usual to consider the *inner* extension, the set of actuals of which the predicate is true.) Let an *extension-sequence S* be an n-tple $\langle X_1, \ldots, X_n \rangle$

such that it is possible that X_i is the extension of R_i for $i = 1, \ldots, n$. Then in the reverse translation, the quantifier $\forall w$ is replaced with the quantifier $\forall S$ over all extension sequences and 'p strictly implies $R_i xy$', in the formula that translates predications, is replaced with '$\langle x, y \rangle \varepsilon$ the i-th component X_i of S', where $S = \langle X_1, \ldots, X_n \rangle$ is the extension-sequence corresponding to w.

The method of substituting extensional entities does not avoid the need for existence assumptions. The use of singletons requires that singletons (and generally, sets) necessarily exist and that each individual determine a singleton. The second assumption follows from ordinary set-theoretic principles, but the first assumption can, and will, be questioned. The use of extensions requires that extensions necessarily exist and that each predicate determines an extension. In this case, both assumptions are questionable since ordinary set-theoretic principles only guarantee the existence of an inner extension (defined over actuals) and not of an outer extension (defined over possibles).

Prior used world-propositions for possible worlds and it is interesting to speculate on why he chose these intensional entities rather than some extensional counterpart. One reason may be that he was prepared to countenance non-nominal quantifiers, i.e. quantifiers whose variables do not occupy the same position as names. Given this stance, the simplest extensions of first-order or sentential modal logic are obtained by adding sentence or predicate quantifiers. And the most natural interpretation of these quantifiers—on logical and semantical grounds—is intensional. The logical ground is that only an intensional interpretation will guarantee the validity of Specification ($(\forall p)A(p) \supset A(B)$ in the case of sentences), and the semantical ground is that the free variable p should be interpreted in the same way as a sentence, namely by assigning an intension. My own view is that all *primitive* quantifiers are nominal and that therefore all non-nominal quantifiers should be analysed away in terms of nominal quantifiers and perhaps other notions too. For me, then, there is not the above reason for preferring an intensional account of possible worlds.

4.   To Be or Not To Be

The previous reverse translations depended upon certain existence assumptions. There were the comprehension-type principles, that the appropriate world-propositions, individual essences, singletons, and extensions exist. These principles were seen to be defensible, although a doubt was recorded for the case of extensions. Then there were the Barcan-type principles, that propositions, properties, and sets necessarily exist. A more accurate formulation of these principles is that necessarily the respective entities necessarily exist. The principles, on the original formulation, would be vacuously true if there were no propositions etc. But this is a subtlety that need not concern us.

The time has come to criticize the second group of assumptions. Any such criticism will have far-reaching consequences, but our main concern will be with those consequences that affect the original programme of reduction. Consider first the case of sets. It is natural to suppose that sets are abstract and that therefore they necessarily exist. But take the set whose sole member is Socrates. Then surely it is necessary that it exists only if Socrates does. Since Socrates enjoys contingent existence, so does the set.

In the above argument, Socrates was only an example; any other contingent existent would have done instead. So what the argument establishes is that if some individual contingently exists then so does some set, i.e. that if all sets necessarily exist then so do all individuals. Since this argument is so critical, it may be worth setting down in formal detail. Let X range over sets and x over sets and individuals, and use E for the existence-predicate and ε for membership. The argument then goes as follows:

(1) $\forall X \square EX$ (supposition)
(2) $\forall x \exists X(x \ \varepsilon \ X)$ (by set theory)
(3) $\forall x \exists X(x \ \varepsilon \ X \wedge \square EX)$ (from 1 and 3 by modal logic)
(4) $\forall x \forall X(x \ \varepsilon \ X \supset \square(EX \supset Ex))$ (assumption)
(5) $\forall x \forall X((x \ \varepsilon \ X \wedge \square EX) \supset \square Ex)$ (from 4 by modal logic)
(6) $\forall x \square Ex$ (from 3 and 5 by classical logic).

The crucial assumption of the argument is (4), which says that if an object belongs to a set then necessarily the object exists if the set does. (4) follows by modal logic from two further assumptions:

(7) $\forall x \forall X(x \ \varepsilon \ X \supset \square(Ex \supset x \ \varepsilon \ X))$
(8) $\forall x \forall X(\square(x \ \varepsilon \ X \wedge EX \supset Ex))$

(7) is a rigidity condition; it says that a member of a set is necessarily a member—at least if the set exists. Thus the members of a set are preserved through possible worlds; they are essential to its identity. (8) says that an existent set cannot have non-existent members. It follows from the conception of sets as constructions; the set cannot be constructed if the members do not exist. The consequences of denying (8) are very strange. For example, Extensionality would fail since two sets might have the same existent members but different non-existent members. A condition would then no longer define a unique set and other means, presumably modal, would be required to define a set. Thus I am inclined to regard both (7) and (8) as correct essentialist principles concerning the nature of sets.

The admission of contingent sets opens up the field of modal set theory; it is no longer an uninteresting and trivial extension of the classical theory. This new field calls for a more careful and detailed treatment than I can give here, but let me consider the question of providing a criterion for the existence of sets. In possible worlds terms, a necessary and sufficient

condition that a set exist in a world is that each of its members exist in that world. Given the Fundierungs-axiom, this is equivalent to saying that a set exists in a world iff each individual in its transitive closure exists in that world. (The transitive closure of a set consists of the members of the set, the members of the members, and so on.) Thus the existence of sets is ultimately traced back to the existence of individuals.

It is more difficult to state a criterion in modal terms. The necessity part of the above criterion is a strong form of (8), namely

$$\Box \forall X \Box \forall x \Box (x \; \varepsilon \; X \land EX \supset Ex).$$

However, the sufficiency part cannot be similarly expressed since no first-order axioms will guarantee that the domain of sets in a world is the powerset of the domain of individuals. Suppose one assumes this and a strong form of rigidity, namely:

$$\Box \forall x \Box \forall X \Box (x \; \varepsilon \; X \supset \Box (x \; \varepsilon \; X)).$$

Then the sufficiency condition is equivalent to the following modal form of Extensionality:

$$\Box \forall X \Box \forall Y (\Box \forall x (x \; \varepsilon \; X \equiv x \; \varepsilon \; Y) \supset X = Y).$$

Some of the above principles may strike the reader as unnecessarily subtle. But it is important to remember that the classical principles are, from the modal point of view, very weak. For example, the necessity of Extensionality merely says that no possible world contains distinct sets whose members, in that possible world, are the same. It says nothing about the membership of sets through possible worlds and so is compatible with sets being like boxes and changing their members from world to world. To state the identity conditions for sets in a satisfactorily strong metaphysical sense, one needs something like the above essentialist principles.

The objections to necessary existence, in the case of propositions, are more problematic. Reconsider the argument (1)–(6), but with X ranging over propositions, x ranging over objects, and 'ε' interpreted as 'occurs in'. Assumption (4) now says that a proposition cannot exist unless any object that occurs in it also exists. This may not be contentious. Unfortunately, the truth of (2)—that each object occurs in a proposition—is no longer clear.

First of all, the sense of 'occurs in' and of 'proposition' is obscure. If propositions are sentences and objects occur in sentences if their names do, then (4) is obviously correct but world-propositions will not, in general, exist and the reduction will not go through. In the context of modal logic and for the purposes of the reduction, it is most natural to say that propositions are identical if they are necessarily equivalent. One can then say that any object occurs in a proposition if the proposition cannot be expressed, in an ideal language, without reference to the object. An ideal language, here, is

one that is not subject to finitary constraints; the alphabet can be of any infinite size and the logical combinations (conjunction, quantification, etc.) can be of any infinite length.

However, (2) is not true under this natural interpretation of propositional identity. It is tempting to suppose that any object x occurs in the proposition to the effect that x exists. But this is because the proposition is presented by means of a reference to the object. It may well be that the same proposition can be expressed without any reference to the object. Suppose, for example, that x is a broom. Then the proposition that the broom exists can be expressed by the sentence 'the brush and handle are attached' or its like, and so reference to the broom itself can be avoided.

There is, then, a fundamental difference between sets and propositions. For propositions, there is a distinction between superficial and deep structure. The objects that appear to be in the proposition may disappear, as it were, on analysis. But for sets, there is no such distinction. The structure of a set is not affected by the internal complexity of its members.

Despite this difference, it does seem reasonable to maintain that *some* contingent individual occurs in a proposition, i.e. that $-\Box Ex \land \exists X (x \, \varepsilon \, X)$ holds for some individual x under the earlier propositional interpretation of 'ε'. The argument (1)–(6), with '$\forall x$' dropped in (2)–(6), would then establish $\Box Ex$. This contradicts the above and so gives $-\forall X \Box EX$ by reduction. The existence of the individual x could be denied, but only by claiming that each proposition can be expressed without reference to any particular contingent individual. Although some metaphysical views imply this claim, it does seem preferable to make our reduction independent of such views.

A criterion of propositional existence can be given, although the formal details are beyond the scope of this chapter.[14] There is a linguistic criterion that is in analogy to the definition of occurrence; it states that a proposition exists in a world if it can be expressed without any reference to possible individuals that do not exist in the world. There is also an equivalent non-linguistic criterion. Say that two worlds are *indistinguishable with respect to a* given world if they are identical but for the identity of their merely possible individuals, i.e. if there is an isomorphism from the domain of the one world on to the domain of the other that keeps fixed those individuals of the given world that exist in either of the other worlds. Then the criterion is that a proposition exists in a world iff the proposition is true in neither or both of any two worlds whenever those worlds are indistinguishable with respect to the given world; the proposition must not distinguish between indistinguishable worlds.

The reduction is principally concerned with the existence of world-propositions; and under certain circumstances it is possible to develop

---

[14] See Fine [1977*b*].

special criteria for the existence of world-propositions. Suppose, for example, that for any possible world there is an isomorphic world which has a disjoint domain of individuals. Thus in this set-up, the individuals are bare; they have no individual essence that can be expressed without reference to the given individual. Then under these circumstances, a world-proposition for v exists in w iff each individual that exists in v also exists in w. In particular, a world-proposition for v exists in the actual world only if each individual of v actually exists. In thinking of world-propositions, it is often helpful to bear the simple set-up above in mind.

There is no need to consider a separate argument for the contingent existence of some properties. For if the property of being identical to x exists then so does the proposition to the effect that x exists. Indeed, for any proposition there is a property whose existence-conditions are the same; if the proposition is q then the property can be the one that holds of o alone if q is true and holds of 1 alone if q is false. The existence criteria for properties can also be developed in analogy to those for propositions.

Although some sets, propositions, and properties contingently exist, there are large classes of these entities that necessarily exist. Call a set *pure* if it is built up from the null set (its transitive closure contains no non-sets) and say that a proposition or property is *purely general* if it can be expressed without reference to any individuals. Then the pure sets and the purely general propositions and properties all necessarily exist. The belief in the necessary existence of sets etc. may arise from an exclusive preoccupation with these special subclasses. One forgets that concrete objects may occur in the construction of an abstract object. In the case of propositions and properties, the preoccupation is further reinforced by the view that intensional entities are meanings that express contingent facts but exist independently of those facts.

5.   Trying Again

Let us consider how the reverse translations should be modified in the light of the previous criticisms. We consider first the reduction of possible individuals and then the reduction of possible worlds. The reductions of possible individuals to individual essences and to singletons both possess unnecessary logical complexity. For the possibilist quantifier $\exists\mathbf{x}$ (= 'there is a possible individual $\mathbf{x}$') can simply be replaced with $\Diamond\exists x$ (= 'possibly there is an actual individual x') or, alternatively, quantification over all possibles $(\forall\mathbf{x})$ can be replaced with necessary quantification over all actuals $(\Box\forall x)$. Thus the external quantifier over possibles is converted into an internal quantifier over actuals. The translation of predications is also simplified, for $R'\mathbf{x}\mathbf{y}w$ becomes 'q strictly implies Rxy', and similarly for the extensional account of possible worlds.

On the face of it, this method only provides a partial elimination of the possibilist quantifier. For $\exists \mathbf{x} B(\mathbf{x})$ says that some possible individual actually has B, while $\Diamond \exists x B(x)$ says that some possible individual has B in a world in which it exists; and so the two formulas may differ in truth-value if no possible individual that has $B(x)$ is actual. Say that a condition $B(x)$ is *strongly rigid* if it is true of a possible individual in every possible world, i.e. if $\Box \forall x \Box (B(x) \supset \Box B(x))$ holds. Then it should be clear that the equivalence of $\exists \mathbf{x}$ and $\Diamond \exists x$ holds when both apply to a strongly rigid condition. But the reverse translations of ω-free classical formulas are always strongly rigid; this merely generalizes the corresponding observation made for the case of a non-quantificational modal language. Therefore, within the context of the translation, the equivalence will always hold. The special properties of the reverse translation enable the partial elimination of $\exists \mathbf{x}$ by $\Diamond \exists x$ to be transformed into a complete elimination of possibilist quantifiers.

This method of elimination differs in a fundamental respect from the earlier reductions. They were *syntactically conservative* in the sense that the syntactic form of the classical sentences was preserved under translation. '$\forall x$' became '$\forall$ individual essences φ', and '$\forall w$' became '$\forall$ world-propositions q' or '$\forall$ extension-sequences Y'. On the other hand, the above method is *syntactically radical*. Syntactic form is not preserved, for the quantifier $\forall x$ is replaced with the hybrid form $\Box \forall x$.

The conservative/radical distinction has a quite general ontological significance. A conservative reduction enables one to identify certain objects in the reduced language with objects (the 'logical constructions') in the reducing language—individuals with essences, possible worlds with extension-sequences, reals with sets of rationals. On the other hand, a radical reduction does not, in general, enable such an identification to be made. Thus Russell's theory of descriptions does not give a construction that corresponds to each description and, similarly, rendering $\exists \mathbf{x}$ as $\Diamond \exists x$ does not enable one to say what possible individuals are.

Similar remarks will apply to the revised reductions for possible worlds: the quantifier $\forall w$ will always be replaced with an expression in which $\Box$ governs an actualist quantifier $\forall a$ (the range of 'a' will depend upon the reduction used). Possible objects have no exact identity for the modal actualist. Before, he was able to treat $\Diamond A$ as if it were existential and $\Diamond \exists x \forall$ as if it were doubly existential. But this pretence can no longer be kept up. $\Diamond A$, when both translations are applied, will become a sentence of the form $\Diamond \exists a A$ and so the original modal form will be preserved. Once a possibility statement, then, always a possibility statement.

The same method of pseudo-quantifiers could be applied directly to the extensional and intensional accounts of possible individuals; *necessary* quantification over all singletons or all individual essences would then replace the non-modal quantifiers. However, this approach would merely

re-create the earlier logical complexity. What motivated the original quan-
tification over singletons and essences was the belief that these entities
enjoyed necessary existence. But without this belief there is no longer any
reason for making the logical climb to another order.

So let us reconsider the reduction of possible worlds, first on the exten-
sional account and then on the intensional account. Before the reverse
translation can be modified, it is necessary to be clearer about the behaviour
of the membership relation ε. The question is: What truth-value, if any,
should $a\,\varepsilon\,b$ have in a world in which one of a and b denotes an individual
that does not exist? A more general question is: What truth-value, if any,
should an atomic sentence have in a world in which one of the possible
individuals named does not exist?

A general principle that helps resolve the above question is this: the truth-
value of any formula, modal or non-modal, should not turn on the
truth-value, if any, that is assigned to an atomic sentence (with a primitive
predicate) in which one of the objects named does not exist. Say that a
formula is *normal* if its truth-value remains fixed when the outer extensions
of the predicates vary but the inner extensions remain the same. Then the
principle states that each sentence should, on analysis, be equivalent to a
normal sentence. This principle might be called the Indifference Principle.[15] It
can be regarded as a consequence of Actualism. For it states that there are no
genuine relations among non-existents or among non-existents and existents,
that any relation which appears to discriminate among non-existents must be
reducible to those that do not.

There are various ways in which this principle might be secured. One is to
adopt the Falsehood Convention. This is the convention that any atomic
sentence should be false in case of empty reference. In modal terms, this
means that $\Box\forall_1 x_1 \ldots \Box\forall x_n (Rx_1 \ldots x_n \supset Ex_1 \wedge \ldots \wedge Ex_n)$ should be laid
down as an axiom for each n-place primitive predicate R. Given these
axioms, it is easy to show that any sentence is equivalent to a normal
sentence. For if x is the sole variable in B, then $\Box B$ may be replaced with
$\Box(Ex \supset B) \wedge (-Ex \supset B^f)$, where B^f is the result of replacing each atomic
sentence containing x in B by $\bot$, the falsehood constant. There is a similar
replacement in case B contains several variables. The result of the replace-
ments is then an equivalent normal formula. One could, of course, equally
well adopt a Truth Convention. But the Falsehood Convention seems to be
more natural since one already understands most predicates as being false of
non-existents. Another way of securing the Indifference Principle is to adopt
the Gap Convention, that atomic sentences should be undefined in case of
empty reference. One crucial difference between the Gap and Falsehood
Convention is that the former does not admit a primitive existence-predicate,

---

[15] This principle is discussed in more detail in Fine [1981*b*].

i.e. a predicate that is true of all existents and false of all non-existents in any given world. We shall, in the sequel, adopt the Falsehood Convention, although the Gap Convention will later be considered in a section on Q.

If the Falsehood Convention is adopted, then all predicates that do not conform to the convention must be defined in terms of predicates that do. This is relevant to the original question about the membership relation. For we may distinguish two senses of membership. There is *strong* membership—symbolized by ε′—that only holds of existents; and there is *weak* membership—symbolized by ε—that holds of **x** and **y** in a world whenever **x** belongs to **y**, regardless of whether **x** or **y** exist in the world. Now the Falsehood Convention requires that the strong should prevail. ε′, rather than ε, should be taken as primitive, and ε should be defined in terms of ε′ and perhaps other notions too, as long as they are false of non-existents. In fact, ε and ε′ are interdefinable, as the following definitions show:

$$x \; \varepsilon \; y \equiv \Diamond(x \; \varepsilon' \; y)$$
$$x \; \varepsilon' \; y \equiv (x \; \varepsilon \; y \land Ex \land Ey).$$

(As definitions, these equivalences should hold for any possible individuals and any possible world, i.e. they should be true when prefixed with $\Box \forall x \Box \forall y \Box$. The same goes for the formulas below.) The compatibility of these definitions depends upon the truth of:

$$x \; \varepsilon' \; y \equiv (\Diamond(x \; \varepsilon' \; y) \land Ex \land Ey).$$

Given that ε′ is only true of existents, this formula states that ε′ is rigid; its inner extension in each possible world is the same. The correctness of the first definition, i.e. that ε is weak membership, depends upon the assumption that each member of a set can coexist with the set. For without this assumption, a set would contain a member and yet in no possible world would both the set and the member exist.

In a similar way, one can distinguish between a strong and weak sense of identity. The strong sense—symbolized by =′—only holds of existents; the weak sense—symbolized by =′—holds of **x** and **y** in a world whenever **x** and **y** are identical, regardless of whether **x** or **y** exist in the world. Analogous interdefinability results hold: the definitions, compatibility, and correctness conditions are obtained from the earlier ones by substituting = for ε and =′ for ε′. In this case, the correctness condition is obviously satisfied and the compatibility condition follows from the appropriate rigidity assumption, namely:

$$x =' y \supset \Box(Ex \supset x =' y).$$

It is customary in modal logic to adopt = rather than =′ as primitive. This may be because the modal axioms then have a simpler formulation. Reflexivity, for example, requires no existential restriction. However, if the

preceding considerations are correct it is $='$, not $=$, that is primitive. This difference is small, but not insignificant. If $='$ is primitive, then $=$ cannot be defined in non-modal terms since the use of $\Diamond$ or its cognates cannot be avoided; $=$, unlike $='$, would then appear to be an essentially modal notion.

The modified extensional reduction of possible worlds can now be given. The predication $R'_i xyw$ is replaced, as before, with '$\langle x, y \rangle \varepsilon$ the ith component of S' or, alternatively, with '$\langle x, y \rangle \varepsilon X_i$'. '$\varepsilon$' is now the defined relation of weak membership. The quantifier $\forall w$ is replaced with $\Box \forall S$, where S ranges over extension-sequences, or, alternatively, $\forall w B$ is replaced with $\Box \forall X_1 \ldots \Box \forall X_n \Box$ ((X_1 is the extension of $R_1 \wedge \ldots \wedge X_n$ is the extension of R_n) $\supset B^*$). In this way, one avoids the problem of putting all of one's extension-sequences in one worldly basket. For the quantifier $\Box \forall X_1 \ldots \Box \forall X_n$ will, in effect, run through all sequences of *possible* extensions.

The revisions do not depend upon the assumption that extensions necessarily exist. However, they still depend upon the assumption that the outer extension of a predicate in a world exists in some (possibly distinct) world. Now although each member of the outer extension exists in some world, there is no reason to suppose that all the members can be found in a single world. Therefore, by our earlier criterion of set existence, there is no reason to suppose that the extension exists in any world. To take an extreme example, the extension in any world of $x = x$ is the set of all possible individuals, but in no world may all possible individuals be actual.

There are two solutions to this problem. One is to exclude from the language any predicate that is true of non-existents. The outer extension of each predicate will then coincide with its inner extension and the above difficulty will not arise. If the Falsehood Convention holds, then each primitive predicate will satisfy the restriction and so any language, when analysed, will submit to such a reduction. The other solution is to introduce a special quantifier PX over quasi-sets. Details will be given in the last part of §6.

Finally, let us reconsider the intensional reduction of possible worlds. It is again necessary to be clear about the primitives of the secondary modal language. We distinguish between sentences (such as 'Socrates is a philosopher') and names of propositions (such as 'the proposition that Socrates is a philosopher' and 'the proposition Socrates last expressed'). The semantical difference is between an expression having a proposition as intension and as extension. Before the propositional variables were sentential; they occupied the same position as sentences. But now the propositional variables are nominal; they occupy the same position as names of propositions.

Three predicates will be used to form sentences from names or variables for propositions: the existence- and truth-predicates and a predicate for strict implication. In conformity with the Falsehood Convention, the

predicates for truth and strict implication will have the strong sense. Thus the strong truth-predicate—symbolized by T'—holds of a proposition ρ in world w iff ρ exists in w and ρ is true in w; and the strong implication predicate—symbolized by →'—holds of propositions ρ and σ in world w iff ρ and σ exist in w and ρ implies σ. (If propositions are identified with sets of possible worlds, then ρ is true in w if w ε ρ and ρ implies σ if ρ ⊆ σ.) Apart from general philosophical considerations, one drawback to using sentential variables is that they presuppose the weak rather than the strong truth-predicate. For under the ordinary semantics, the sentential variable p, like any sentence, will be true in a world independently of whether the proposition expressed by p exists.[16]

The behaviour of these predicates within modal contexts is quite distinctive. Many commonly accepted principles do not hold once the contingent existence of propositions is allowed. Let ξ be the operator 'the proposition that'; thus ξ applies to a sentence to form a name of the proposition expressed by the sentence. Then the scheme $\Box(A \equiv T' \xi A)$, i.e. that A is necessarily equivalent to the proposition that A is true, is not, in general, correct. For let P be the sentence 'Socrates does not exist'. Then P is true in a world in which Socrates does not exist, but $T' \xi P$ is false since ξP does not exist. Indeed, in this case, the sentence $\Box(-P \equiv T' \xi P)$ is true. Strictly speaking, these conclusions hold only under certain assumptions about the existence of Socrates; they hold, for example, under the previously presented picture of bare individuals.

The rejection of this and related principles may appear strange. But it is perfectly natural once it is appreciated that $T'\xi A$ carries an existential commitment to the proposition that A. Since A need not carry this commitment, it can be true without $T'\xi A$ being true. If the commitment is made explicit, then the principles will stand. For example, the modified principle:

$$\Box(E \xi A \supset (T \xi A \equiv A))$$

is correct.

The existence-predicate enjoys peculiarities of its own. What is important to remember is that propositions may enjoy the same vicissitudes of existence as ordinary individuals. For example, take the principle of possible coexistence. This says that any two possible individuals both exist in some possible world or, to put it in modal terms, that $\Box\forall x\Box\forall y\Diamond(Ex \wedge Ey)$ holds. Now this principle is not especially plausible. But neither is the corresponding principle for propositions, that $\Box\forall\rho\Box\forall\sigma(E\rho \wedge E\sigma)$ holds. Indeed, on the simple picture of bare individuals, the two forms of the principle are equivalence, since then the proposition that x exists will exist iff x does.

---

[16] At the end of Fine [1970] sentential variables are combined with a variable domain of propositions. But I now consider this an incorrect approach.

If such existential failures are allowed, then many definitions of standard notions in the theory of propositions must be formulated with great care. Let us illustrate with the definitions of negation, disjunction, necessity and possibility. That ρ is the negation of σ —ρ neg$'\sigma$—can be defined by

$$E\rho \wedge E\sigma \wedge (\forall\tau)(\rho \rightarrow' \tau \wedge \sigma \rightarrow' \tau \supset \Box T'\tau).$$

But this defines strong negation. *Weak negation*—ρ neg σ—can be defined by ($\Diamond\rho$ neg$'$ σ). (The correctness of this definition depends upon the quite reasonable assumption that a proposition and its negation can always coexist.) Given certain facts about neg, it is then quite easy to justify the use of a negation operation '*n*' on propositions. That ρ is the *disjunction* of σ and τ—ρ dis$'$ σ, τ—can be defined by $E\rho \wedge E\sigma \wedge E\tau \wedge \forall\mu((\sigma \rightarrow' \mu \wedge \tau \rightarrow' \mu) \equiv \rho \rightarrow' \mu)$. However, weak disjunction—ρ dis σ, τ—cannot be defined as $\Diamond(\rho$ dis$'\sigma$, $\tau)$, for there may be no world in which both σ and τ exist. Consequently, there is no justification of a disjunction-operation '*d*' in analogy to that for '*n*'. The *strong necessity predicate* on propositions—$N\rho$—can be defined by $\Box T'\rho$. Note that this is also the weak predicate, since necessary propositions necessarily exist. However, the *weak possibility predicate*—$P\rho$—cannot be defined by $\Diamond T'\rho$, for this definition requires that ρ both be true and exist in a possible world. Rather, the correct definition is $-\Box T'n(\rho)$.

These peculiarities would be avoided if implication or truth were taken in the weak sense, as free from an existential commitment to propositions. However, if the Falsehood Convention is accepted, these predicates must somehow be defined in terms of other notions. Our experience with identity and membership might suggest that it is an easy matter to define the weak predicates from the strong. But the success of the earlier definitions depended upon certain rigidity and existence conditions being satisfied. The truth-predicate is not rigid, it is a contingent predicate *par excellence*, and so the compatibility conditions will not even be satisfied. The implication-predicate is rigid (on existents), but the correctness conditions need not be satisfied. A proposition and one of its implications may fail to coexist and $\Diamond(\rho \rightarrow' \sigma)$ would then be false of that pair of propositions. One might attempt to define $\rightarrow$ as necessity of the material implication, so that $(\rho \rightarrow \sigma) \equiv Nd(n(\rho),\sigma)$; but, as we have already noted, there is no justification for this use of '*d*'. It is, in fact, very difficult to define the weak notions in terms of the strong, and so we will follow the easier course of giving the reduction directly in terms of the strong predicates.

In giving the reverse translation, the definition of world-proposition must be modified to suit the nominal status of propositions. Thus ρ is a *world-proposition*—in symbols, $Q\rho$—if $\Diamond(T'\rho \wedge \forall\sigma(T'\sigma \supset \rho \rightarrow' \sigma))$ holds. The reverse translation can now proceed as follows: the predication $Rxyw$ is replaced with $\Box(\rho \supset Rxy)$; $\forall x$ is replaced with $\Box\forall x$; and $\forall w$ is replaced with $\Box\forall\rho$, where ρ is a quantifier over world-propositions.

That Q defines the correct notion requires a certain amount of justification. After all, might not a world contain no true proposition that distinguishes it from another world? This possibility can be ruled out on the basis of a metaphysical postulate. It says that each world can be described in terms of its actuals, i.e. that each world possesses a complete description which only makes reference to the individuals existing in that world. Thus, for a logical atomist, the description might be the conjunction of all atomic sentences of the form $Ra_1 \ldots a_n$, where $a_1, \ldots, a_n$ name actuals, the negations of such atomic sentences, and perhaps also the statement $\forall x(x = a \vee x = b \vee \ldots)$, where the list a, b, … includes names for all actuals. For him, then, no two worlds could agree on the actuals they contain and on the relations that held among them and yet differ on the relations involving non-actuals. This postulate would seem to be a *sine qua non* of any reasonable actualist position. But once it is granted, the correctness of the definition of Q then follows. For the complete actualist description will express a proposition that both exists and is true in the given world.

Prior defines a world-proposition as one that is *maximally possible*;[17] it is possible and implies any proposition or its negation—in symbols, $P\rho \wedge \forall\sigma(\rho \rightarrow' \sigma \vee \rho \rightarrow' n(\sigma))$. The two definitions are equivalent if propositions necessarily exist. But not otherwise. Every world-proposition p is maximally possible or, at least, possibly maximally possible. For in the world in which ρ is true any proposition or its negation is true. However, not every maximally possible proposition need be a world-proposition. Pick a world v in which Socrates does not exist and form the conjunction of all propositions in v that are true of the actual world. Then the proposition expressed by the conjunction is maximally possible (in v), but it may not be a world-proposition. For although it implies that there is an individual that has the properties in fact possessed by Socrates, it may not imply that any particular individual has those properties. But any world-proposition must imply an instance of any existential proposition that it implies.

The use of the expression $\Box\forall\rho$ overcomes the problem of contingent propositional existence in the same way that the earlier use of $\Box\forall x$ and $\Box\forall X$ overcame the problems of contingent individual and set existence respectively. In effect, $\Box\forall\rho$ searches through each possible world for world-propositions. Since each world will, at least, contain its own world-proposition, all world-propositions will eventually be found.

It is interesting to consider the result of using $\forall\rho$ instead of $\Box\forall\rho$ in the original reduction. The use of $\Box\forall\rho$ was based on the equivalence $\Box A \equiv \Box\forall\rho(Q\rho \supset \Box(\rho \supset A))$. If the second '$\Box$' is dropped, we obtain the definition of a new operator $\boxdot$ as follows: $\boxdot A \equiv \forall\rho(Q\rho \supset \Box(\rho \supset A))$. Now the right-hand side is true in a given world iff A is true in all

world-propositions that exist in the world. Say that v is *accessible from w* if the world-proposition for v exists in w. Then $\square A$ is true in a world iff A is true in all accessible worlds.

This definition is of technical and philosophical interest. The technical interest is that it enables a weak modal operator with an accessibility semantics to be embedded in an extension of S5. The accessibility relation is reflexive, for each world contains its own world-proposition. But there is no reason why it should be transitive or symmetric. In the set-up of bare individuals, v will be accessible from w iff each individual in v is in w. In this case, the accessibility relation is also transitive and if certain reasonable assumptions are made about the distribution of individuals through possible worlds then $\square$ will have exactly the logic of S±4.

The philosophical interest is in relation to the thesis that necessity is truth in all possible worlds. It is often claimed that this thesis leads to the modal logic S5. But the above definition of $\square$ is an instance of this thesis if $\square(\rho \supset A)$ is interpreted as 'true in', and, as we have seen, the definition may lead to a logic weaker than S5. The explanation of this anomaly is that the thesis about necessity only holds when the domain of possible worlds is taken to be fixed. If the domain of possible worlds is allowed to vary from world to world, then a statement might be true in all actual possible-worlds, say, and yet not true in one of the possible-worlds of an actual possible-world. The Law $\square A \supset \square\square A$ would then fail.

6. Refinements and Extensions

The previous sections gave a modal reduction of a classical first-order language for possible worlds and individuals. The present section shows how the principles of the reduction can be applied to languages that are both weaker and stronger than the original classical language.

Identity for worlds. The classical language contains a relative identity predicate I' if the original modal language contains an identity predicate I. The ordinary identity predicate $=$ can be defined in terms of I' for $\mathbf{x} = \mathbf{y}$ iff $\exists$ $w I' \mathbf{x} \mathbf{y} w$. However, the classical language contains no identity predicate on possible worlds or, indeed, any predicate that is defined exclusively on possible worlds. Therefore it is natural to add such a predicate to the classical language even though it is not required for the classical translation.

In translating back from this enriched language, there are no difficulties when possible worlds are reduced to world-propositions. For then $w = v$ can be replaced with $\square(p \equiv q)$ if sentential variables are used and with $\rho = \sigma$ (or $\square(T'\rho \equiv T'\sigma)$) if propositional variables are used. However, there are difficulties when possible worlds are reduced to extensions. For, in this case, two worlds are identified if they confer the same extension on each predicate of the language, and so there is no way of saying that *distinct* worlds confer the

same extension on the different predicate. If the language is such that this last possibility does not arise, then world identity does not call for special treatment. It may be defined in the classical language by conjoining formulas of the form $\forall \mathbf{x} \forall \mathbf{y} (R'\mathbf{xy}w \equiv R'\mathbf{xy}v)$.

This difference in the reductions shows that the intensional account is more basic than the extensional. The latter account works only when there are certain restrictions or conditions on the classical language. Once they are lifted, the intensional account must be used. This may be another reason why Prior did not consider an extensional account.

No sorts. The classical language is two-sorted. It contains a sort for possible worlds and a sort for possible individuals. Any many-sorted language can be reduced to a one-sorted language by collapsing all the sorts into a universal sort and then recovering the original sorts by means of sortal predicates. In the present case, the universal sort would be for possible objects (= possible worlds or possible individuals) and the sortal predicates would be for being a possible individual and for being a possible world.

It is sometimes possible to translate the de-sorted language back into the original many-sorted language. Suppose, for example, that crossing sorts always results in falsehood, i.e. that a predicate in the one-sorted language is false of objects $a_1, \ldots, a_n$ if one of them is not of the correct sort. Then a backward translation can be obtained by first replacing quantification $\forall x A(x)$ over the universal sort with $\forall x^1 A(x^1) \wedge \ldots \wedge \forall x^m A(x^m)$, where $x^1, \ldots, x^m$ are variables for each different sort, and then replacing predications $R y_1 \ldots y_n$ with $\bot$, the falsehood constant, if the variables $y_1, \ldots, y_n$ are not properly sorted. In the present case, the backward translation is possible if possible worlds and possible individuals are distinct and if R' is true of possible objects $a_1, \ldots, a_n, a_{n+1}$ only when $a_1, \ldots, a_n$ are possible individuals and a_{n+1} is a possible world.

However, without special assumptions, the backward translation will not be possible. Consider the sentence that says something is of two sorts—for example, that some object is both a possible world and possible individual. Then, in the absence of special assumptions, no sentence can express this. The same difficulty arises if one adds to the sorted language a universal identity-predicate, i.e. one that does not discriminate between different sorts. For then the above sentence can be rendered as $\exists x^1 \exists x^2 (x^1 = x^2)$.

All of the earlier reverse translations were for a two-sorted classical language. But suppose that the language is de-sorted or enriched with identity and that each possible world is counted as a possible individual. How, then, can the reverse translation be given? One method is to construct a first-order modal language in which the first-order variables range over both worlds and individuals. Thus the domain D_w associated with each world w includes both worlds and individuals proper. There is no need for the set of all worlds to be included in any domain, but one can at least

require that $w \varepsilon D_w$, i.e. that each world belongs to its own domain. If earlier existence-conditions for world-propositions are carried over to worlds, then one could also require that $D_v \subseteq D_w$ implies $v \varepsilon D_w$, i.e. that any world contains a world constructed from its own individuals.

To this modal language is added an actuality predicate G. G is true of an object **x** in world w if (a) **x** is w and (b) **x** exists in w. Clause (b) is, of course, redundant if the condition $w \varepsilon D_w$ holds. Note that the classical correlate G' of G is the identity predicate on worlds. Thus if one accepts the principle that R is a logical predicate if R' is, then G is a logical predicate. This, in a sense, justifies the introduction of G into the modal language. The reduction now proceeds as follows: $R'\mathbf{xyz}$ is replaced with $\Box(Gz \supset Rxy)$ and $W\mathbf{x} =$ '**x** is a world' with $\Diamond Gx$; $\forall \mathbf{x}$ is replaced with $\Box \forall x$, as before; and $A(\omega)$ is replaced with $\exists x(Gx \wedge A^*(x))$, where $A^*(x)$ is the modal formula corresponding to $A(x)$.

The previous translations all involved a jump in expressive power: the classical language was stronger than the primary modal language; and the secondary modal language was, in its turn, stronger than the classical language. However, in this case, the modal and classical languages are on an expressive par. For a classical translation of the new modal language merely returns the original one-sorted classical language.

However, from the point of view of modal actualism, the new language is not above suspicion. It is not that it offends against actualism; the domain of each possible world need only contain the existing individuals and worlds. Rather, it goes against the spirit of modalism. Possible worlds quantifiers and modal operators are usually regarded as two competing means of expressing modal truths. The new language is a curious hybrid in which both means of expression are combined. Since the modalist does not need both modal operators and possible worlds, it is only natural for him to eliminate the latter along lines already indicated. The classical possibilist, too, may attempt to eliminate the modal operators in favour of possible worlds; but for him there may be special difficulties, since not all of the possible worlds may exist in the actual world.

Rigidity operators. Our main concern, so far, has been to translate a classical language of possible objects into a modal language. However, the same principles of translation may be used in translating one modal language into another. We now show that certain non-standard operators can be defined in terms of the standard ones. What distinguishes the non-standard operators is that a reference to the actual world is involved in their truth-conditions.

First we consider the case in which the operator $\downharpoonleft$ for 'it is actually the case that' is introduced into the language.[18] The semantics for this operator is

---

[18] The temporal counterpart 'Now' is considered in Kamp [1971].

that $\downarrow A$ is true at a possible world if A is true in the actual world. Thus if $A^*(w)$ corresponds to A, then $A^*(\omega)$ corresponds to $\downarrow A$. This operator increases the expressive power of ordinary first-order modal logic. For example, the sentence $\Diamond \forall x - \downarrow Ex$, saying that it is possible that all actual individuals do not exist, cannot be said without the use of $\downarrow$.

The language with $\downarrow$ can be translated into the classical language and so it can be translated into the secondary modal language. A direct translation can also be given. Thus if A is of the form $A(\downarrow B)$ it may be translated as $\exists\, p(p \wedge Qp \wedge A(\Box(p \supset B)))$ or as $\exists\, \rho(T'\rho \wedge Q\rho \wedge A(\Box(T'\rho \to B)))$ if nominal quantifiers for propositions are used. Similarly, if A contains several formulas of the form $\downarrow B$. In effect, $\downarrow B$ is replaced by 'The true world-proposition strictly implies B', where the description is given widest scope. The translation could also be based upon the extensional account of possible worlds. For example, the sentence $\Diamond \forall x(- \downarrow Ex)$ would then become $\exists X(\forall x(x \,\varepsilon\, X) \wedge \Diamond \forall x(- x \,\varepsilon\, X))$.

The behaviour of $\downarrow$ may, in some ways, appear anomalous. The formula $A \supset \Box \downarrow A$ is valid, but the necessary truth of $\downarrow A$ does not follow from the truth of A. However, this behaviour is merely the result of scope ambiguity. The conditional formula, upon translation, becomes $\exists p(p \wedge Qp \wedge (A \supset \Box\Box(p \supset A)))$; and the validity of this formula does not imply that $\exists p(p \wedge Qp \wedge \Box(p \supset A))$ is necessarily true if A is true. Admittedly, it is odd to regard a connective as being subject to scope ambiguity, but the use of the description 'the true world-proposition' shows how this ambiguity can be reduced to the more familiar case of descriptions.

Kaplan[19] so interprets $\downarrow$ that $\downarrow A$ is necessarily true if A is true. This interpretation has some extreme consequences: it leads, for example, to a complete breakdown in the correlation between a priori and necessary truths. What the above analysis shows is that the logical behaviour of $\downarrow$ can be respected and yet the extreme consequences avoided.

Vlach [1973] has introduced a companion † for $\downarrow$. The tense-logical readings for † and $\downarrow$ are 'once' and 'now' (or 'then'). †, however, has no modal reading. The semantics for † and $\downarrow$ can be explained in the following way. Usually, when evaluating a formula at a world, one has to evaluate its subformulas at different worlds. But once this is done, the reference to the original world is lost. Now † allows one to keep or 'store' a reference to the world of evaluation, while $\downarrow$ enables one to pick up this reference. Thus † and $\downarrow$ work in pairs, with $\downarrow$ picking up the reference fixed by †. In evaluating a formula, one must consequently keep track of two worlds: the 'floating' world at which ordinary formulas are evaluated and the (temporarily) 'fixed' point. The rules for † and $\downarrow$ are then: $\dagger A$ is true when the floating world is w if

---

[19] 'The Logic of Demonstratives' (unpublished at the time but subsequently published in considerably elaborated form in Kaplan [1989]).

A is true when the fixed world is *w* (keeping the floating world the same); and ↓*A* is true when the fixed world is *w* if *A* is true when the floating world is *w* (keeping the fixed world the same).

The above rules permit a translation of the language with † and ↓ into the classical language and thereby into the secondary modal language as well. Rather than give details, let us simply give the direct translation of † and ↓ into the secondary modal language. If *C* is a formula †*A*(↓*B*) that only contains the two designated occurrences of † and ↓, then its translation is $\exists p(p \wedge Qp \wedge A(\square(p \supset B)))$. This replacement can take place when *C* itself is part of a broader formula. Thus the effect of † is merely to limit the scope of the world-description implicit in ↓; † is, in terms of the preceding discussion, a marker of scope. Any adjacent pairs of † and ↓ can be eliminated in this way. If any occurrences of ↓ remain, they may be singly eliminated as above. (In effect, a sufficient number of occurrences of † are placed at the beginning of the formula). If any occurrences of † remain, they may be dropped since the references secured by † are never picked up by a subsequent occurrence of ↓.

Possibilist quantifiers. We now consider a modal language in which the quantifiers range over all possible individuals. Such a language is in conformity with modalism, but not with actualism, and so the possibilist quantifiers must be removed. We have already observed that $\Diamond \exists x A(x)$ is not, in general, a correct translation of $\exists x A(\mathbf{x})$. However, the translation may be repaired with † and ↓; $\exists x A(\mathbf{x})$ is equivalent to $\dagger \Diamond \exists x \downarrow A(x)$. For when we come to evaluate ↓*A*(*x*), ↓ takes us back to the point at which we evaluated $\dagger \Diamond \exists x \downarrow A(x)$. Since † and ↓ can be eliminated, this gives a direct translation: $\exists x A(x)$ goes into $\exists p(p \wedge Qp \wedge \Diamond \exists x \square(p \supset A(x)))$. A similar result could also be obtained, of course, by looking at the classical translation of the language with possibilist quantifiers.

It is sometimes claimed that certain English sentences require possibles for their analysis and that this constitutes an argument for the possibilist's position.[20] Such arguments are quite naive, since they presuppose that the logical form of the sentences will closely follow their grammatical form. It is true that some uses of the possibilist quantifier are not eliminable within a first-order actualist language. The formula $\square \exists \mathbf{x} P\mathbf{x}$ provides a particularly simple example. However, as the preceding translation shows, all the uses are eliminable once propositional quantifiers are added to the actualist language.

Higher-order logic. Prior's reduction of possibilist talk was, in one respect, over-modest. He was content to give a modal reduction of that language required for the classical translation of modal discourse. But the resulting modal language could also be translated into a classical language which, in

---

[20] Cf. Prior [1967], 149–51.

its turn, could be given a modal reduction, and so on *ad infinitum*. In general, the problem remains of reducing a higher-order classical language, with quantifiers over sets of possible worlds, sets of possible individuals, sets of such sets and so on.

In case propositions, properties and sets enjoy necessary existence, this problem has a simple solution. For possible worlds and individuals were identified with objects of a certain sort and so sets of the former entities could be identified with sets of the latter entities; for example, sets of possible worlds could be identified with sets of world-propositions. But without necessary existence, the reverse translation is syntactically radical and so there is no such automatic extension to higher-order logic. For example, what corresponds to the possible individuals quantifier is the modalized quantifier $\Box\forall x$, where x ranges over actual individuals. But what is to correspond to the quantifier over sets of possible individuals? Should it be $\Box\forall X$, where X ranges over (actual) sets of individuals? But then one would have to assume that each set of possible individuals exists in some possible world or, in terms of our criterion for set existence, that each set of possible individuals is included in the domain of some world. And this is not a reasonable assumption.

There may be better solutions. For example, sets of possible individuals may be replaced with properties and membership regarded as possible possession of the property. There will, however, be no general solution within the type of modal language so far considered. Suppose that all possible worlds are isomorphic and that their domains are finite but disjoint. Then in each possible world the domains of sets, properties, propositions, etc., will be extremely limited; and it would appear to be impossible to say various things—such as that there are an infinite number of possible individuals.

There are, as with the analogous problem over the extensional account of possible worlds, two responses to this difficulty. First, one might rule out those situations that give rise to this limitation in expressive power. It may be that no reasonable metaphysical view would tolerate the type of situation described above, but all the same it would be preferable to free the reduction from all but the weakest metaphysical presuppositions. The other response is that the higher-order extensions of a modal language have been too narrowly conceived and that some distinctively modal types of variable-binding operator need to be introduced. More specifically, we need a variable-binding operator P on quasi-classes X and a relation-symbol 'In' that holds between individuals and quasi-classes. $(PX)B$ is true in a world iff for some *arbitrary* class of possible individuals, B is true in each world; and 'x In X' is true in a world when the object assigned to x belongs to the class assigned to X.

Now the natural objection to this move is that PX is merely a possibilist quantifier over all sets and that therefore the new language is not in

conformity with actualist principles. But I am not sure that we need accept this objection. Let B^+ be the result of replacing 'x In X' in B by the infinite disjunction $x = x_1 \vee x = x_2 \vee \ldots$ for arbitrarily many variables $x_1, x_2, \ldots$ Then $(Px)B$ can be regarded as an abbreviation for the infinitary but actualist formula $\Diamond \exists x_1 \Diamond \exists x_2 \ldots \Box B^+$. For this latter formula says that there are possible individuals $\mathbf{x}_1, \mathbf{x}_2, \ldots$ (not necessarily all in one world) such that B^+ holds, and $x = x_1 \vee x = x_2 \vee \ldots$ holds of $\mathbf{x}$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots$ if 'x In X' holds of $\mathbf{x}$ and the class $\{\mathbf{x}_1, \mathbf{x}_2, \ldots\}$. In the same way, one might regard the quantifier $\exists X$ over classes of individuals as an abbreviation for $\exists x_1, \exists x_2, \ldots$, with $x \, \varepsilon \, X$ replacing $x = x_1 \vee x = x_2 \ldots$ In both these cases, a new notation is introduced in order to give finitary expression to a formula that is infinitary but subject to certain constraints. In the first case, the constraint is that the quantifiers be actualist and, in the second case, it is that the quantifiers be first-order.

The notation $\Diamond \exists x_1 \Diamond \exists x_2 \ldots \Box$ of the expansion is, I think, very natural within the context of an infinitary modal language with quantifiers. For suppose that we wish to give a complete description of all the modal facts, i.e. of the universe of possible worlds as opposed to any of its members. Then it is natural for the modal actualist to proceed in the following way. First, he gives a complete description $B_1, B_2, \ldots$ of each world as if he had names for all possible individuals. Second, he replaces those names by variables $x_1, x_2, \ldots$ to give the formulas $C_1, C_2, \ldots$ Finally, he gives the formula $A = \Diamond \exists x_1 \Diamond \exists x_2 \ldots \Box (x_1 \neq x_2 \wedge x_1 \neq x_3 \wedge \ldots \wedge \Diamond C_1 \wedge \Diamond C_2 \wedge \ldots$ $\Box (C_1 \vee C_2 \vee \ldots))$ as the complete description of the modal universe. The use of names in the first stage is, of course, a mere pretence used in defining the final formula A. For the modal actualist, there may be no specific individuals $a_1, a_2, \ldots$ in virtue of which the formula A is true.

The use of the notation Px (or $\forall X$) is justified in terms of the infinitary expansion, and it is necessary that any extension of the notation be justified in terms of a corresponding extension of the expansion. For example, it is easy to extend the notation to quasi-classes of ordered pairs (whose components can coexist), but there is no basis for allowing x to appear to the left of 'In' (or allowing X to appear to the left of 'ε'). Thus the uses of Px are severely limited. Intuitively speaking, one should always be able to regard a quasi-class as distributing over all possible worlds, as picking out possible individuals from each of the worlds in turn.

As long as this limited use of quasi-classes is accepted, one can resolve the above difficulty of expressing that there are an infinite number of possible individuals. For one may say that there is a quasi-set x of pairs $\langle n, \mathbf{x} \rangle$, n a natural number and $\mathbf{x}$ a possible individual, such that it defines a one–one function on the natural numbers. One can also solve the problem in §5 over the extensional account of possible worlds. For the quantifiers over extensions can be replaced with quantifiers over quasi-sets.

Indeed, it would appear that quasi-sets provide a general reduction of higher-order classical languages. For we may say that there exists a quasi-set χ of pairs $\langle \alpha, t \rangle$, α an ordinal and t a world-proposition or individual, that is a one–one function from a class of ordinals on to the class of world-propositions and individuals. In terms of this one–one correspondence, one may define relations R^* on ordinals that are isomorphic to the relations $\Box(T'\rho \supset Rxy)$ between world-propositions and individuals. One then has, in any possible world, a structure that is isomorphic to the one provided by the classical language, and so one may use any resources from the classical language in describing this structure.

7.   Q-ish difficulties

We single out for separate treatment the problem of reduction for Prior's system Q[21]. The system Q is an actualist modal logic. What distinguishes it from other such logics is its account of sentences which contain names for individuals not existing in a given world. All such sentences are said to be undefined or truth-valueless. This might be called the Gap Convention, in analogy to the Falsehood Convention which states that atomic sentences are false in case of empty reference.

The Gap Convention might be broken down into two parts: first, that atomic sentences are undefined in case of empty reference; and second, that gaps are preserved under the logical operations, be they truth-functional, modal or quantificational. It would then be possible to accept the first part of the convention but not the second (the second part is vacuous, of course, if there are no gaps). For example, gaps could be treated by the method of supervaluations or by some other set of three-valued truth-tables. However, such a differential approach goes against the whole spirit of Q. An empty name in any sentence, be it atomic or not, is a source of gaps. For in a world in which reference fails, the sentence is not 'statable', it states no fact of the world, and so cannot be given a truth-value. Segerberg [1970] introduces an operator T (taking a gap into false) in proving completeness for Q-type system. But if the above remarks are correct, such operators have no place in a proper presentation of the system Q.

The system is, in many respects, unnatural. Because of gaps, many standard modal principles are only valid under statability restrictions. Again, gaps create the need for distinguishing between a strong and weak sense of the modal operators. S is strongly necessary ($\boxed{s}$) if it is always true and weakly necessary ($\boxed{w}$) if never false; and similarly, S is strongly possible ($\diamondsuit$) if it is sometimes true and weakly possible ($\boxed{w}$) if sometimes not false. Perhaps odder than these two features are certain anomalies in what can be expressed

---

[21] First presented in Prior [1957].

and how. One would naturally express 'Possibly I do not exist' with the sentence $\Diamond - Ea$, where $\Diamond$ has the strong sense. But, in Q, E is undefined for non-existents and so $\Diamond - Ea$ is false. What $\Diamond - Ea$ expresses in our earlier modal language would be expressed in Q by $\diamondsuit\!\!\!\!\diamondsuit - Ea$. However, this is a kind of accident. For what is earlier expressed by $\Diamond (- Ea \wedge - Eb)$—e.g. 'Possibly I and you do not exist'—cannot be expressed in Q at all. Only in higher-order extensions of Q could this sentence be expressed. In first-order Q, there are definite limitations on what uses of E can be captured.

Prior was well aware of the first two oddities, though perhaps not the third. All the same, Prior was firmly convinced, on philosophical grounds, that Q was the 'correct' modal system. I do not think his arguments for this claim can be sustained, but a full discussion of the issue would take us too far afield. Let me merely discuss the issue in so far as it relates to the preceding account of propositional existence.

The question is whether a predicate can be true (or false) of non-existents. In particular, we may ask whether the existence-predicate can be false of a non-existent, or, to put it in modal terms, whether $\exists x \Diamond - Ex$ can be true for the strong sense of $\Diamond$. For the sake of an example, let S be the sentence 'Socrates does not exist'. Then Prior would argue:[22]

    (1) $\Diamond S$ is true only if the proposition that S is true in some possible world ($\Diamond S \supset \Diamond T' \xi S$).

    (2) The proposition ξS is true in a world only if it exists in that world ($\Box(T' \xi S \supset E \xi S)$).

    (3) The proposition ξS exists in a world only if Socrates exists in that world ($\Box(E \xi S \supset Ea)$, for a = 'Socrates').

    (4) The proposition ξS is true in a world only if Socrates does not exist in that world ($\Box(T' \xi S \supset - Ea)$).

    (5) From (1)–(4) it follows that $\Diamond S$ is false ($- \Diamond S$).

Now (4) is certainly correct, and (3) we are prepared to accept. (2) is also correct, at least for a strong sense of truth. (1), however, is not acceptable. For, as we have already remarked, $\Box(S \supset T' \xi S)$ is not true, and so there is no reason why the truth of $\Diamond S$ should imply the possible truth of ξS.[23]

The matter cannot, however, rest there. For something like (1) must be correct if the possible worlds semantics is to be applicable to modal languages. One solution to this problem is to adopt the principle that $\Diamond S$ is true if the sentence S is true in some possible world, and yet not require that when a sentence is true in a world then so is the proposition expressed. A related solution is to adopt (1) but for a weak sense of 'true in' or 'true' (T, not T'), so that a proposition can be true in a world without the proposition existing

---

22 [1967], 149–51.
23 This question is briefly discussed in Prior [1968], 128–31.

in that world. If either solution is adopted then the relevant senses of 'true in' must be made out. Now that S is true in a world might be expressed in terms of the truth of $\Box(T'\rho \supset S)$, where ρ is a world-proposition. The Q-theorist could, of course, rejoin that the truth of $\boxed{\text{s}}(T'\rho \supset S)$ implies the necessary existence of ξS. But as an argument *for* Q, this is circular. The most that the Q-theorist can argue is that Q is self-sustaining, that the relevant sense of 'true in' cannot be made out once a Q-ish language is adopted. But this is not to argue for Q on the basis of a neutral position in which only actualism and the legitimacy of the possible worlds semantics are taken for granted.

Even though I reject the arguments for Q, I do think that the system embodies correct principles and distinctions. For, first, Q is in conformity with the Indifference Principle, the principle that the truth-value of a formula should not turn upon relations involving non-existents. In Q, the principle is secured with the Gap Convention; whereas I should prefer to reconcile the principle with the absence of gaps. And, second, the distinctions arising from gaps can be re-expressed in terms of propositional existence. For example, $\boxed{\text{s}}A$ is, in our terms, $\Box T' \xi A$, whereas $\boxed{\text{w}}A$ is $\Box(E\,\xi A \supset T'\,\xi A)$.[24]

Let us now turn to the translations from and to Q. In the first direction, there are various classical translations and consequently various classical languages that might be used. First, the classical language might be taken to contain gaps that correspond to the gaps in the original modal language. Thus Rxy would go into $R'\mathbf{x}\mathbf{y}w$, where the second formula is truth-valueless whenever one of the individuals assigned to x or y did not belong to the world assigned to w; $\vee$ and $-$ would represent themselves, though gaps would need to be preserved in the classical language too; $\boxed{\text{s}}B$ would go into $\forall w\,TB^* \wedge SB^*$ and $\boxed{\text{w}}B$ into $\forall w - T - B^* \wedge SB^*$, where T is the truth-operator (true when the sentence is and false otherwise) and where SA abbreviates $A \vee -A$; and, finally, $\forall xB$ would go into $\forall \mathrm{x}(TE'\mathrm{x}w \supset TB^*) \wedge SB^*$. Second, to each modal formula A there might be taken to correspond two conditions $A^t(w)$ and $A^s(w)$ that give the truth and stability conditions of A respectively. $A^s(w)$ would be the conjunction of the formulas $E'\mathbf{x}w$ for $\mathbf{x}$ free in A and R^sw (meaning R is stable in w) for R a predicate of A. As for $A^t(w)$, Rxy would go into $R^t\mathbf{x}\mathbf{y}w$, where this is now a bivalent formula; $(-B)^t$ into $-(B^t)$; $(B\vee C)$ into $(B)^t v (C)^t$; $\boxed{\text{s}}B$ into $\forall w B^t(w)$; $\boxed{\text{w}}B$ into $\forall w(B^s(w) \supset B^t(w))$; and, finally, $\forall xB$ into $\forall \mathrm{x}(E'\mathrm{x}w \supset B^t)$. Finally, Q could be translated into *our* first-order modal language and thereby into the standard classical language. Everything but the modal operators would be untouched. If A contained only the variables x, y, z, then $\boxed{\text{s}}A$ would go into $\Box(Ex \wedge Ey \wedge Ez \wedge A)$ and $\boxed{\text{w}}A$ would go into $\Box(Ex \wedge Ey \wedge Ez \supset A)$. In this translation, unlike the others, it would need to be assumed that no variables (or, generally, names and variables) were

---

[24] Cf. the last paragraph of Prior [1969*b*].

implicit in the simple predicates. Thus a formula would be stable just when the explicitly named individuals existed.

The secondary modal language is obtained from the original one by adding sentential variables $\forall p$. However, propositions can no longer be regarded as sets of possible worlds; for a proposition may be true, false or non-existent (unstatable) in any given world. Instead, a proposition may be regarded as a *partial* function from the set of possible worlds into the two truth-values, True and False. The quantifier $\forall p$ will then range, in each world, over all propositions that are defined for that world. There is, in Q, no reason for letting the domain of propositions (i.e. partial functions) vary from world to world. For there are no propositions, like our earlier proposition that Socrates does not exist, which are true or false in a world in which they do not exist. The contingency of propositional existence is already captured, as it were, in the possible worlds functions being partial.

In Q, there are two kinds of world-propositions. On the one hand, there are those that are true in one world and false in all others. These are definable by $\diamondsuit(p \wedge \forall q((q \wedge \boxed{S}(q \supset q)) \supset \boxed{S}(p \supset q)))$ as long as they exist; but there is no reason, in general, to suppose that such world-propositions exist. On the other hand, there are the propositions that are true in one world and not true in all others. These propositions exist, but are not definable. The formula $\diamondsuit(p \wedge \forall q(q \supset \boxed{W}(p \supset q)))$ defines them under the assumption that any two worlds are sharply distinguishable, i.e. that there is a proposition true in one of them and false in the other. However, this assumption is questionable (consider two isomorphic worlds with disjoint domains of individuals). There seems to be no other definition that would avoid this difficulty; and so let us suppose that some assumption or new primitive enables the second world-propositions to be defined by a formula Qp. The need for these props might be taken as an objection, albeit of a technical and recondite sort, against the system Q.

A reverse translation can now be given for the various classical languages, although we shall only consider the last. First, $R'\mathbf{x}y w$ is replaced with $\diamondsuit(p \wedge Rxy)$ and $\forall wA$ with $\boxed{W}\forall p(Qp \supset A^*)$. Note that $\boxed{W}(p \supset Rxy)$ cannot be used for atomic sentences since it is true when Rxy is unstatable in the given world. Finally, if A is a formula whose sole free variables are x and y and A is the result of replacing each atomic formula containing a free occurrence of y in A^* by $\bot$, then $\forall xA$ is replaced with $\boxed{W}\forall xA^* \wedge \boxed{W}\forall p((Qp \wedge -\diamondsuit(p \wedge Ey)) \supset \diamondsuit(p \wedge \forall xA^f))$; there is a similar replacement when more variables occur freely in A. The first conjunct only says that A is true of all individuals x that coexist with y, and this explains the need for another conjunct. The replacement A^f in this conjunct is justified as long as the classical language conforms to the Falsehood Convention. The translations of $\forall w$ and $\forall x$ are only correct under the assumption that for any finite set of individuals and world-proposition there is a world in which the

individuals exist and the proposition is statable. Again, this assumption appears to be both questionable and unavoidable.

It is worth noting that the translation only uses Ⓦ and its companion operator $\Diamond\kern-1.6ex\text{—}\, = -\boxed{\text{w}}\,-$. This shows that Ⓢ (or Ⓜ) must be definable in terms of Ⓦ and sentential quantifiers. Suppose that A contains the free variable x alone. Then a direct translation of ⓈA is $\boxed{\text{w}}A \wedge \boxed{\text{w}}\forall p(Qp \supset \Diamond\kern-1.6ex\text{—}\,(p \wedge Ex))$. A similar translation can be given in case A contains several variables.

8.   Problems with time

The preceding sections have been entirely occupied with modal logic. Many of the considerations apply, *mutatis mutandis*, to tense logic and so a separate detailed account is not required. What we shall do is consider the special problems when tense logic is substituted for modal logic and when it is combined with modal logic.

The original tense-logical language is like the original modal language, but with the two tensed operators G (always will be) and H (always has been) in place of the single modal operator $\Box$. The corresponding classical language is still two-sorted, but now one sort is for individuals—past, present or future—and the other sort is for instants of time. We use $\mathbf{x}$, $\mathbf{y}$, $\mathbf{z}, \ldots$ as variables for the first sort and $t, u, v, \ldots$ as variables for the second. The language contains an earlier–later relation $<$ in addition to the predicates R' with an extra argument for instants. The classical translation is then as follows: Rxy goes into $R'\mathbf{xyt}$; the connectives $\vee$ and $-$ represent themselves; GA goes into $\forall u(t < u \supset A^*(u))$ and HB into $\forall u(u < t \supset B^*(u))$; and $\forall xB$ goes into $\forall x(E'xt \supset B^*(t))$.

Let us note that there are now two interpretations of E (or E'). On the first interpretations, only present individuals exist. The domains of individuals at different instants could then vary in size, although one might suppose that individuals don't exist after ceasing to exist, i.e. that $D_t \cap D_v \subseteq D_u$ for $t < u$ and $u < v$. On the second interpretation, past and present individuals exist. The domains of individuals are then cumulative, i.e. $D_t \subseteq D_u$ for $t < u$. E could, of course, be used for present existents alone and then the cumulative interpretation of the quantifiers would require the translation $\forall x(E'xt \vee \exists u(u < t \wedge E'xu) \supset B^*(t))$. I myself find the second form of tense-logical Actualism more plausible than the first. For certain causal facts seem to involve a 'genuine' relation between individuals at different instants.

For the reverse translation, it is necessary to define instant-propositions. Let us use LA for 'it is always the case that A' and $MA(= -L-A)$ for 'it is sometimes the case that A'. If $<$ is linear, then LA can be defined as $HA \wedge A \wedge GA$. If time is not linear, then a new definition or a new primitive may be required. That p is an *instant-proposition*—in symbols, Ip—can now

be defined as $M(p \wedge \forall q(q \supset L(p \supset q)))$. If time is linear, then Ip can be defined without quantifiers as $M(p \wedge G - p \wedge H - p)$. However, this is not of much help since quantifiers will eventually need to be introduced. The reverse translation is then as follows: $R'\mathbf{x}\mathbf{y}t$ is replaced with $L(p \supset Rxy)$ and $t < u$ with $L(p \supset Fq)$; $\forall t$ is replaced with the quantifier $\forall p$ over instant-propositions; and $\forall \mathbf{x}$ is replaced with $L\forall x$.

The only new feature of the account is the definition of $<$ as $L(p \supset Fq)$. An equally good definition is $L(q \supset Pp)$. These definitions are interesting in that they are the only place in the translation where one of the unidirectional operators P or F is used in preference to the bidirectional operator L. Since each definition uses only one of the operators P or F, it should be possible, via the two translations, to define either operator in terms of the other and L. For example, a direct definition of FA is $\exists p(p \wedge Qp \wedge M(A \wedge Pp))$: A is future if at some time it is the case when the present is past. These definitions may suit those who are inclined to regard one of the tense-logical operators as more basic than the other, though it is hard to justify the choice of L as primitive.

This account will raise the same questions of existence for instant-propositions as previously arose for world-propositions. First, there will be the question of showing that instant-propositions always exist. It is tempting to follow the earlier arguments and say that an instant-proposition is the conjunction of all (tensed) propositions true at a given instant. However, the nature of the propositions that need to be chosen will be different in the cases of time and modality. To describe a possible world, one need use only propositions that are expressed by non-modal sentences, the modalities are simply not required in describing empirical reality. However, to describe an instant one may need to use propositions that are expressed by sentences containing tense-logical operators. One may need to say that such and such *will* happen or *has* happened. There do, indeed, seem to be cases in which no use of the tense-logical operators is required—e.g. 'Jesus Christ is born'. But, first, it is not clear that this sentence picks out a single instant and, second, the analysis of 'born' may well require the use of tense-logical operators. It follows that Prior's reduction of instants may be substantially different from one in terms of events, for future and past-tensed sentences are not usually regarded as describing present events.

Secondly, there is the question of whether instant-propositions (or tensed propositions, generally) always exist. Just as the modal actualist can claim that propositions about Socrates do not exist in worlds in which Socrates does not exist, so his tense-logical counterpart can claim that (tensed) propositions about Socrates do not exist at instants at which Socrates does not and, perhaps, has not existed. The solutions for the two cases are also similar, and so I shall not give separate details. Let us merely note that the cumulative interpretation of the quantifiers provides a special bonus; for if

two objects exist at different instants they also exist at the same instant, namely the later of the two instants, and so there will be simple definitions of the weak notions of §5 above in terms of the strong.

The tense-logical case also raises questions of its own. Can one be sure that instant-propositions, as defined, are true at one instant alone? (If the quantifier-free definition of I is adopted, then the question is one of existence.) A similar problem arose for world-propositions, but only in the context of contingent propositional existence. If propositions necessarily exist then there is no harm in identifying worlds at which the same propositions are true. But the same manœuvre is not available in tense-logic, for the properties of the earlier–later may be disturbed. In other words, there is no guarantee that $-M(p \land Fp)$ holds for instant-propositions.

There are two familiar situations in which this failure can occur. First, there is the case of an interval in which no change occurs. Now means of distinguishing the different instants of the interval may be open to the tense-logical theorist which are not open to someone who thinks in terms of events. Suppose A occurs throughout the interval, but not before. Then $A \land -PA$ occurs at the first instant of the interval, $A \land PA \land -PPA$ at the second instant, and so on. But this is to suppose that there *is* a first or second instant. If the interval were open and dense, then no such description could be given. Secondly, there is the case of cycles. Suppose time is like the integers and the instants alternate between A- and B-instants thus:

$$
\begin{array}{ccccc}
A & B & A & B & A \\
\cdot & \cdot & \cdot & \cdot & \cdot \\
-2 & -1 & 0 & 1 & 2
\end{array}
$$

Alternatively, and more plausibly, A and B could represent longish periods of history. Then, in this case too, there is no tense-logical sentence that distinguishes between the even or between the odd numbered instants.

There are various responses one can make to this difficulty. One is to deny that such situations can occur. Another is to tie individuals to times. Thus if Bill is smoking a pipe at Time 0 in the cyclic situation above then, strictly speaking, it must be another Bill smoking his pipe at Time 2. There is some plausibility in having time-bound individuals (though none, as far as I can see, in having them world-bound). However, time-bound individuals are not in keeping with the general approach of tense-logic and, in any case, they are of no help when the domain is empty at the instants to be distinguished. Yet a third response is to identify the indiscernible instants and then restructure the earlier–later relation accordingly. Thus the dead stretch above would correspond to an instant being earlier than itself ($M(p \land Fp)$) and the cycle would correspond to two instants being earlier than one another ($M(p \land Fp) \land M(q \land Fp)$). The earlier–later relation would, at least, be transitive and connected since there are tense-logical laws that guarantee

these properties. Various transformations (suggested by the completeness proofs in tense-logic) could then be used to obtain linear orderings, though these orderings would be far from unique. Another response altogether is to distinguish the different instants by Now. For example, in the cyclic situation, what distinguishes instants 0 and 2 is that at 0 *A* is happening *now*. I treat this response with some misgiving. For, first, I do not accept the corresponding argument in modal logic—that things are *actually* happening does not distinguish the actual world from other possible worlds; and, second, the use of 'now' smacks too much of a reference to an instant. Perhaps one should just accept that instant-propositions have the desired properties, even though no expressible propositions have those properties.

Let us now consider a language in which both tense-logical and modal operators are used together.[25] If the Priority Thesis for time and modality is accepted, then such a language is very natural. For modal operators will apply to tensed sentences, unless sentences subject to modal qualification are specially restricted; and similarly for the application of tense-logical operators. The original language, then, will be like the earlier modal languages, but with the three intensional operators G, H, and $\Box$. The classical language will now contain three sorts—for individuals, possible worlds, and instants. To each n-place predicate R, there will be an $(n + 2)$ place predicate R' with extra arguments for a world and instant. The classical translation is then as follows: Rxy goes into $R'\mathbf{xy}wt$; GB into $\forall u(t < u \supset B'(w, u))$, and similarly for H; $\Box B$ into $\forall v B'(v, t)$; and $\forall x B$ into $\forall x(E'wt \supset B'(\mathbf{x}, w, t))$. This semantics is sometimes called *two-dimensional* since it states the truth-conditions relative to a world-instant pair.

It is perhaps worth emphasizing that all modal notions are now suitably tensed. Thus the modal operators apply to tensed sentences to form tensed sentences, as in 'Possibly I will not exist'. Also, the domain of individuals is relative to the world at a time. There are not now two ontologies, one for time and the other for modality, but a single ontology for both. That only actuals or that only present individuals exist are now but two aspects of this one ontology.

In reversing the translation, the obvious strategy is to synthesize the earlier translations for time and modality. Thus world-propositions are defined by Q, instant-propositions by I, and the predication $R'\mathbf{xy}wt$ is replaced with $L\Box(p \land q \supset Rxy)$, where p is a world-proposition and q is an instant-proposition. But this will not do. For within the context of our two-dimensional language, the Q-propositions are those true in one world at the *present* instant and the *I*-propositions are those true at one instant in the

---

[25] Such a language is briefly mentioned in Prior [1968a], 133.

actual world; and these definitions have unfortunate consequences, both technical and philosophical.

The technical misfortune is that the account of predication requires that $p \wedge q$ is true at a single world-instant if p is a world-proposition and q is an instant-proposition. But once we stray from the actual world or the present instant there is no guarantee that this is so on the definitions above.

The philosophical misfortune is that the definitions do not respect the necessary features of instants or the sempiternal features of worlds. The *only* necessary features of instants seem to be that: (a) they are instants, (b) they exist, and (c) they have a certain structure, i.e. that if the instant is earlier than another then it is necessarily earlier. In a more sophisticated account of temporal structure, (b) might be disputed and (c) be made conditional on the existence of the instances. However, on our perhaps over-simple account, (a)–(c) should certainly hold. Now instant-propositions, as defined above, may not possess these necessary features. (a) corresponds to an instant-proposition p being necessarily an instant-proposition, (b) to p being necessarily true; and it is clear that neither (a) nor (b) need hold. (c) corresponds to $L(p \supset Fq)$ being necessarily true when true. Now if Fq is part of the description of p, as when the tensed-descriptions are complete, then (c) will hold; but not otherwise.

World-propositions possess corresponding sempiternal features, though (c) does not apply. However, the *necessary* features of possible worlds are not determined by the counterparts to (a) and (b). For, first, possible worlds (i.e. world-propositions) do not necessarily exist. But, second and more importantly, possible worlds possess other necessary features. What goes on in a possible world necessarily goes on in that world; indeed, the identity of a world is determined by its content. But what goes on at an instant is purely, or largely, accidental; the identity of an instant is determined, if at all, by its position in the temporal structure.

In order to overcome these difficulties, it is necessary to revise the definitions of instant- and world-propositions. An instant-proposition p is now to be true of one and the same instant in *each* possible world. In modal-tense-logical terms, the new definition of Ip is $M\square(p \wedge G -p \wedge H -p)$; p is an instant-proposition if at some time it is necessarily present and necessarily not future or past. If time is not linear, then the definition could be $M\square(p \wedge \forall q(q \supset L(p \supset q)))$. Similarly, a world-proposition is now true of one and the same world at each instant. The new definition of Qp is $\diamondsuit L(p \wedge \forall q(q \supset \square(p \supset q)))$, in exact analogy to the second definition of Ip. Thus instant-propositions and world-propositions are the temporal and modal cross-sections, respectively, of the two-dimensional instant-world manifold.

The reverse translation can now go as follows: $Rxywt$ is replaced with $L\square(p \wedge q \supset Rxy)$, where p and q are the world- and instant-propositions

corresponding to w and t respectively; $t < u$ is replaced with $L\square(q \supset Fr)$, where q and r are the instant-propositions corresponding to t and u; $\forall \mathbf{x}$ is replaced with $L\square\forall x$; and $\forall w$ and $\forall t$ are replaced with quantifiers $\forall p$ and $\forall q$ over world- and instant-propositions respectively.

An alternative, but equivalent, account can be given by defining the propositional counterpart to world-instant pairs. Say that p is a *state-proposition* if it is true at exactly one world-instant pair, i.e. if $M\diamondsuit(p \wedge \forall q(q \supset L\square(p \supset q)))$ holds. Thus a state-proposition corresponds to the state of the world at an instant. The state is not a time-slice, since that can reappear in other worlds, but *everything* that is presently the case at the given instant. Say that state-propositions p and *q agree on worlds* if $\diamondsuit(p \wedge Mq)$ holds and that they *agree on instants* if $M(p \wedge \diamondsuit p)$ holds. Then instead of quantifying over instant- or world-propositions in the reverse translation, one can quantify over all state-propositions that agree on instants or on worlds with the given state-proposition. It is as if instants and worlds are identified with suitable equivalence classes of world-instant pairs.

Any world- or instant-proposition in the new sense is also such a proposition in the old sense. Therefore there will be the same problems of existence and, also, the same solutions to those problems. However, the new sense brings its own existence-problems too. An instant-proposition is true in all possible worlds at a given instant; it is necessarily true at that instant. So any sentence expressing the proposition cannot refer to any contingent feature of the world; it cannot say anything significant about what is happening at the instant. But the only content-free sentences of this sort are those to the effect that things are happening *now*. Thus even without cycles or dead time, the present approach is committed to now-sentences or the inexpressibility of instant-propositions.

The world-propositions are subject to new problems over individual existence. Suppose I attempt to express the true world-proposition, i.e. I attempt to give a description of the world that, at all times, implies all truths. Now I may need to say that at some time an individual visits Mars, and yet not be able to say which individual since there have been no past or present Martian visitors. But assume that another possible world is the same as this one but for the identity of the merely future individuals and that, in particular the first visitor to Mars is a different individual. Then at the time of the first Martian visit, my description will no longer imply all truths since it will not imply that one individual was the visitor rather than another. Thus the world-proposition may change over time. As the world becomes more definite, then so do the propositions that describe it. But at no time may there be a proposition that describes the world, in its temporal entirety, with complete definiteness.

Under these circumstances, the earlier reduction will not work. For at a given time, a world-proposition will not pick out a single world but also all the other future 'realizations' of that world. The reduction can be repaired if the quantifiers are cumulative and the falsehood convention is adopted. Instead of determining whether a formula A is true at all instants in a given world, one deals with the instants at which the individuals involved in A exist by ordinary means and then deals with the other instants by means of the Falsehood Convention. Formally speaking, if B is a classical formula whose sole variable is $\mathbf{x}$ then the translation of $\forall w B$ is $L \forall p (Q p \supset (\Box(p \supset Ex) \supset B^*) \wedge (- \Box(p \supset Ex) \supset B^f))$, where B^f is the result of replacing atomic sentences containing free occurrences of x in B^* by $\perp$.

Our interest in modal tensed logic has centred on the reduction problem. However, many problems in the metaphysics of time involve possibility and are naturally discussed within the context of such a language. My aim is not to discuss the nature of time as such, but let me end with one illustration of this type of application. It is that there is an intimate connection between tense-logical priority and the view that time is absolute. Time may be absolute in two senses. It may be absolute in the sense that simultaneity is well-defined without reference to any physical system (the intra-world sense). Or it may be absolute in the sense that simultaneity is well-defined without reference to a possible world, i.e. that there is simultaneity between events of different possible worlds (the inter-world sense).

Now once the tense-logical theorist combines tense and modality he is almost committed to the view that time is absolute in the second sense.[26] For in order to interpret 'Possibly I am not sitting', he must suppose that in some possible world I am not sitting *at the same time* as I am sitting in the actual world. One could interpret this sentence as saying that another possible world first diverges from the actual world in the fact that I am not sitting. But this type of interpretation would not work for more complicated possibility sentences. To take an extreme example, the tense-logical theorist may claim that in the previous cycle it is possible that B occurs in place of when A does and A occurs in place of when B does. For at Time o, he may claim that it is possible that: B occurs, A occurs tomorrow, A occurs yesterday, and so on.

For the tense-logical theorist, the present has objective significance and he may secure cross-world simultaneity by supposing that at each time the same present runs through each possible world. In this respect, he has the edge over someone who believes that space is absolute; for it is most implausible to suppose that there is an objective 'here' or that spatial indexicals are primitives.

---

[26] Consult Prior [1968a], 133–4, for a discussion of how tense-logic may be reconciled with relativity physics.

Technical Appendix

This appendix gives a technical footing to some of the translations. For reasons of space, only the key translations are considered.

The First Translations

The translations combine an intensional account of possible worlds with a pseudo-quantifier approach to possible individuals. The propositional quantifiers are sentential and the propositions are assumed to exist necessarily.

Languages. There are three languages: the original modal language $\mathscr{L}$, the classical language $\mathscr{L}'$, and the secondary modal language $\mathscr{L}^*$. $\mathscr{L}$ is the language of quantified modal logic. It contains various predicates of specified degree, the individual variables $x_1, x_2, \ldots$, the universal quantifier $\forall$, the truth-functional connectives $\vee$ and $-$, the modal connective $\Box$, and brackets. We assume that the predicates include the one-place predicate E of existence and the two-place predicate I of identity. Formulas are defined in the usual way.

$\mathscr{L}'$ is a classical two-sorted first-order language. It contains: two sorts of variables—$x_1, x_2, \ldots$ for possible individuals and $w_1, w_2, \ldots$ for possible worlds; the constant ω for the actual world; for each predicate R of degree n in $\mathscr{L}$, a predicate R' of degree $(n + \pm 1)$ whose first n arguments are for individuals and whose last argument is for worlds; an identity predicate on individuals; the logical symbols $\vee$, $-$, and $\forall$, and brackets. Again, formulas are defined in the usual way.

$\mathscr{L}^*$ is obtained from $\mathscr{L}$ by adding propositional variables $p_1, p_2, \ldots$ To the definition of formula are added the clauses: p_i is a formula for $i = 1, 2, \ldots$; if A is a formula then so is $\forall p_i A$.

Translations. If x (or y_i) is the variable x_j, let $\mathbf{x}$ (or $\mathbf{y}_i$) be the variable $\mathbf{x}_j$. The classical translation t takes an $\mathscr{L}$-formula A and a world-variable or constant w into an $\mathscr{L}'$-formula $t(A,w)$. t is defined by the following clauses:

   (i) $t(Ry_1 \ldots y_n, w) = R'\mathbf{y}_1 \ldots \mathbf{y}_n w$
   (ii) $t(-B, w) = -t(B, w)$
   (iii) $t((B \vee C), w) = (t(B, w) \vee t(C, w))$
   (iv) $t(\forall x B, w) = \forall \mathbf{x}(E'\mathbf{x}w \supset t(B, w))$
   (v) $t(\Box B, w) = \forall w_1 t(B, w_1)$

t is defined on formulas alone by the clause:

   (vi) $t(A) = t(A, \omega)$

For A in $\mathscr{L}^*$, let QA abbreviate $\Diamond(A \wedge \forall p(p \supset \Box(A \supset p)))$ where p is the first propositional variable not to occur free in A. The reverse translation s takes ω-free formulas of $\mathscr{L}'$ into $\mathscr{L}^*$-formulas. Its defining clauses are:

(i) (a) $s(R'\mathbf{y}_1 \ldots \mathbf{y}_n w_i) = \square(p_i \supset R\mathbf{y}_1 \ldots y_n)$
   (b) $s(\mathbf{y}_1 = \mathbf{y}_2) = y_1 = y_2$

(ii) $s(-B) = -s(B)$

(iii) $s((B \vee C)) = (S(B) \vee s(C))$

(iv) $s(\forall w_i B) = \forall p_i (Qp_i \supset s(B))$

(v) $s(\forall \mathbf{x} B) = \square \forall x s(B)$

All subsequent translations will satisfy the analogues of clauses (ii)–(iii) and therefore this part of the definition will be omitted.

s is extended to formulas A containing ω in the following way. Let A^g be the result of replacing each occurrence of ω in A by the first variable w_i not to occur free in A. Then:

$$s(A) = \exists p_i(p_i \wedge Qp_i \wedge s(A^g)).$$

This translation is equivalent to treating ω as a description $\iota p_i(p_i \wedge Qp_i)$ with widest scope, for the uniqueness condition is automatically satisfied.

Semantics. A *model* for $\mathscr{L}$ is a quadruple $\mathscr{M} = (W, D, w_o, \phi)$ such that:

(a) W (worlds) and D (possible individuals) are non-empty sets, w_o (actual world) εW, and $\phi(R, w) \subseteq D^n$ for each n-place predicate R and $w \varepsilon W$;

(b) $\phi(I, w) = \{\langle a, a \rangle : a \varepsilon D\}$ for each $w \varepsilon W$;

(c) $(\forall a \varepsilon D)(\exists w \varepsilon W)(a \varepsilon \phi(E, w))$.

Let $D_w = \{a : a \varepsilon \phi(E, w)\}$ and call θ an *assignment* if it is a function from the set of individual variables into D. Then the definition of satisfaction contains the following clauses:

(i) $\mathscr{M} \models_\theta Ry_1 \ldots y_n$ iff $\langle \theta(y_1), \ldots, \theta(y_n) \rangle \varepsilon \phi(R, w_o)$

(ii) $\mathscr{M} \models_\theta \square B$ iff $\mathscr{M}' \models_\theta B$ for all models $\mathscr{M}'$ differing from $\mathscr{M}$ in the third component only

(iii) $\mathscr{M} \models_\theta \forall x_i B$ iff $\mathscr{M} \models_{\theta'} B$ for all θ' such that $\theta'(x_i) \varepsilon D_{w_o}$ and $\theta'(x_j) = \theta(x_j)$ for all $j \neq i$.

The clauses for $\vee$ and $-$ are standard.

Given an $\mathscr{L}$-model $\mathscr{M} = (W, D, w_o, \phi)$, let ϕ' be the function ψ such that $\psi(R') = \{\langle a_1, \ldots, a_n, w \rangle : \langle a_1, \ldots, a_n \rangle \varepsilon \phi(R, w)\}$ for each predicate R' in $\mathscr{L}'$ and let $\mathscr{M}' = (W, D, w_o, \phi')$. Then a *model for* $\mathscr{L}'$ is a model $\mathscr{M}'$ where $\mathscr{M}$ is a model for $\mathscr{L}$. In $\mathscr{M}'$, W is now regarded as the domain of the world-variables, D as the domain of the individual-variables, w_o as the denotation of ω, and ψ as the valuation. The truth-definition then proceeds along classical lines.

A *model for* $\mathscr{L}^*$ is a quintuple $\mathscr{C} = (W, D, P, w_o, \phi)$, where $\mathscr{M} = (W, D, w_o, \phi)$ is an $\mathscr{L}$-model and P (propositions) is a subset of $\mathscr{P}(W)$. Such a model is *complete* if $P = \mathscr{B}(W)$. θ, α is an *assignment for* $\mathscr{C}$ if

θ is an assignment for $\mathcal{M}$, as above, and α is a function from the set of propositional variables p_1, $p_2, \ldots$ into P. In the definition of satisfaction, the clause for $\forall p_i$ is:

  $\mathcal{C} \models_{\theta,\,\alpha} \forall p_i B$ iff $\mathcal{C} \models_{\theta,\,\alpha'} B$ for all α' such that $\alpha'(p_i)\varepsilon\, P$ and $\alpha'(p_j) = \alpha(p_j)$ for all $j \neq i$.

The other clauses are as before.

  Correctness of translations. We state results which show that the translations are correct and compatible with one another.

  Given an $\mathcal{L}$-model $\mathcal{M} = (W, D, w_o, \phi)$, let $\mathcal{M}^*$ be the $\mathcal{L}^*$-model $(W, D, \mathscr{P}(W), w_o, \phi)$.

  Theorem 1. (a) For any $\mathcal{L}$-sentence A and $\mathcal{L}$-model $\mathcal{M}$, $\mathcal{M} \models A$ iff $\mathcal{M}' \models t(A)$.
  (b) For any $\mathcal{L}'$-sentence A and $\mathcal{L}'$-model $\mathcal{M}'$, $\mathcal{M}' \models A$ iff $\mathcal{M}^* \models s(A)$.
  (c) For any $\mathcal{L}$-sentence A and $\mathcal{L}$-model $\mathcal{M}$, $\mathcal{M} \models A$ iff $\mathcal{M}^* \models s(t(A))$.

Similar results hold for formulas; indeed, these are required for an inductive proof of the above theorem.

  Say that a sentence A of $\mathcal{L}(\mathcal{L}', \mathcal{L}^*)$ is *valid* if it is true in all $\mathcal{L}$-models (respectively: $\mathcal{L}'$-models, complete $\mathcal{L}^*$-models). Then the theorem has the following consequences:

  Corollary 2. For sentences A from the appropriate language:
  (a) A is $\mathcal{L}$-valid iff $t(A)$ is $\mathcal{L}'$-valid
  (b) A is $\mathcal{L}'$-valid iff $s(A)$ is $\mathcal{L}^*$-valid
  (c) A is $\mathcal{L}$-valid iff $s(t(A))$ is $\mathcal{L}^*$-valid
  (d) $A \equiv s(t(A))$ is $\mathcal{L}^*$-valid.

The above results also hold if complete $\mathcal{L}^*$-models are replaced with *atomic* models, i.e. ones in which P is an atomic Boolean algebra.

  Axiomatization. The sets of $\mathcal{L}$-valid and $\mathcal{L}'$-valid formulas can be axiomatized. An axiomatization of the $\mathcal{L}$-valid formulas can be found in Fine [1978]. An axiomatization of the $\mathcal{L}'$-valid formulas can be obtained by adding the following axioms to the appropriate two-sorted predicate calculus:

  $I'\mathbf{x}_1\mathbf{x}_2 w_1 \equiv \mathbf{x}_1 = \mathbf{x}_2$
  $\forall \mathbf{x}_1 \exists w_1 E'\mathbf{x}_1 w$

These axioms express the conditions (b) and (c) on an $\mathcal{L}$-model; and so completeness follows from the completeness result for the predicate calculus.

Corollary 2 now implies a syntactic counterpart to (a):

> *Corollary* 3. *A* is a theorem of the system for $\mathscr{L}$ iff $t(A)$ is a theorem of the system for $\mathscr{L}'$.

This result also has, of course, a direct syntactic proof.

The set of $\mathscr{L}^*$-valid formulas is not axiomatizable. However, it may be determined which axioms are required for the syntactic counterparts of corollary 2 (b)–(d) to hold. First, add to the system for $\mathscr{L}$ the standard quantificational principles for $\forall p_i$, namely $\forall p_i A \supset A p_i / B$, B free for p_i in A, and $\forall p_i (B \supset C) \supset (B \supset \forall p_i C)$, p_i not free in B. Then add these non-logical axioms:

> *Existence of world-propositions* $\Box \exists p(p \wedge Qp)$
> *Barcan formula* $\Box \forall p A \equiv \forall p \Box A$
> *ω-completeness* $\Box \forall p(Qp \wedge \Box(p \supset \exists x B) \supset \Diamond \exists x \Box (p \supset B))$

The results then follow:

> *Theorem* 4. For sentences A from the appropriate language:
> (a) A is an $\mathscr{L}'$-theorem iff $s(A)$ is an $\mathscr{L}^*$-theorem
> (b) A is an $\mathscr{L}$-theorem iff $s(t(A))$ is an $\mathscr{L}^*$-theorem
> (c) $A \equiv s(t(A))$ is an $\mathscr{L}^*$-theorem.

Dropping individual variables. Further results can be obtained if individual variables are dropped and consequent changes are made in the languages, models and translations. $\mathscr{L}$ is now the language of modal sentential logic, $\mathscr{L}'$ the language of the monadic predicate calculus (with world-variables only), and $\mathscr{L}^*$ the language of $\mathscr{L}$ with sentence variables.

First, the logic for $\mathscr{L}^*$ is now decidable and has a particularly simple axiomatization.[27]

Second, a reverse translation into $\mathscr{L}$ is now possible. Say that an $\mathscr{L}'$-formula is *uniform* if it contains no world-variables other than w_1. Then note that $t(A)$ is always a uniform formula. In case individual variables are dropped, a translation of uniform $\mathscr{L}'$-formulas into $\mathscr{L}$ can be obtained by replacing $\forall w_1$ with $\Box$ and $P' \omega$ with P (but after $P' \omega$ has been moved outside the scope of a $\forall w$ quantifier).[28] Since each formula of the monadic predicate calculus is equivalent to a uniform one, this gives a full reverse translation. However, if individual variables are retained, then the translation of uniform $\mathscr{L}'$-formulas into $\mathscr{L}$ is no longer possible. For example, $\forall w_1 \exists x R' \mathbf{x} w$ has no translation in $\mathscr{L}$.

Translations without Necessarily Existing Propositions

These translations extend the earlier ones to the case in which propositions do not necessarily exist.

---

[27] See Fine [1970], and the references therein.
[28] See Prior and Fine [1977], last sections of ch. 2.

Languages. $\mathscr{L}$ and $\mathscr{L}'$ are as before. The secondary modal language, now dubbed $\mathscr{L}^{\circ}$, is obtained from $\mathscr{L}$ by adding: propositional variables $\rho_1, \rho_2, \ldots$; and the predicates T' for existence and $\rightarrow'$ for strict implication. The variables ρ_i are now nominal and so to the definition of formula in $\mathscr{L}$ is added the clause:

> $T'\rho$, $\rho \rightarrow' \sigma$, $E\rho$ and $\rho = \sigma$ are formulas when ρ and σ are propositional variables.

The proposition–that operator ξ is not required for the reduction. However, if it were added to the language, one would need to distinguish between terms and formulas. Propositional variables are terms, and so is ξA for A a formula; $T'm$, $m \rightarrow' n$, etc. are formulas for arbitrary terms m and n.

Translations. Let $Q\rho_i$ now abbreviate

$$\Diamond(T'\rho_i \wedge \forall\rho_{i+1}(T'\rho_{i+1} \supset (\rho_i \rightarrow' \rho_{i+1}))).$$

The reverse translation r from ω-free formulas of $\mathscr{L}'$ into $\mathscr{L}^{\circ}$ is then defined by:

(i) (a) $\mathrm{r}(R'\mathbf{y}_1 \ldots \mathbf{y}_n w_i) = \Box(T'\rho_i \supset R y_1 \ldots y_n)$
   (b) $\mathrm{r}(\mathbf{y}_1 = \mathbf{y}_2) = y_1 = y_2$
(ii) $\mathrm{r}(\forall w_i B) = \Box\forall\rho_i(Q\rho_i \supset \mathrm{r}(B))$
(v) $\mathrm{r}(\forall \mathbf{x}B) = \Box\forall x \mathrm{r}(B)$.

As before, let A^g be the result of replacing each occurrence of ω in A by the first world variable w_i not to occur free in A. Then r is extended to all $\mathscr{L}'$-formulas A by the clause:

$$\mathrm{r}(A) = \exists\rho_i(T'\rho_i \wedge Q\rho_i \wedge \mathrm{r}(A^g)).$$

Semantics. A *model for* $\mathscr{L}^{\circ}$ is a sextuple $\mathscr{C} = (W, D, P, w_o, \phi, x)$, where (W, D, P, w_o, ϕ) is an $\mathscr{L}^*$-model and X is a function from W into $\mathscr{P}(P)$ subject to the condition that $(\forall X \,\varepsilon\, P)(\exists w \,\varepsilon\, W)(X \,\varepsilon\, X(w))$. Intuitively, $X(w)$ is the domain of propositions that exist in world w. The modal $\mathscr{C}$ is *complete* if each $X(w)$ is a complete Boolean algebra and *atomic* if $\{w\} \,\varepsilon\, X(w)$ for each w. Note that a complete model need not be atomic.

An *assignment* θ, α *for* $\mathscr{C}$ is defined as before, but with $\rho_1, \rho_2, \ldots$ replacing $p_1, p_2, \ldots$ in the domain of α. The noteworthy clauses in the truth-definition are:

(i) (a) $\mathscr{C} \models_{\theta,\alpha} T'\rho_i$ iff $\alpha(\rho_i) \,\varepsilon\, X(w_o)$ and $w_o \,\varepsilon\, \alpha(\rho_i)$
   (b) $\mathscr{C} \models_{\theta,\alpha} \rho_i \rightarrow' \rho_j$ iff $\alpha(\rho_i), \alpha(\rho_j) \,\varepsilon\, X(w_o)$ and $\alpha(\rho_i) \subseteq \alpha(\rho_j)$
   (c) $\mathscr{C} \models_{\theta,\alpha} E\rho_i$ iff $\alpha(\rho_i) \,\varepsilon\, X(w_o)$
   (d) $\mathscr{C} \models_{\theta,\alpha} \rho_i = \rho_j$ iff $\alpha(\rho_i) = \alpha(\rho_j)$
(ii) $\mathscr{C} \models_{\theta,\alpha} \forall\rho_i B$ iff $\mathscr{C} \models_{\theta,\alpha'} B$ for all α' such that $\alpha'(\rho_i) \,\varepsilon\, X(w_o)$ and $\alpha'(\rho_j) = \alpha(\rho_j)$ for all $j \neq i$.

If the language $\mathscr{L}^\circ$ contained ξ, then the denotation $d(m)$ of each term m would need to be determined: $d(\rho_i) = X(\rho_i)$; and $d(\xi A) = \{w \; \varepsilon \; W: \mathscr{C} \models_{\theta,\alpha} A\}$. Clause (i) could then be formulated with denotations of arbitrary terms replacing the assignment function α.

Correctness. Say then an $\mathscr{L}^\circ$-model $\mathscr{M}$ *extends* an $\mathscr{L}'$-model $\mathscr{M}$ if $\mathscr{M}$ is $\mathscr{N}$ without its last component. Then in analogy to theorem 1 (b)–(c), we have:

Theorem 5. Suppose $\mathscr{M}$ is an $\mathscr{L}$-model and $\mathscr{N}$ an atomic $\mathscr{L}^\circ$-model that extends $\mathscr{M}$. Then:

      (a) $\mathscr{M}' \models A$ iff $\mathscr{N} \models r(A)$ for any $\mathscr{L}'$-sentence A
      (b) $\mathscr{M} \models A$ iff $\mathscr{N} \models r(t(A))$ for any $\mathscr{L}$-sentence A.

It would also be possible to establish analogues of Corollaries 2 and 3 and Theorem 4. However, I shall not go into details.

5

Plantinga on the Reduction of Possibilist Discourse

Plantinga is what I call a modal actualist. He believes that the idioms of necessity and possibility are to be taken as primitive in preference to talk of possible worlds and that only actuals, as opposed to possibles, are to be granted ontological status. On these two issues, he and I agree.

The modal actualist faces a challenge. Talk of possible worlds and of possible individuals appears to make perfectly good sense. There seems to be a clear meaning, for example, in the claim that some possible object does not exist. So the modal actualist, once he grants that possibilist discourse makes sense, must somehow give it sense. It is on this question of how such a challenge is to be met that Plantinga and I disagree.

He favours a reduction of possibilist discourse in which possible worlds and possible individuals give way to propositions and properties, respectively; I favour a reduction in which reference to possibles becomes a modal manner of reference to actuals. In this chapter, I shall attempt to adjudicate between these rival positions.

In the first section, I shall set out the problem of reduction and Plantinga's favoured solution. In the second, I shall present my central criticism of the reduction, namely that it is question-begging. In the next three sections, I shall consider the related question of whether properties and propositions exist necessarily, first presenting an argument against and then disposing of an argument for their necessary existence. In the final section, I shall present my own reduction and the reasons for preferring it to Plantinga's.

The central theme of this chapter is the question of reduction; but it should have a broader significance than such a theme might suggest. Partly this is because other issues, of independent interest, are raised: the connection between existence and predication; the necessary existence of propositions; the Priorian stand on modality. But perhaps more important than this question of particular issues is the question of how the issues are to be approached, of what is to count as a plausible consideration one way or another. Even when I have found myself in agreement with Plantinga on a certain view, I have often also found myself unhappy with the reasons he

adduces in its favour. It is in this difference of approach, then, that the chapter may also have a broader significance.

1. The Plantinga Reduction

Modal actualism finds its natural setting in a language for which the modal connectives are primitive and the quantifiers are actualist, ranging, in each world, over the actuals of that world. Such languages are familiar from the study of modal logic. Possibilism, on the other hand, finds its natural setting in a language for which the connectives are truth-functional and the quantifiers are possibilist, ranging either over possible worlds or possible individuals. These languages are familiar from the study of classical logic.

As is well known, the possibilist is able to translate the language of the modal actualist into his own in such a way as to preserve truth. First, he explains what it is for a statement of the modal actualist to be true *in a world*. He says, for example, that $\forall x A(x)$ is true in a world if and only if $A(x)$ is true in that world of all of the (possible) individuals that exist in that world; and he says that necessarily A ($\Box$ A) is true in a world if and only if A is true in all worlds. He then takes the modal actualist's statement A to assert that A is true in the actual world. If, for example, A is the statement 'something is a philosopher but possibly not a philosopher' ($\exists x (Px \land \Diamond -Px)$), then the possibilist's translation is 'some (possible) individual exists in the actual world, is a philosopher in the actual world, and is not a philosopher in some (other) possible world' ($\Sigma w \Sigma x (Aw \land E^* xw \land P^* xw \land \Sigma v - P^* xv)$); where **A** is the predicate for being actual).

The question now arises as to whether the modal actualist can return the favour, whether he can translate the possibilist's language back into his own. Once we go into the details of the possibilist's language, we see that three key locutions are involved:[1] first, the predicate for the actual world; second, the atomic predications used by the possibilist—these comprise the statements of identity between possible individuals ($x = y$), the ordinary world-relative predications such as 'x is a philosopher in w' ($P^* xw$), and the special world-relative predication 'x exists in w' ($E^* xw$); finally, the quantifiers (Σw and Σx) over possible worlds and possible individuals. The modal actualist must somehow render these locutions in acceptable terms.

This is the problem. What is Plantinga's solution? He proposes an answer along essentially the following lines: identify possible worlds with world-propositions, propositions true in one world alone, and possible individuals with individual essences, properties true of a single possible individual in

---

[1] I assume that the language is first-order. Additional problems, ignored in this chapter, can arise once the possibilist is allowed to quantify over sets of possibles or other higher-order entities.

each possible world; then treat properties of possible worlds and individuals as corresponding properties of world-propositions and essences. The main text for the account is Plantinga [1976] though there are hints of it to be found elsewhere in his work.

The history of such an account goes back to Prior [1968*b*], who worked out the reduction of possible worlds to world-propositions, but not of possible individuals to essences. Essentially the same account was later given by Adams [1974]. The extension to possible individuals was independently proposed by myself (Ch. 4 above) and by Plantinga; although, unlike Plantinga, I only raised my own account in order to shoot it down.

This approach to the problem may be seen as having its origin in a general strategy for dealing with problems of reduction. This is first to identify the disputed entities with suitable surrogates and then to treat the properties of the entities as corresponding properties of the surrogates. (See Quine [1964, 1969].) Once one applies this strategy to the case at hand, there is perhaps no more natural choice for the surrogates than the world-propositions and the individual essences. Indeed, this choice of surrogates can be regarded as arising from a common source. For one can treat propositions as a special kind of property, true, in each world, *of* the worlds in which they are true. The world-propositions then become special cases of individual essences, and the identification as a whole becomes an instance of a 'bundle theory', with disputed entities giving way to their properties.

For the reduction to be complete, it must be shown what properties of the surrogates are to correspond to the properties of the given entities. This means that the modal actualist must give an account of the counterparts to the three key locutions of the possibilist. Plantinga [1976] is a little short on details here; he merely works through a few examples. But it is clear from what he says how a systematic account might go. Instead of saying that a possible world is actual, say that a world-proposition is true. Instead of saying that an individual has P in a world (P^*xw), for P an ordinary predicate or the existence-predicate, say that possibly some (actual) individual has an essence and necessarily P's whenever the world-proposition is true. ($\diamondsuit \exists x(x\mathrm{H}\varphi \wedge \square(\mathrm{T}\rho \supset \mathrm{P}x)$, where H is the predicate of having between an object and a property).[2] Instead of saying that two possible individuals *are identical*, say that the individual essences are possibly co-exemplified ($\diamondsuit\exists x(x\mathrm{H}\varphi \wedge x\mathrm{H}\Psi)$).[3] Replace quantification over possible worlds with quantification over world-propositions, which are now taken to be proposi-

---

[2] Plantinga's suggestion in [1976] (Loux [1979], 268) is rather different; it is to adopt $\square(\mathrm{T}\rho \supset \exists x(x\mathrm{H}\varphi \wedge \mathrm{P}x))$ as the paraphrase. However, this account is only correct when the predicate P is never true of non-existents in a world. Our account has the advantage of not depending upon such an assumption for its correctness.

[3] Plantinga omits to state any counterpart to identity between possible individuals; but I take it that there is no real alternative to the present proposal.

tions that possibly are true and imply all truths ($\Diamond(T\rho \wedge \forall\sigma(T\sigma \supset \Box(T\rho \supset T\sigma)))$, where T is the truth-predicate).[4] And, finally, replace quantification over possible individuals with quantification over individual essences, which are now taken to be properties for which it is possible that some (actual) individual is necessarily its sole possessor ($\Diamond\exists x\,\Box(xH\varphi \wedge \forall y\Box(yH_\varphi \supset y = x)))$.[5]

With these stipulations, the reduction is then complete.

2.   The Central Charge

The central charge to be made against the preceding reduction is that it is circular. Its correctness requires that use be made of essences and world-propositions that already presuppose the possible individuals in question.

Crucial to this question of circularity is a certain assumption. Call a property *actualist* if it only 'involves' actual individuals, no merely possible individuals. Then the assumption is that each possible individual has an *actualist* individual essence.[6] Given that properties are closed under arbitrarily long conjunctions, this assumption is equivalent to the claim that for any two distinct possible individuals there is an actualist property essentially possessed by the one but not by the other.[7] Thus the assumption is a version of the principle of the identity of indiscernibles, what might be called the 'Discernibility Doctrine'.

These principles have been formulated using the problematic notion of 'involves'. But a purely syntactic formulation of the thesis may also be given. Call a formula A(x), of one free variable x, *actualist* if it contains no (genuine) names for merely possible individuals. Then the assumption states that for each possible individual there is an actualist formula true of exactly that individual in each possible world. On the other hand, the Discernibility

---

[4] Plantinga's definition, as suggested by [1974], 45, and elsewhere, is $\Diamond T\rho \wedge \forall\sigma(\Box(T\rho \supset T\sigma)\vee \Box(T\rho \supset -T\sigma))$. However, our definition has the advantage of not depending for its correctness on the necessary existence of propositions.

[5] Plantinga's definition, as suggested by [1974], 72, and elsewhere, is $\Diamond\exists x\Box\forall y(yH\varphi \equiv x = y)$. Again, our definition has the advantage of not depending for its correctness upon the assumption that properties are never true of non-existents.

[6] A stronger assumption is that each possible individual has a purely qualitative individual essence, one not involving any individuals at all, either possible or actual. It is this stronger assumption that Plantinga introduces in [1979], 111, in connection with the necessary existence of properties; but it is the weaker assumption that is most relevant to his question, as it is to ours. It must be allowed, though, that there is some sort of metaphysical oddity involved in all individuals having an actualist, but not a qualitative, essence, and that therefore qualitative discernibility at least inherits some plausibility from actualist discernibility.

[7] Grant the assumption. Given distinct possibles x and y, let the differentiating essential property be an actualist individual essence of x. Grant the claim. Then let the individual essence of x be the conjunction of its actualist essential properties (including existence).

Doctrine states that for any two distinct possible individuals there is an actualist formula $A(x)$, true of the one in each world in which it exists, but not true of the other in each world in which it exists. The formulas $A(x)$ are to belong to an ideal language, one that God might speak. Thus there will be no gap between the properties that exist and the properties expressible in the language.

Granted the Discernibility Doctrine, the circularity can be avoided. The original reduction might still be circular; for it quantifies over all essences and all world-propositions, and some of them might still involve the possible individuals. But require the essences and world-propositions to be actualist, and all circularity is removed.

It is far from clear, however, that the Discernibility Doctrine holds. Presumably, there could have been different elementary particles from the ones there in fact are. Take now two merely possible elementary particles of the same kind. Then it is hard to see how these particles could be distinguished in terms of their actualist essential properties. There seems to be no basis upon which such a distinction could be made.[8]

But even if the doctrine is undisputed, it seems objectionable that the reduction should be made to depend upon it. For one thing, the reduction thereby becomes more vulnerable to criticism. But also, more importantly, we feel that the reduction of possibilist disourse should not depend upon any particular modal views, that if the reduction is possible then that possibility should be written into the very nature of the discourse itself.[9]

---

[8] Various counterexamples have been levelled against the identity of indiscernibles in the literature; and Plantinga presents his own counterexamples against what I have called the Discernibility Doctrine, but with purely qualitative properties, in [1979], 111–114. In one respect, though, his discussion is misleading; for he also seems to presuppose that the essence is non-modal. It is only in this way that one can make sense of his 'complete descriptions'. But one needs to distinguish between the doctrines of discernibility, by purely qualitative and by purely qualitative *and* non-modal means. The latter is essentially a question of whether there could be distinct individuals such that one in *some* world in which it existed was just like the other in a corresponding world in which it existed, of whether the two individuals could share their world-roles. The former is substantially (though not equivalently) a question of whether there could be distinct individuals such that the one in any *arbitrary* world in which it existed was just like the other in some corresponding world in which it existed, of whether the two individuals could share their *portfolios* of world-roles. The two questions are not the same, since two distinct individuals might share given world-roles without sharing their portfolios. Probably what breeds the confusion is a conflation between the present problem of individuation and the more familiar problem of transworld identity. To identify an individual 'across possible worlds' one must use non-modal essences; but in the present context, the transworld identity of individuals is not in question and so modal means of individuation may also be used. (It is clear from this footnote and n. 6 how much care must be taken in properly formulating the relevant version of the identity of indiscernibles.)

[9] It is also for these reasons that I proposed the changes mentioned in nn. 2, 4, and 5. Plantinga might believe in the assumptions upon which the correctness of his definitions depend; but it is better that the reduction not be made to depend upon them.

In any case, Plantinga ([1979], 112–14) rejects the Discernibility Doctrine (in its qualitative form); and so in criticizing him, it would appear reasonable to assume that the actualist form of the doctrine is false.

Once the Discernibility Doctrine is denied, it follows that an essence for each possible individual cannot be specified in actualist terms alone. I take a *specification* of a property to be essentially a matter of producing a formula $A(x)$ that expresses the property. The notion of 'expresses' may be problematic here; but all I require of the notion is that a formula and the property it expresses be true of the same individuals in each possible world. Suppose now that the Discernibility Doctrine is false. Then some possible individual will have no actualist essence; there will be no actualist formula $A(x)$ true of that individual in each possible world. But then it is trivial that no actualist formula can express an essence for the individual.

It will be impossible, *even in principle*, to specify an individual essence for each possible individual without referring to some merely possible individuals. The circularity of the reduction is then apparent; an adequate supply of surrogates for the merely possible individuals requires that we already presuppose some of those individuals.

The circularity may be brought out in starker form as follows. Take the reduction to be one in which each possible individual is identified with its *identity property*,[10] the property of being identical to the given individual. The circularity now seems especially blatant. But the reduction is really no less acceptable than any other reduction in which the possible individuals are identified with their essences; for although, in specifying an essence for a given individual, we may not need to refer to that individual, eventually, in specifying essences for all the merely possible individuals, we must make reference to some of those individuals. The reference may be postponed, but not avoided.

Now normally I would be content to let the argument rest there, to let the circularity speak for itself. But since the question is so crucial to the success of Plantinga's reduction, we would do well to consider it further.

There are two different ways in which we can push the case against Plantinga. First, we can try to bring out how objectionable the reduction is by pointing to the consequences of adopting reductions of this kind. Imagine a philosopher (a platonic idealist) worried by material things, but not by properties or sets as such. He then proposes to get rid of material things in favour of their identity properties or singleton sets. Surely there is some sort

---

[10] Plantinga ([1983], 1) follows Adams ([1981], 4) in calling identity-properties 'thisnesses' or 'haecceities'. I do not like this terminology. The terms 'thisness' and 'haecceity' suggest something grander than mere identity properties. These terms are naturally taken to refer to that which explains or underlies the specific individuality of things; and this also seems closer to Scotian usage. On the other hand, identity properties do not explain individuality; they presuppose it.

of trick or cheat here. In the same way, it is not as if someone who accepts sets or properties can thereby rid himself of all unwanted entities in favour of their identity properties or singleton sets. There must be something wrong in a procedure that so trivializes the ontological enterprise.

Secondly, we may articulate more clearly what the objectionable character of the reduction consists in. In order to say what an identity property or singleton set is, we must make mention of the given individual. What is singleton Socrates? The set whose sole member is *Socrates*. What is the identity property for Socrates? The property of being identical to *this man*, Socrates. (Perhaps, on certain views, we can say what the identity property for Socrates is without bringing in Socrates. But a similar point will still hold, namely that in explaining the identity of *all* identity properties we will need to bring in some merely possible individual.) Let us say that y *presupposes* x if the identity of y stands in need of explanation and x is required in order to explain that identity. Then, in this technical sense, we may say that singleton sets presuppose their members and identity properties presuppose their bearers.

It now seems reasonable that if y belongs to an ontology and y presupposes x, then x should belong to that ontology; the objects of the ontology should be closed under presupposition. For suppose y belonged to the ontology and that y presupposed x, Then, in so far as the identity of y stands in need of explanation, it should be possible to explain its identity in terms of the elements of the ontology. So if x is required to explain its identity, it too must belong to the ontology.

We now reach the conclusion that an ontology containing singleton sets or identity properties must also contain their members or bearers. Therefore an actualist ontology, one containing only actual individuals, cannot contain the singleton sets or identity properties of merely possible individuals.

The point might be put picturesquely as follows. The actualist must make possibilist discourse intelligible from a standpoint that assumes complete scepticism with regard to merely possible individuals. But from such a standpoint, a world of actual individuals alone, we could make no sense of what the singleton sets or identity properties for mere possibles were. They would be like ciphers for sets or properties, not sets or properties themselves.

What misleads us here is a picture of the platonic realm as ontologically autonomous, as forming a world apart. It then seems reasonable that one who accepts this world could remain neutral on what else there is. But this picture should be replaced by one in which the platonic entities may *grip onto* the rest of reality. To the extent that they do, a commitment to the platonic realm will then carry with it a commitment to the other entities.

We see that Plantinga faces a double challenge: he must explain why it is acceptable for him to replace entities with their identity properties,

yet not acceptable in general; and he must dispose of the argument from presupposition.

There are various ways he might deal with the argument. Perhaps the most plausible is to argue that the identity properties do not stand in need of explanation, that they are primitive elements of the ontology. But this objection is weak. Presumably other relational properties—being a wife of Socrates, being a pupil of Plato—are not primitive. So why not being identical to Socrates? And if this identity property, why not also the property of being identical to a given merely possible individual? There is the further difficulty that we want to explain the application conditions for a property in terms of its identity. It is because an identity property is what it is that, necessarily, something has the property iff it is identical to the given individual. But the application conditions for an identity property become a mystery once its identity is taken as primitive. (Related considerations against essences or Haecceities being primitive are given in Adams [1981], 12–18).

Perhaps this objection, or another, can somehow be sustained. But we need not pursue the matter, since the second challenge, of discriminating between the reductions of Plantinga and of the idealist, would appear to be unanswerable. It is no use saying here that the identity properties of possibles are primitive, since the platonic idealist can equally well appeal to the primitiveness of the identity properties in support of his own reduction. What then is Plantinga to say?

There are two ways in which Plantinga might attempt to explain why his own reduction is acceptable, yet not that of the platonic idealist. The first is to argue that the idealist cannot explain in general what a singleton set or identity property for a material thing is without bringing in the general concept of a material thing; such a set or property is one whose sole member or possessor is a material thing. On the other hand, the actualist can explain what an identity property is, at least up to necessary coextensiveness, without appeal to the general concept of a possible object; for an identity property is then simply an individual essence, which, as we have seen, can be characterized in terms acceptable to the modal actualist.

But this difference between the idealist and actualist is, even if it exists, not a relevant difference. There are idealists who find the *concept* of a material thing perfectly intelligible. Perhaps they are prepared to define it as a non-mental non-abstract thing or in some other way. But it is not as if the reduction of material things to identity properties suddenly becomes acceptable to such an idealist. What the idealist is primarily concerned to get rid of is not the general concept of a material thing, but of the things that answer to that concept; and if he is not happy with the things, he is not going to be happy with the entities, such as the sets or properties, that presuppose those things. In the same way, our actualist should be worried not so much by the

general concept of a possible individual, but by an ontology that includes possible individuals or entities that presuppose them.

The other difference that Plantinga might appeal to lies in the predicates that the actualist and idealist attach to their respective surrogate entities. The idealist must explain the predicate 'is red' as applied to material things, let us say, in terms of a corresponding predicate for singleton sets or identity properties. But how is such a predicate to be analysed? Presumably in terms of some such expression as 'is a set containing (or a property had by) a material thing that is red', which presupposes the entities whose ontological status is in question. On the other hand, the actualist can analyse the predicates of essences corresponding to predicates of possibles without presupposing those possibles. What corresponds to the predicate 'is red in the world w', for example, is the predicate 'is an essence such that it is possible that some (actual) individual with the essence necessarily is red whenever the world-proposition corresponding to w is true'.

It must be conceded that there is a big difference between the reductions of the idealist and the actualist in this respect; the predicates for the surrogates of the one are circular, of the other not. But does this clear the actualist reduction of all circularity? I think not. The idealist reduction, it seems, suffers from a double circularity; one located in the predicates and the other in its ontology. If the one could be cleared, the other would still remain.

To see this, let us suppose that our idealist is a bundle theorist who holds that the relation of coexemplification or compresence is primitive. This view has had a long history in philosophy and is not without its plausibility. But under such a view, the circularity in the predicates can be removed, for a material thing's being red can be analysed in terms of the coexemplification of the identity property for the material thing and the property of being red. Yet surely, even if we accept the primitiveness of coexemplification, we are not going to be happy with the proposed reduction. The circularity involved in positing the identity property remains.

As far as I can see, there are no other plausible ways of distinguishing between the reductions of the actualist and the idealist; and so our second challenge remains unanswered.

There is a rather different line of response that Plantinga might make to our arguments and that is perhaps implicit in his defence of the necessary existence of properties. Under this response, it is maintained that the identity properties for possible individuals are all actual or existent. For, firstly, it is necessary for each (actual) property that there exists an identity property for that object; and, secondly, it is necessary that each (actual) property necessarily exists. But from these two assumptions it follows that each identity property for a possible individual exists; for it will exist in the world in which the individual is actual and so, from the necessity of its existence, it will also exist in the actual world. We may bring the property down, as it

were, from the possible world to the actual world. Therefore, if it is indeed true that the identity properties for possibles are among the actuals, then the actualist should have no qualms about using them in his reduction.

It should be noticed that this response seems to sidestep our previous considerations altogether. We have, on the one side, an argument against the acceptability of identity properties in an actualist reduction and, on the other, an argument for their acceptability. But at no point do the two arguments appear to meet.

What has gone wrong? I would suggest that the question of necessary existence is irrelevant, as such, to the ontological issue. Let us *grant* that the identity properties or singleton sets necessarily exist. If the actualist is defined as one who embraces all and only existents, the identity properties or singletons should be acceptable to him. But the point remains that actualism, as so defined, is an incoherent ontological position. The charge of circularity did not turn at all upon the status of properties or sets as actual or possible; and so the charge will still apply, regardless of one's views on that question.

A label, even of a broadly ontological character, does not legitimate an exclusive commitment to the items so labelled. The commitment to certain items forces one to make a commitment to other items. The platonist, for example, is committed to the members of his sets or to the constituents of his properties regardless of whether these members or constituents are themselves abstract. In the same way, the actualist is committed to the members of sets or to the objects involved in properties regardless of whether they are actual or not.

Indeed, it is not just that the assumption of necessary existence for properties is insufficient to make the proposed reduction acceptable. The assumption of contingent existence for properties also fails *in itself* to make the reduction unacceptable. Of course, it would then be unacceptable to a thoroughgoing actualist, one who scorns possibles altogether. But we can imagine a more discriminating form of actualism, that lets in some of the possibles, just as we can allow for a qualified nominalism that lets in some of the abstract objects. On such a view, the merely possible properties might be considered acceptable, but the merely possible individuals not.

The question of circularity is at least separable in principle from the question of necessary existence. The fundamental objection to be made against the reduction is not that the essences are contingently existing entities but that they presuppose the very objects, the possible individuals, whose status is in question. To this extent, then, my criticisms of the reduction in [1977*a*], 125–30, and Ch. 4 §4 above, are misdirected; and so too are Plantinga's arguments for necessary existence in [1979] and [1983], at least in so far as they are construed as a defence of the reduction.

What encourages confusion here is a certain use of the phrase 'ontologic-ally dependent'. One naturally takes this phrase to signify presupposition as

previously explained, so that y is ontologically dependent upon x when its *identity* is dependent upon that of x. But philosophers, including Plantinga (see [1979], 111), have been tempted to define the phrase in terms of modal dependence, so that y is ontologically dependent upon x when its existence depends upon that of $x(\square(Ey \supset Ex))$. Thus when it is asked whether y is dependent upon x, the question is read as one concerning ontological dependence but answered as one concerning modal dependence. It would be better if the two senses of 'dependence' were sharply distinguished, perhaps by means of the terms 'ontological' and 'modal' as I have used them.

But although considerations of ontological dependence and the charge of circularity, in particular, can stand on their own, they tie in naturally with considerations of modal dependence and with the objection from contingent existence. It should also be noted that conclusions concerning necessary existence have a direct bearing on how we should view actualism and the proposed reduction. If it is concluded that the identity properties exist necessarily, then actualism will not be a coherent ontological position. On the other hand, if it is concluded that some of them exist contingently, their use in the reduction will be objectionable to the thoroughgoing actualist not just for presupposing possibles but for being possibles themselves.

For these reasons, and also because of its intrinsic interest, we shall take up the topic of the necessary existence of properties in the next three sections.

3.   The Argument Against Necessary Existence

Do properties exist necessarily? In considering this question, let us take the existence of properties for granted. So the only question is whether, among the properties that exist, there are those that exist contingently.

Our discussion will be in two parts. In the first, we shall put forward an argument for the contingent existence of certain properties; in the second, we shall try to meet an argument of Plantinga's against the contingent existence of properties.

The question of the existence-conditions for properties would seem to depend crucially upon one's conception of properties. Let us first of all consider this question under the aspect of what one might call the *objectual-structuralist* stance on properties; later, we shall generalize the answer. According to structuralism, properties are either simple or else are structured complexes, built up in a definite way from their constituents. According to *objectual* structuralism, individuals, as opposed to their intensions, may occur as the constituent of properties. An identity property, for example, would most naturally be regarded as a construct from the identity

relation, obtained by plugging the given individual into one of its argument places.[11]

Under such a conception of properties, it is very plausible to hold to what might be called the *contingency thesis*, namely that some properties exist contingently. If the argument is spelled out, it goes somewhat as follows.[12] First, under structuralism it is plausible that:

> *Existence* A property exists in a world only if all its (individual) constituents do.[13]

For since a property is actually built up from its constituents, it is difficult to see how it can exist unless they do. Under structuralism, it is also plausible that:

> *Rigidity* If a property has an (individual) constituent in one world it has that constituent in any world in which it exists.

For the structure of a property is not something that changes from world to world; it is, if you like, a *rigid* feature of the property. Finally, under Objectualism one will have:

> *Objectualist Premiss* Some property has a contingently existing individual as constituent.

The property of being identical to Socrates is an example.

From these three assumptions it follows that some property exists contingently. For, given the Objectualist Premiss, we may take a property with a contingently existing individual as constituent. Now go to a world in which that individual does not exist. Then the property will not exist in that world either; for otherwise, it would still have the individual as a constituent in the world, by Rigidity, and so the individual would exist in the world after all, by Existence.

Plantinga is well aware of this argument. He mentions it in [1979], 111, and attempts to meet it in [1983] 7–9. In one respect, though, his discussion is rather curious; for while he rejects the argument as stated for properties, he accepts a similar argument for sets, but with membership in place of constituency (see [1976] or Loux [1979], 260). Thus we find him rejecting the Objectualist Premiss ('If an abstract object like a proposition has constituents, wouldn't they themselves have to be abstract?' [1983], 9), and also

---

[11] Such a stance is also described in Plantinga [1983], 9–11; though it is there tied (unnecessarily) to the view that the role of proper names in ordinary language is to pick out the objectual constituents of proposition. There is a further discussion of the stance in my [1980].

[12] This argument is stated, in a somewhat confused form, in my [1977a] §4; it is also essential to the development of the modal theory of propositions in [1980].

[13] Plantinga jokingly calls this, or something like it, 'existentialism' (see [1979], 115 and [1983], 1–3). We avoid terminological trespass by using the term 'Existence Principle'.

Rigidity ('Perhaps it [the proposition without the constituent] would have been slightly ill-formed or even maimed; but couldn't it exist nonetheless?' [ibid.]).

But there seems to be little or no basis for discriminating between sets and properties in this way. The premises of each argument have about equal force. If it is asked how a person can be a constituent of a property, the answer is: in the same way that a person can be a member of a set. And if it is suggested that a property might exist in a maimed or ill-formed state, then the proper response is that the structure of properties is rigid and that a property without one of its constituents could no more retain its identity than could a set without one of its members.

Indeed, the premises in the two arguments have their basis in the same underlying intuitions. They derive their plausibility from being instances of more general truths concerning the relationship of a constituent to a complex.

There are perhaps only two considerations that might favour the set-theoretic as opposed to the property-theoretic premises. The first is that properties, unlike sets, are not structured entities at all. Plantinga talks of the obscurity in the notion of constituent ([1983], 7); and perhaps he has something like that in mind. But it strikes me that, on the contrary, the general notion of structure or of constituent is as clear in its application to properties as to sets. Let it be noted that the view of sets as having members as *constituents* is no part of the mathematical theory of the subject, but arises from philosophical reflection. We recognize a certain distinctive operation, the set-builder, that *constructs* or *builds up* a set from its members. But there is no reason a priori to suppose that this is the only such operation; and it is equally plausible that relational properties, let us say, should be built up from subjects and relations by means of a characteristic operation of plugging in. Any greater obscurity in the case of properties, it seems to me, attaches not so much to the idea that properties have structure as to the problem of saying exactly what that structure is.

But still, it must be admitted that my view is not a common one and that the structuralist conception of properties has in recent times been held in increasing disfavour. For this reason alone, and quite apart from the merits of the case, it will be important for us later to develop an argument for contingency that makes no appeal to structure.

The other consideration concerns intensionality. Properties, it may be argued, are intensional entities and, as such, can have only intensional entities as constituents. This view has had a long history and is most commonly, though perhaps mistakenly, associated with the name of Frege. But whatever the plausibility of the view, and this is not the place to debate its merits, it is hard to see how in the present context it can be maintained. For given the falsehood of the Discernibility Doctrine, there will be identity

properties not necessarily coextensive with any purely qualitative property. So either the intelligibility of such properties must be denied, in which case the reduction would have to be given up, or else the Fregean view itself must be given up.

Some minor differences aside, the argument for contingent existence is about as plausible in the case of properties as in the case of sets. But still the question arises as to its cogency in either case. And here I must admit that, even though I accept the argument, it is not as compelling as one might like.

Consider a parcel of matter c made up of two other parcels, a and b. It then seems absolutely clear that c cannot exist unless a and b exist. What else could the existence of c consist in? However, when we turn to the claim that a property cannot exist without its constituents or a set without its members, the intuition, though still there, is less firm.

What appears to account for this infirmity is the presence of an opposing intuition, namely, that sets and properties exist necessarily. We are inclined to think that as abstract entities, set in a platonic heaven, there is no way in which their existence can be circumscribed by empirical circumstance. What can cabbages, or even kings, have to do with the existence of sets and properties?

In so far as one is pulled in the direction of the opposing intuition, as Plantinga clearly is, the premises of our argument will seem less secure. But where can they be wrong? It is most tempting, I think, to lay the blame on the notion of constituency. When we say that properties or sets are 'built up' from their constituents, this is a mere metaphor. We are then taken in by this metaphor and assume, as in the physical case, that the existence of the parts is essential to the existence of the whole. Plantinga does not himself mention this point, but it might be treated as a relevant respect in which the notion of constituency is 'obscure'.

I myself am not moved by this objection. It seems to me that there is a clear and general sense of constituent, that applies equally well to abstract and concrete entities and that conforms, in either case, to the Existence Principle. But still, it may be worthwhile to try to restate the argument in a way that avoids all appeal to the notion of constituency. One could thereby undermine the present objection; and one could also avoid the dependence on a view, the structural conception of properties, that many have found so unattractive.

When one is asked to give a broader account of the relationship of constituents to a complex, it seems natural to say that the constituents are required in order to explain the identity of the complex, that, in our previous phrase, the complex *presupposes* the constituents. Thus to say what a particular set is, we must say what its members are and how they are put together by means of the set-builder to form the set; while in explaining what

an identity property is, we may say that it is formed from the identity relation and the given individual by means of the operation of plugging in.

It is now a natural actualist requirement that only existents should be required to explain the identity of existents. Call this the *Generalized Existence Principle*. What it says, granted that the presuppositions of an object are essential to it, is that any case of ontological dependence is a case of modal dependence.

The Existence Principle itself can then be seen to have its basis in this more general requirement. It is because constituents are required to explain the identity of the complexes, that the complexes cannot exist without their constituents. The considerations concerning constituents can likewise be bypassed in our argument for contingent existence; for in that argument the notion of presupposition may be substituted throughout for the notion of constituent. Singleton sets and identity properties will not exist without their members or bearers, since otherwise there would be no explanation of their identity.

An interesting aspect of the present argument is how well it ties in with our previous objection to the Plantinga-style reduction. The heart of that objection was that a given entity (the identity property) should belong to an ontology only if any entity required to explain its identity (the individual) belongs to the ontology. The present argument turns on the actualist requirements that an entity belong to the domain of existents only if any entity required to explain its identity also belongs to that domain. Thus we can see the actualist requirement as arising from the desire that the existents should be capable of forming an ontology, from the desire, if one likes, that actualism should be a coherent ontological doctrine.

Another interesting aspect of the new argument is its independence from any particular conception of sets or properties. The discussion has so far been predicated upon a structuralist conception. But what if one were to hold to the identity of necessarily coextensive properties or to some identity criteria intermediate, in strength, between the sharing of extensions and the sharing of structure?

It would then be unclear what the existence conditions for properties should be. Take, for example, the property of being identical to Socrates under the identity criterion of necessary coextensiveness. This property is, let us say, necessarily coextensive with and hence identical to the property of being a person that issues from such and such an egg and sperm cell. But which specification of the property do we take to determine the existence conditions of the property? The one requiring the existence of Socrates, the one requiring only the existence of the sperm and egg, or neither?[14]

---

[14] It is therefore misleading for Plantinga to claim, as he does in [1979], 111, that quiddative properties, those involving individuals, will be necessary existents if 'essences can be constructed

It is commonly thought that principles such as identity under shared extensions adequately explain the identity of properties. But this is not true. The principle concerning shared extensions no more explains the identity of properties than does the extensionality principle explain the identity of sets. For just as the latter is compatible with sets being like boxes and changing their members from world to world, so the former is compatible with properties being like blanks for structural properties, which may change in what fills the blank from world to world in any way compatible with the given extensions.

An answer to our question about existence conditions depends upon a more exact determination of the identity of properties. But even without such a determination, it is possible to say something about the existence conditions for properties. For it is in the spirit of the Generalized Existence Principle that an entity should exist in a given world only if the means of explaining the identity of that entity exists in the world. In explaining the identity of an existent, there should be no need to make a detour through the non-existent. Now, at the very least, an explanation of the identity of the property should require a specification of its application conditions, of the conditions under which it is had by an object. So in each world there should in principle be a formula $A(x)$, constructible from the names for individuals in that world, for which it is necessary that an object have the property if and only if it satisfies the formula.

Given such a minimal condition for the existence of properties, it is then exceedingly plausible that some properties should exist contingently. This does not quite follow from the denial of the Discernibility Doctrine; for even though a property has no actualist counterpart, one involving actual individuals alone, there may, in each world, exist individuals that suffice to specify its application conditions. This, though, would be a kind of modal freak. Once given the denial of the Discernibility Doctrine, there is little reason for ruling out properties whose application conditions cannot be specified in certain worlds and which therefore fail to exist in those worlds.

Resort to anti-structuralism therefore offers no escape from the contingency thesis.

4.   The Argument for Necessary Existence: The Classical Response

In addition to criticizing the arguments against the necessary existence of properties or propositions, Plantinga puts forward an argument of his own

from qualitative properties.' For on a criterion in which an identity property and a necessarily coextensive qualitative property are distinct, it may be argued that the identity property exists contingently even though its qualitative counterpart exists necessarily; and on a criterion in which the two are the same, it is not clear, as I have pointed out, which specification of the property is to be taken to determine the existence conditions.

in favour of their necessary existence. This first appears in [1979], and is later elaborated in [1983].

One version of the kind of argument Plantinga wishes to propound goes as follows:

(1) Possibly Socrates does not exist;
(2) Necessarily, if Socrates does not exist then the proposition that Socrates does not exist is true;
(3) Necessarily, if the proposition that Socrates does not exist is true, then the proposition that Socrates does not exist exists;
(4) ∴ It is possible that Socrates does not exist and the proposition that Socrates does not exist does exist.

Or, in symbols:

(1) $\Diamond -Ea$
(2) $\Box(-Ea \supset T\xi -Ea)$
(3) $\Box(T\xi -Ea \supset E\xi -Ea)$
(4) ∴ $\Diamond(-Ea \wedge E\xi -Ea)$,

where 'ξA' denotes the proposition that A.

I say this is a version of the argument because Plantinga's actual argument, as presented in [1983], 9–10, is somewhat different. He argues first from 'Possibly Socrates does not exist' ($\Diamond -Ea$) to 'The proposition that Socrates does not exist is possible' ($P\xi -Ea$, for **P** the possibility-predicate), and from the latter statement to 'Possibly the proposition that Socrates does not exist is true' ($\Diamond T\xi -Ea$). This then justifies the inference from the first to last statement. Now the intermediate steps will be of little independent interest, and the mediate inference, from 'Possibly Socrates does not exist' ($\Diamond -Ea$) to 'Possibly the proposition Socrates does not exist is true' ($\Diamond T\xi -Ea$), must be seen to depend upon the more general principle that necessarily if S (S a sentence) then the proposition that S is true ($\Box(S \supset T\xi S)$). Thus the effect of the first stage of Plantinga's argument can be achieved by adopting (1) and (2) as premisses. But once this is done, his premiss (7) becomes redundant and we obtain his conclusion with a single additional premiss (my (3), Plantinga's (6)).

In fact, Plantinga's original argument has a precursor. Drop the intermediate steps mentioned above and we almost obtain the argument of my [1977], 149–50 (Ch. 4 §4 above), which, in its turn, was derived from Prior. The only difference is that the conclusion is inverted with the first of the premisses, thereby illustrating the maxim that one man's modus ponens is another man's modus tollens.

The conclusion of the argument is not quite what Plantinga wants. This is that all propositions necessarily exist, or better, that necessarily they

necessarily exist. But the given conclusion merely states that a particular proposition can exist in certain circumstances. This difference is not too serious, though; for we may take it that if the given conclusion holds, then the principal and perhaps sole objection to the stronger conclusion has been met. Indeed, independent arguments for the stronger conclusion are readily constructed. From $\Box(S \supset T\xi S)$, $\Box(T\xi S \supset E\xi S)$ and $\Box(E\xi - S \supset E\xi S)$, S any sentence, it follows that $\Box E\xi S$; and similar arguments can be constructed to establish $\Box\forall\rho\Box E\rho$.

A more serious shortcoming in the conclusion is that it only relates to the necessary existence of propositions, not of properties. But it is the necessary existence of properties that most concerns Plantinga, and it is this that is apparently required for the success of his reduction. Now normally this discrepancy would not be thought to be of great moment. Properties are the same kind of things as propositions; what goes for one, in regard to existence, should go for the other. But it is not clear that Plantinga can afford to be so cavalier in the matter. After all, properties are in some respects more like sets than propositions. So why not let them share in the contingent existence of sets rather than the necessary existence of propositions?

There is a more specific argument for tying the existence of properties to that of propositions. It is this: a property φ exists if a proposition constructed from φ, say that $\varphi = \varphi$, exists; so if the proposition exists necessarily, the property does too. But Plantinga cannot avail himself of this argument; for the plausibility of the first premiss depends upon something like the Existence Principle, that a complex can exist only if its constituents do. A similar doubt would seem to infect the premiss $\Box(E\xi - S \supset E\xi S)$ in our more general argument for the necessary existence of propositions, a premiss that Plantinga elsewhere cites with approval.

Nor is it easy to develop an argument for properties, independent of the argument for propositions but comparable to it. The most natural analogue goes as follows (with η perhaps the degenerate property $\lambda x - Ea$ of a's not existing):

(5) $-Ea$
(6) $\Box(-Ea \supset aH\eta)$
(7) $\Box(aH\eta \supset E\eta)$
(8) $\therefore \Diamond(-Ea \wedge E\eta)$

But Plantinga is debarred from accepting the second premiss by his espousal of the doctrine, later to be discussed, that all properties are existence-entailing. Let us put these difficulties on one side, though. The argument for the necessary existence of propositions retains its interest, regardless of what goes for properties; and the difficulties might, in any case, be somewhat mitigated by Plantinga's abandoning his differential stand on properties and

sets. At the very least, he would then be better able to appeal to some general analogy with propositions in support of the necessary existence of properties.

I shall consider two objections to Plantinga's argument. One charges it with equivocation, the other with incoherence. My reason for considering each of the two responses is rather different. I consider the first because it is right, the second because of the light it can throw on Prior's philosophy of modality. Plantinga's evaluation of the two responses is rather different from mine. He regards the second as the more serious; the first does not even appear in [1979] and is relegated to second place in [1983]. This suggests that the only real alternative to the Plantinga stand on propositions is the Priorian stand on modality. But I hope it becomes clear that no such choice need be made.

Now there is no question as to the validity of the argument, at least if the modal notions are taken in their usual sense. So the only question concerns the truth of the premisses.

But there is here an obvious and immediate objection. One should distinguish between two notions of truth for propositions, the *inner* and the *outer*. According to the outer notion, a proposition is true in a possible world regardless of whether it exists in that world; according to the inner notion, a proposition is true in a possible world only if it exists in that world. We may put the distinction in terms of perspective. According to the outer notion, we can stand outside a world and compare the proposition with what goes on in the world in order to ascertain whether it is true. But according to the inner notion, we must first enter with the proposition into the world before ascertaining its truth.

Now if truth bears the inner sense, the third premiss holds and the second fails; while if truth bears the outer sense, the second premiss holds and the third fails. There is, however, no single sense of truth for which both premisses hold. The argument rests on a fallacy of equivocation.

Even though this objection is so obvious and is, in my opinion, the most fundamental of the objections to the argument, it is not one that Plantinga mentions, either in [1979] or in [1983]. This is because he treats the argument as if the sense of its key terms were unproblematic. Thus the only objections he considers are those that relate to the truth of particular premisses.

This makes it hard to know what Plantinga would want to say in response to the objection. It is, however, possible to reconstruct some sort of answer. For we may take it that there is a sense, indeed a single sense, of truth for which the second premiss holds. Fix on that as the intended sense of truth. Then the objection is one that is directed towards the third premiss.

This is, indeed, an objection that Plantinga considers, not in [1979], but in [1983]. He there tries to show that the truth of the third premiss is required by what he calls *Serious Actualism* and which I would prefer to call *Property*

Actualism. (After all, it is not as if actualism itself were not serious.)[15] This principle plays a prominent role in Plantinga's thought. It is invoked in [1974], 150, under the heading 'Restricted Ontological Principle', and also in [1979], 109, and [1983], 11. As stated in the latter, it says that 'for any object x and property p, it is not possible that x should have had p but not existed'; or, in symbols:

(5) $\forall x \forall \varphi \Box (x H \varphi \supset Ex)$[16]

For the moment, let us leave aside the question of whether the third premiss follows from the principle. It is still far from clear that the principle itself is true. The issue may be put in terms of the property of not existing. Let us grant that possibly Socrates does not exist. Then the question is whether we can infer from this that possibly Socrates has the property of not existing.

The issue may be stated in more general terms as one about the proper extension of the abstraction principle to modal contexts. Let $\lambda x A(x)$ denote the property of being an x such that $A(x)$.[17] The classical principle of abstraction states that:

(6) for any x, x has the property $\lambda x A(x)$ iff $A(x)$ ($\forall x(x H \lambda x A(x) \equiv A(x))$)

Clearly it is intended that this principle hold of necessity.

(7) $\Box \forall x(x H \lambda x A(x) \equiv A(x))$.

But this leaves open the question of how $\lambda x A(x)$ behaves in regard to the non-existents of each world. One answer is provided by the *unrestricted principle of abstraction*:

(8) Necessarily, for any x it is necessary that x has $\lambda x A(x)$ iff $A(x)$ ($\Box \forall x \Box (x H \lambda x A(x) \equiv A(x))$).

---

[15] Even what Plantinga calls 'actualism' is not properly so-called. He takes it to be the doctrine that everything there is exists or, better, *necessarily* everything there is exists ($\Box \forall x Ex$). But when the actualist speaks with the possibilist he rejects this claim; it is just that he then renders it in other terms. There seems to be a common confusion here. For example, the idealist is commonly represented as saying that everything is mental. But when he speaks with the common man he rejects this claim. How best to represent ontological claims is a large and difficult topic. But for the purposes of this chapter, actualism may be taken to be the view that the quantifiers in the unreduced or most basic language are actualist. It is for them that $\Box \forall x Ex$ holds.

[16] A better formulation of what Plantinga has in mind would be $\Box \forall x \Box \forall \varphi \Box (x H \varphi \supset Ex)$. The interlacing of quantifiers and modalities is something that is often omitted in the formulation of modal theses; but it need not concern us here.

[17] Here, and elsewhere, I ignore the complications that arise from $A(x)$ containing free variables other than x.

This would support the inference from 'possibly Socrates does not exist' ($\Diamond-Ea$) to 'possibly Socrates has the property of not existing' ($\Diamond aH\lambda x-Ex$). Another answer is provided by the *restricted abstraction principle*:

> (9) Necessarily, for any x it is necessary that x has $\lambda xA(x)$ iff A(x) *and* x exists ($\Box\forall x\Box(xH\lambda xA(x) \equiv A(x) \wedge Ex)$).

What Restricted Abstraction adds to the neutral principle (7) is just the application of Property Actualism to abstracts:

> (10) $\Box\forall x\Box(xH\lambda xA(x) \supset Ex)$.

Under this principle, the previous inference would not go though.

There is no doubt that property abstracts or the copula might be so construed that they conformed to Restricted Abstraction. Property abstracts might be understood to have a built-in existential presupposition, as signifying something like $\lambda x(Ex \wedge A(x))$, so that, regardless of one's views on the unadorned abstract, it would follow that the adorned abstract conformed to Property Actualism. Alternatively, the copulative relationship $xH\varphi$ might be thought to carry an existential load, to have Ex and Eφ as part of its meaning, so that, regardless of one's view on the abstracts, it would again follow that Property Actualism held.

But although one *might* use property abstracts or the copula in this way, there is, as I insisted in Fine [1976], 564, a perfectly intelligible use in which they conform to the unrestricted principle. One can so understand property talk, that to say Socrates has the property of not existing is to say no more, in modal contexts, than that Socrates does not exist. It is then trivial, if it is possible that Socrates not exist, that it is possible that Socrates possess the property of not existing.

This use of property abstracts, in which 'x has $\lambda xA(x)$' says no more than A(x), is not only perfectly intelligible: it also appears to be that most appropriate way of understanding abstracts. We want to know the application conditions for the complex property $\lambda xA(x)$, the conditions under which it is true of an arbitrary object a. These conditions are most naturally provided by the statement A(a). Thus what we want to assert is $\Box(aH\lambda xA(x) \equiv A(a))$, not $\Box(aH\lambda xA(x) \equiv (A(a) \wedge Ea))$, as would be required by the restricted principle. Indeed, from this perspective, there seems to be something quite arbitrary and gratuitous in adding Ea to the application conditions. If not Ea, why not $-$ Ea or some other statement altogether?

For Plantinga, there is an additional difficulty. One wants to be able to express the application conditions for the property $\lambda xA(x)$ not merely in terms of the statement A(a), but also in terms of the truth of the proposition $\xi A(a)$. One wants to accept not merely $\Box(aH\lambda xA(x) \equiv A(a))$, but also $\Box(aH\lambda xA(x) \equiv T\xi A(a))$. Since we accept Abstraction for propositions:

$\Box(T\xi A \equiv A),$

the one will follow from the other. But since Plantinga also accepts propositional abstraction, both will fail together.

The arguments against Property Actualism appear strong. So what reasons can Plantinga have to accept it? It seems there are mainly two. The first is that the principle follows from what Plantinga calls Actualism, the doctrine that necessarily everything exists ($\Box\forall xEx$) (see [1979], 108–9). Such an argument would indeed have added lustre to the principle; but Plantinga has since conceded that it is wrong ([1983], 11–12).

There is no need to go into the error, but we may note that the confusion between the two forms of actualism seems common. One finds Adams ([1981], 7, 18), for example, lumping the two together in the account of actualism. It should be recognized, however, that there are two independent doctrines here. One can require the quantifiers to be actualist and yet allow the predicates to be true of non-existents, and one can allow the quantifiers to be possibilist and yet require the predicates to be false of non-existents.

The confusion may have no deeper source than the conflation of $\forall x\Box(xH\varphi \supset Ex)$, which is independent of Actualism, with $\Box\forall x(xH\varphi \supset Ex)$, which is a trivial consequence of it. But there is a deep reason why both doctrines are called 'actualist', why both are placed under the broad umbrella of actualist opinion. We have a picture of the world as consisting entirely of actuals in relation to one another. The application of the relations to non-existents must then at least be irrelevant to the identity of the world; it cannot be that two worlds agree on existents but differ on non-existents. But this condition is most naturally secured by having the relations *false* of the non-existents. So there is some affinity in the two doctrines, if not a logical connection.

Plantinga's second reason for espousing Property Actualism is based upon the distinction between predicative and impredicative propositions ([1974], 149–51, [1983], 13–14, and elsewhere). *Predicative* propositions predicate a property of their subject (are of the form $aH\varphi$, let us assume); *impredicative* propositions do not. Now an alleged counterexample to property actualism is provided by the proposition that Socrates does not exist ($\xi-Ea$); for the proposition can be true even though Socrates does not exist. But Plantinga argues that this is no counterexample since the proposition is not predicative in form, it does not predicate a property of its subject.

There is some sort of shift here between making a statement such as '$aH\varphi$', and declaring that the corresponding proposition is true ($T\xi aH\varphi$). But this is harmless enough if abstraction holds for propositions. If the point is made directly in terms of statements, it is that the sentence 'Socrates does not exist' ($-Ea$) fails to provide a counterexample to Property Actualism since it is not of the form '$aH\varphi$'.

We may put the point in terms of the distinction between Property Actualism:

   (5) $\forall x \forall \varphi \Box (x H \varphi \supset Ex)$;

and what one might call *Formula Actualism*:

   (6) $\forall x \Box (A(x) \supset Ex)$, for any formula $A(x)$.

(Formula Actualism would seem to correspond to Plantinga's Unrestricted Ontological Principle in Fine [1977a] 151.) There is no doubt that Formula Actualism is false; and this may be shown by means of the formula $-Ex$. But from this it does not follow that Property Actualism is false, since the offending formula $-Ex$ is not of the form '$xH\varphi$'.

So Plantinga's objection to the counterexample is correct. But to dispose of a bad reason against a view is hardly to put forward a good reason for it. After all, it is not as if rejection of the principle had to be based upon the kind of confusion Plantinga alludes to. We may concede that the statement 'Socrates does not exist' is not predicative in form and yet still maintain, for the reasons already given, that it is necessarily equivalent to a statement that is predicative in form. The counterexample would then go through, Plantinga's distinction notwithstanding.

We see that Plantinga has not really provided any good arguments in favour of Property Actualism; and, indeed, in [1983], 14–15, he himself comes very close to giving the principle up. He there seems prepared to concede that 'conditions', as opposed to properties, might conform to Unrestricted Abstraction. Now I take it that conditions are not merely formulas but, like properties, are some kind of abstract object. But then the distinction between properties and conditions becomes quite tenuous. Properties behave just as one would expect conditions of the form $\lambda x(Ex \wedge A(x))$ to behave. Is it really to be supposed that the two kinds of entity are distinct, or is it not rather that one is merely a special case of the other?

Plantinga admits, it is true, to some extra reservations over conditions. He writes, 'it isn't at all easy to see what sort of thing a *condition* is, or to state the conditions under which an object meets a condition in a world' ([1983], 14). But these reservations are easily taken care of. The question of the identity of the complex $\lambda x A(x)$ is largely independent of the issue over abstraction and so whatever view one takes of the identity of properties may readily be transferred to the case of conditions. As for the application conditions, these are as well specified by Unrestricted Abstraction for conditions as they are by Restricted Abstraction for properties.

I have not actually disproved the principle of Property Actualism; but the case against it seems strong, the case for it weak. In so far, then, as the charge of equivocation can be met only by appeal to the principle, the argument also becomes weak. This is perhaps somewhat obscured by the line of

reasoning in [1979], since there the principle is still thought to be derivable from actualism *per se*. But with the admission of underivability in [1983], the vulnerability of the argument becomes clear.

But matters are worse than that. *Even with the principle*, the argument will not go through. There is an obvious gap in the inference from principle (5) of Property Actualism to premiss (3) of the argument. The antecedent in (5) predicates a property of its subject; it is of the form $a\mathrm{H}\varphi$. On the other hand, the antecedent of (3) merely attaches a predicate to its subject; it is not of the form $a\mathrm{H}\varphi$, but of the form Pa. In (3) we merely talk of a proposition being true. But application of (5) requires that we talk of a proposition having the property of being true.

It is surprising that Plantinga himself is not more alive to this difference, since it is the very kind of difference that he points to in disposing of the alleged counterexample to Property Actualism. But perhaps I have misunderstood his position in some way. The application of the principle requires that the antecedents in (3) and in (5) have the same logical form. Maybe I have got the form of either or both of these antecedents wrong. One possibility is that Property Actualism should be left alone, but that, in the original formulation of the argument, 'The proposition is true' (Tξ−Ea) should be replaced throughout by 'the proposition has the property of being true' (ξ−EaHτ, for τ the property of being true). The application of Property Actualism is then unproblematic; but the truth of the second premiss ($\Box(=\mathrm{E}a \supset \xi-\mathrm{E}a\mathrm{H}\tau)$) becomes doubtful. In the original formulation of the argument, Plantinga had to rely on our accepting a sense of truth for which the second premiss held. Now it does indeed seem reasonable that if it is possible that Socrates does not exist ($\Diamond-\mathrm{E}a$) then, for some sense of truth, it is possible that the proposition that Socrates does not exist is true ($\Diamond\mathrm{T}\xi-\mathrm{E}a$). But given that predications of truth are to be subject to Property Actualism, why also accept that it is possible that the proposition that Socrates does not exist has the property of being true ($\Diamond\xi-\mathrm{E}a\mathrm{H}\tau$)? The intuitive appeal that gave us one will not also give us the other.

A more interesting possibility lies in realigning the antecedent of the principle, not the premiss. It will not do to formulate Property Actualism with arbitrary formulas as antecedents: that gives us Formula Actualism, which is false. But there is an alternative. Formulate the principle with statements Px as antecedent, where P is now either a simple or a complex predicate. It has hitherto been supposed that predicative statements are of the form '$a\mathrm{H}\varphi$', with explicit copula; but perhaps they are best taken to be of the form Pa, with implicit copula. It will then be to predicative statements in this sense that Property Actualism applies.

It is hard to be sure what kind of predicative statement Plantinga has in mind; his verbal formulations tend to obscure the relevant difference. But whatever his intentions on this point, it seems worth exploring the

possibility that predicative statements are to be taken as having an implicit copula.

When they are taken in this way, Property Actualism becomes:

(11) $\forall x \ \Box(Px \supset Ex)$

for P a simple or complex predicate (perhaps we can quantify over the P).

We might call the principle in this form *Predicate Actualism*. Now let us concede the coherence of predicational forms Pa with complex P, although this raises difficult questions in philosophical logic; and let us grant the truth of Predicate Actualism. The application of the principle to (3) then appears unproblematic, since the antecedent of (3) is of the form Pa.

But the matter is not so simple. Suppose I were to argue as follows: I use the phrase '*x* is a non-existent' (N*x*) simply as an *abbreviation* for the longer expression 'it is not the case that *x* exists' (−E*x*). Plugging in '*x* is non-existent' (N*x*) into Predicate Actualism then gives $\forall x \Box(Nx \supset Ex)$. But N*x* abbreviates −E*x*. So we obtain $\forall x \Box(-Ex \supset Ex)$, from which it follows that everything necessarily exists ($\forall x \Box Ex$).

To this argument, Plantinga would have to reply that the statement '*x* is non-existent' (N*x*) is not really predicative in form. In abbreviating the statement 'it is not the case that *x* exists' (−E*x*), one has let the abbreviation, perhaps somewhat misleadingly, assume the logical form of a predicative statement. But the logical form of an abbreviation is given, not by its outward appearance, but by the logical form of the statement that it abbreviates.

So appearances can be deceptive. An expression can merely pose as a predicative statement. But if the expression for non-existence can so pose, why cannot the expression for truth? It is not necessary, of course, that the phrase 'is true' should actually be used in ordinary language as an abbreviation for a complex formula. It suffices that it should behave, in regard to its logical form, as if it were such an abbreviation.

But is this plausible? The notion of non-existence bears its complexity on its face; but is it plausible to suppose that the concept of truth submits to analysis? We now seem to have a firm basis for the application of Predicate Actualism. There must be a sense of truth for which the second premiss holds; this sense is simple; and so it must be expressible by a predicate (rather than an abbreviatory device) to which the principle will be applicable. Indeed, we do not need the full principle here, but merely its restriction:

(12) $\forall x \Box(Px \supset Ex)$, for P a simple predicate.

Now I am inclined to accept the restricted form (12) of Predicate Actualism. But I do not believe that the assumption of the simplicity of truth can be sustained. It is possible that there is no simple predicate of truth, but that any such predicate is to be defined by a Tarski-style recursion on propositions.

But let us grant, for the sake of argument, that there is a simple truth-predicate. By Predicate Actualism, this will be existence-entailing and so will express what I previously called an inner concept. However, it will now be possible to find a rather long formula that defines the outer concept of truth in terms of the inner concept. (This is by no means easy; see my [1980] §8, for details).

So there will be two concepts of truth: the inner, which is expressed by a predicate; and the outer, which is expressed by a formula. The contingency theorist can now argue that the concept of truth for which the second premiss holds is one that is expressed by a formula or by a pseudo-predicate for that formula. Thus even if the truth of Predicate Actualism is granted, its application in obtaining the third premiss cannot be guaranteed.

The application of some form of Property Actualism is again forestalled; and the attempt to justify the third premiss on its basis would seem to fail. It should be noted that a similar fate befalls other attempts to use the principle. On p. 5 of [1983], for example, Plantinga argues: 'sets with contingent members, for example, are not necessary beings—not, at least, if serious actualism is correct. For if it is, then if Quine had not existed, Quine's singleton would not have contained him.' But to this an opponent can legitimately reply that even if the truth of Property or Predicate Actualism is granted, its application in this case is not justified, since membership is being used as a part of a non-predicative expression.

The principle is in general of little help in coming to a substantial modal conclusion. The essential poverty of *Property* Actualism can be appreciated once it is observed that its truth would arise merely from the decision to restrict the term 'property' to conditions, in Plantinga's sense, of the form $\lambda x(Ex \land A(x))$. It is clear that nothing of any importance could follow from such a stipulation; and so no more can follow from the principle. Of course, the reasons for holding the principle may go beyond a mere stipulation; but the consequences will be the same, whatever the reasons.

The principle of *Predicate* Actualism may be more fruitful. Its most secure applications would depend upon judgements of simplicity, though, since without the simplicity of the alleged predicate it would always be possible that it was merely serving as an abbreviatory device. However, judgements of simplicity are notoriously difficult to make. Thus one finds Adams ([1981], 11) arguing that the unique relation between an identity property and its individual is primitive. But one could with equal, or greater, plausibility argue that the relation was not primitive, but was to be analysed in terms of the construction of the identity property from the individual.

We should be frank in facing up to the basis of our essentialist opinions. A set can only exist with its members, not because of Property Actualism, but because of the constructive nature of sets. It may be satisfying to derive particular truths from general principles; but we should recognize that most

essentialist claims have their basis, not in such general principles, but in peculiarities of the subject-matter at hand.

At one point in [1983], after the discussion of 'conditions', Plantinga himself comes close to recognizing the futility of any appeal of Property Actualism. He there writes (p. 21), 'Our question is really whether *being true* is existence-entailing.' All pretence of deriving the premiss from general principles is abandoned, and we have something like a direct appeal to intuition.

Plantinga's answer to his own question is 'no' and this answer strikes him as 'obvious'. But to give this answer is completely to overlook the possibility of equivocation. What is obvious is that the second premiss holds for the outer concept of truth and that the third premiss holds for the inner concept, but the question is whether there is a *single* concept of truth for which both premisses hold. I, for one, fail to see it; and I doubt whether anyone else, who had not already prejudged the issue, would see it either.

Appeals to general principle or intuition fail to justify the third premiss. In the absence of any other countervailing reason, the change of equivocation would appear to stand.

5. The Argument for Necessary Existence: The Priorian Response

Let us now consider the other response to Plantinga's argument. This is of the kind that Prior would give on the basis of the philosophy of modality underlying his system Q. (See [1957] and [1967].) Let me repeat: I consider this objection, not because I would want to give it, but because of the light it may throw on Prior's philosophy and its relationship to Plantinga's views.

Plantinga formulates the Priorian response as ([1979], 116): 'This proposition [that Socrates does not exist], clearly enough, is possible; Socrates is not a necessary being. On the other hand, the existentialist is obliged to claim that (24) [the proposition in question] is not possibly true.' In other words, the sense in which the given proposition is possible is not one in which it is possibly true.

In [1983], Plantinga takes this response to be an objection to his premiss (5):

> (13) If the proposition that Socrates does not exist is possible, then possibly that proposition is true ($P\xi-Ea \supset \diamondsuit T\xi-Ea$).

Since this premiss merely serves as an intermediate step in establishing:

> (14) If it is possible that Socrates does not exist, then it is possible that the proposition that Socrates does not exist is true ($\diamondsuit-Ea \supset \diamondsuit T\xi-Ea$);

we may take it that the response is also an objection to (14) (or also to our second premiss, since it is that which sanctions (14)).

But it is hard to see why Prior should object to (14). Presumably in (14) the two ◇'s have the same sense. Now for Prior there are only two senses of 'possible': the weak (not always false) and the strong (sometimes true). But in the weak sense, the consequent of (14) is true (since the propositions $\xi-Ea$ fails to exist in some world); and in the strong sense, the antecedent is false (since the individual a fails to exist in some world). So in either case, (14) is true.

Indeed, if Prior is allowed to choose the interpretation of the modalities, there is no reason why he should object to the argument at all. If ◇ and □ bear the weak sense in my version (1)–(4) at the beginning of section 4, then the premisses are true and the argument a valid one of the system Q.

It might be thought odd that Prior should so readily accede to the conclusion ◇($-Ea \wedge E\xi-Ea$). But when ◇ has the weak sense, its meaning is quite innocuous. Its truth will simply fall out from the fact that Socrates fails to exist in some world; and, indeed, the above argument is then quite fatuous, since ◇($-Ea \wedge E\xi-Ea$) will follow directly from $-Ea$.

When Plantinga credits Prior with the denial of (14), he does not intend that the two 'possibles' should have the same sense, as would be required if the denial were to be an objection to the premiss (13). What he seems to have in mind is that the sense in which it is possible that Socrates does not exist for Prior is not that of 'possibly true' in its customary or intended use. So the response, if it is an objection to any of the premisses, is an objection to the first, not the second.

But what is this customary or intended sense of 'possibly true'? In [1979], 119, Plantinga treats the phrase as commensurate with 'truth in some possible world'. But that latter expression is ambiguous. Like the absolute notion, the world-relative notion of truth can have both an inner and an outer sense, the inner sense requiring existence of the proposition at the world in question, the outer sense not. Of course, someone who denies the contingent existence of propositions will dispute the distinction. But even he can make the distinction in his own terms: a proposition (or statement) will be true at a world in the inner sense conditionally upon the individuals it mentions existing at that world, and true at a world in the outer sense regardless of the existence of those individuals in the world.

When possibility is understood as truth in some possible world, it is correspondingly ambiguous. Should relative truth bear the inner sense, we obtain Prior's strong sense of possibility (true at some world in which the individuals exist). Should relative truth bear the outer sense, we obtain what we might call the *classical* or *standard* concept of possibility (truth at a world, regardless of the existence of the individuals). Thus three senses of possibility might be distinguished: the two above; and Prior's weak sense (not always false *in the inner sense*).

Now Prior accepts the first premiss of the argument for the weak sense of possibility, and there is no doubt that it fails for the strong sense. So Prior can only be taken to be denying the third premiss for the classical concept of possibility.

But how can that be? Surely it is part of our understanding of the classical concept that it should be possible in this sense that Socrates does not exist. What Prior wishes to deny, it seems, is not the truth of the first premiss but its intelligibility. The classical concept of possibility is, for him, simply not coherent.

Thus his objection is not to *any* of the premisses of the argument, but to the terms in which it is stated. We see that Plantinga has again gone astray in taking the key terms of the argument to be beyond dispute. Where Plantinga finds one sense, the first objector finds two and the second sees none. In neither case can the objector's position be sensibly construed as relating to the truth of any one of the premisses.

This is not the place to discuss all the implications of Prior's position. But we may note that an alternative formulation of it can be given in terms of the outer concept of truth. If the world-relative outer concept were coherent, then the classical concept of possibility could be explained, in the usual way, as truth in some possible world. Conversely, anyone who finds the classical concept of possibility acceptable should also find the relative outer concept of truth acceptable. In some sense, the outer concept of truth is already presupposed in the possible worlds semantics for the classical concept. But this is unlikely to convince the modal actualist, who sees the possible worlds semantics as non-basic or unreduced. There is, however, a more rigorous route from the one concept to the other. Following my [1980], §8, we may define absolute outer truth from absolute inner truth, making essential use of the classical concept of possibility, and then define relative outer truth, or its surrogate, from absolute outer truth.

So one might understand Prior as denying the coherence of the outer concept of truth. I previously tried to illuminate the distinction between inner and outer truth in terms of a difference in perspective. One might therefore also think of Prior as denying the coherence of the external or transcendental perspective on worlds. There is, for him, some kind of metaphysical illusion involved in supposing that we can stand outside all the worlds and, from this privileged vantage point, survey their various contents.

It is important to appreciate how radical is Prior's position. His views run directly counter to what is commonly presupposed in both philosophical and technical discussions of modal logic. In treating Prior's views, philosophers have usually concentrated on its more positive aspects and, especially, on the distinctions among the different kinds of modality. Thus one finds Plantinga ([1979], 116–17) saddling Prior with three kinds of possibility: the weak

and the strong; and the notion of truth in some existing possible world (one whose individuals are actual). But these positive aspects are relatively innocuous; anyone can make sense, for example, of the three kinds of possibility distinguished by Plantinga. The true force of Prior's position lies in its negative theses, in the rejection of the classical concept of possibility and the outer concept of truth. As with intuitionism in the philosophy of mathematics, the position is best understood by what its proponent refuses to say, not by what he does say.

Having got Prior's response to the argument straight, let us discuss Plantinga's objections to it ([1979] 119–21, and [1983], 15–20). There are three in all; but since they have a common form, they may to some extent be considered together. Plantinga wants to insist, first of all, that the proposition that Socrates does not exist is *genuinely* possible. He then points out that the proposition is only possible for Prior in the weak sense (sometimes not false). Finally he argues that any genuine notion of possibility must satisfy certain desiderata that Prior's weak notion fails to satisfy.

The three objections differ merely on what desideratum is specified. According to the first objection, it should not be genuinely possible that Socrates $\neq$ Socrates, though in Prior's weak sense it is.

According to the second objection, any genuine concept of possibility should be closed under logical consequence, though Prior's weak concept is not. Plantinga's example here is curious. It is provided by the inference from Socrates $\neq$ Socrates to $\exists x(x \neq x)$, with the first possible in the weak sense and the second not. But even the standard concept of possibility is not closed under the rule of Existential Introduction. It is possible that Socrates does not exist ($\Diamond -Ea$), but not possible that something does not exist ($\Diamond \exists x -Ex$). Indeed, the study of modal logic leads one to doubt whether Existential Introduction is a valid rule of inference at all.

There is, however, a better example to hand. From A $\wedge$ B one may infer A. But, under the weak concept of possibility, $\Diamond(2 \neq 2 \wedge$ Socrates $\neq$ Socrates$)$ is true while $\Diamond 2 \neq 2$ not.

The third objection is that a 'crucial insight' of the existentialist should be expressible by $-\Diamond(E\xi -Ea \wedge -Ea)$, whereas under the weak concept of possibility, this formula is simply false.

As desiderata on a genuine concept of truth, the three cases have different value. The first, that it not be possible that Socrates $\neq$ Socrates, is the most convincing. To the second, it may be objected that what is fundamental to a genuine concept of possibility is that it be closed under necessary implication, that from possibly A and necessarily if A and B, it should follow that possibly B$((\Diamond A \wedge \Box(A \supset B)) \supset \Diamond B$. Now here necessity ($\Box$) should have the sense $-\Diamond -$, cognate to the sense of possibility ($\Diamond$). But when it does, the desideratum is satisfied by the weak notion of possibility. The third desideratum is the least convincing, even though Plantinga regards this

objection as the 'clinching point'. Why should it be a requirement on a genuine concept of possibility that the existentialist insight be expressible by $-\Diamond(\mathrm{E}\xi - \mathrm{E}a \wedge -\mathrm{E}a)$? If there is a desideratum here, it is presumably because of some more basic semantic property of possibility that has yet to be specified.

Nor is it clear to me why Prior should be so concerned to admit that it is *genuinely* possible that Socrates does not exist; that it is possible in some loose or slightly improper sense should be enough. Plantinga may have misunderstood the fundamental purpose that a Priorian should have in introducing the weak notion of possibility. This is not so that it should still in some sense be possible that Socrates does not exist. Rather, it is that what is expressible using the classical concept of possibility should be expressible by other means. To this end, it is immaterial whether another concept of possibility be used or not.

There is, however, no need to get bogged down in this issue of what pertains to a *genuine* concept of possibility. Plantinga can make his objections to Prior without bringing it up. He can simply point out that there *is* a sense of possibility for which the first premiss is true and the respective desiderata are satisfied. Since for Prior there is only one sense of possibility, the weak sense, for which the first premiss is true, he must deny that there is any sense of possibility that also satisfies the desiderata. It may be conceded that weak possibility is genuine enough and yet the objections still be made.

But to these objections, as thus reformulated, Prior has a ready reply: they all presuppose what is at issue, namely, the intelligibility of the classical concept of possibility. The only reason we have to suppose that there is a concept of possibility that makes the first premiss true and satisfies any one of the desiderata is that classical possibility is such a concept. Unless we had already presupposed the intelligibility of classical possibility, we would no longer have any reason to suppose there was such a concept. If one were to argue against the intuitionist on the grounds that there must be a concept of negation that conforms to the classical rules of natural deduction, the argument would immediately be acknowledged to be circular. That there is such a concept is something to be justified in other terms, not taken for granted. But in this respect the present case is no different.

Plantinga seems to have overlooked the radical nature of Prior's position. He makes his criticisms from a perspective that is only intelligible once that position is rejected.

But to be fair to Plantinga, it may be possible to push some of his objections a little further. In the case of the first objection, though not with much force in the cases of the others, he may say: 'Look, can you not just *see* that there is a sense of possibility for which it is possible that Socrates does not exist yet not possible that Socrates ≠ Socrates. I know you will not let me explain classical possibility in terms of the outer concept of truth. But we

have here an independent way of explaining what the concept is, and one that makes it perfectly intelligible.'

I have a great deal of sympathy for this objection; there does, indeed, appear to be a concept of possibility that discriminates between Socrates' non-existence and his non-identity. But without wishing to vindicate Prior, I might point out that he does have an answer here. The question is whether there is a modal distinction between 'Socrates = Socrates' and 'Socrates exists', with one somehow necessary and the other not. But in saying Socrates = Socrates, we have already presupposed that Socrates exists; there can be no more information in the one than in the other. Therefore if it is necessary that Socrates = Socrates, it must, in the same sense, be necessary that Socrates exists.

It is worth noting that Prior is not just forced to adopt this stand because of his other views. It has in itself a great deal of plausibility and perhaps constitutes the most intuitive entry point into his overall position. The independent plausibility of the stand can be brought out by its connection with that philosophical tradition in which existence is denied to be a predicate. Prior would say that in supposing it to be possible that Socrates does not exist and yet not possible that Socrates ≠ Socrates, we had ascribed an illusory content to the existence-predicate; we had mistaken a vacuous for a non-vacuous predicate. Thus he would want to assert what others had asserted quite independently of any modal considerations.

It is surprising that Plantinga himself is not more sympathetic to this line of reasoning, for it has its roots in something very like the principles of Property or Predicate Actualism. The reason existence statements cannot arise above claims of self-identity in their content is that all statements concerning an individual are existence-entailing.

Indeed, it seems that the only reasonable defence of Property Actualism depends upon the adoption of a Priorian position. It was a difficulty in Plantinga's views that a predicative statement $x\mathrm{H}\lambda x\mathrm{A}(x)$ (or $[\lambda\, x\mathrm{A}(x)]x$) was not interchangeable in all modal contexts with the corresponding formula $\mathrm{A}(x)$. That difficulty disappears once a Priorian position is adopted; since the relevant counterexamples depend upon the use of concepts, such as non-vacuous existence, that are to be eschewed. Thus properties may be understood to behave in accordance with the principle of Property Actualism and yet the unconstrained substitution of $\mathrm{A}(x)$ for $x\mathrm{H}\lambda\, x\mathrm{A}(x)$ allowed.

It is in this way ironic that a critical element in Plantinga's defence of his argument should have its most systematic basis in a position that he later rejects.

Plantinga's third objection may also be bolstered up. Let us allow that Prior is under no obligation to express the existentialist insight by means of the formula $-\diamond(\mathrm{E}\xi - \mathrm{E}a \wedge -\mathrm{E}a)$. Still, the problem arises as to how it is to be

expressed. Existentialism seems sayable; so how is Prior to say it? This is not quite Plantinga's third objection, but it arises naturally from it.

Normally, Prior expresses existentialism in terms of possible worlds: a proposition exists (is statable) in a world only if the individuals it concerns exist in that world. But such a formulation seems to presuppose an outer concept of truth and is, in any case, suspect for a modal actualist. A direct formulation of the existentialist thesis is not available in quantified Q; since it is a peculiarity of Q that whereas what is classically expressible by $-\Diamond-Ey$ is expressible within that system, what is classically expressible by $-\Diamond(Ex \land -Ey)$ is not. Nor does the introduction of propositional quantifiers seem to help, except under rather special assumptions.[18] It seems that Prior must introduce a new connective. The most natural choice, in the present context, is a new form of implication $\rightarrow$, subject to the condition that A $\rightarrow$ B is true in a world iff it is stable in that world and B is true in any world in which A is. Modal dependence might then be expressed by $Eb \rightarrow Ea$ and the existentialist insight by the result of substituting $\xi-Ea$ for b.

As far as I can see, such a connective is perfectly legitimate from the Priorian standpoint. It would have been preferable for Prior to have done without it and to have made do with the weak and strong notions of possibility; but it should come as no surprise to us that he cannot. The Priorian is in the same position as the intuitionist, needing several concepts where the classical logician only needs one. But there is nothing sacrosanct about the weak and strong modalities in this respect. If he needs these concepts, then he may well need others.

There is, however, a deeper problem over expressibility. What the classical logician wishes to express by $\Diamond\exists x \ \Diamond\exists y - \Diamond(Ex \land Ey)$ (there are two possibles that cannot coexist) is surely sayable. But it seems that Prior, even with the introduction of new notions, cannot legitimately say it; there would appear to be some sort of block in principle.[19] If this is so, it would point to some fundamental flaw in the Priorian philosophy, though without revealing what it is.

Something of value may be extracted from Plantinga's objections to Prior; but they fail to get to the heart of the matter. In order to evaluate Prior's position, we require a better understanding of the basis for his views, of what lies behind the doctrines of the vacuity of existence, the unintelligibility of classical possibility, and the incoherence of the transcendental standpoint.

---

[18] I omit the proof. A much more thorough investigation needs to be made of what connectives are definable in Q and of what connectives need to be added to get a reasonably expressive language.

[19] The same goes for the possibilist's $\Sigma x \Sigma y \ \Pi w - (E^*xw \land E^*yw)$. The Priorian is thus incapable of translating all of what the possibilist wants to say into his own modal discourse. What I referred to as a 'technical and recondite' difficulty in [1977], 152, now strikes me as rather deep.

It is only at this deeper, more theoretical, level that the cogency of his views can properly be determined.

This completes our account of the responses to Plantinga's argument for the necessary existence of propositions. We have found no reason why a contingency theorist should give any weight to the argument. If he adopts a classical position, he will dismiss it as a fallacy of equivocation; and if he adopts a Priorian position, he will dismiss it as unintelligible.

It is surprising that Plantinga should have thought that the argument would carry much weight in the first place; for what is lacking is any standpoint, concerning the nature of propositions or some other relevant matter, that would make intelligible why certain statements may be adopted as premises rather than others. The choice of premises is natural once it is assumed that propositions exist necessarily; but without that assumption, the choice seems arbitrary.

This is brought out upon considering which of the following statements should be adopted:

(i) Unrestricted Abstraction for Propositions: $\Box(T\xi S \equiv S)$,
(ii) Unrestricted Abstraction for Properties: $\Box\forall x\ \Box(xH\lambda\ xA(x) \equiv A(x))$,
(iii) Predicate Actualism for Truth: $\Box\forall\rho\ \Box(T\rho \supset E\rho)$,
(iv) Predicate Actualism for the Copula: $\Box\forall x\Box\forall\varphi\ \Box(xH\varphi \supset Ex \wedge E\varphi)$,
(v) The Necessary Existence of Propositions: $\Box\forall\rho\ \Box E\rho$,
(vi) The Necessary Existence of Individuals: $\Box\forall x\ \Box Ex$.

Plantinga adopts (i), Unrestricted Propositional Abstraction, and (iii), Predicate Actualism for Truth. He then draws conclusions that make (v), the Necessary Existence of Propositions, plausible. But why should he not equally well adopt (ii), Unrestricted Property Abstraction and (iv), Predicate Actualism for the Copula? He could then model the following reductio on his original argument:

(1′) $\Diamond -Ea$
(2′) $\Box(-Ea \supset aH\eta)(\eta = \lambda\ x-Ex)$
(3′) $\Box(aH\eta \supset Ea)$
(4′) $\therefore \Diamond(-Ea \wedge Ea)$,

thereby making (v), the Necessary Existence of Individuals, plausible.

One suspects that the only reason Plantinga is prepared to accept the one pair of statements and not the other is that he can get away with it. The necessary existence of propositions is a moot point, and so there is room for the acceptance of (2) and (3) (or of (i) and (iii)). On the other hand, we all know that individuals do not exist necessarily; and so there is no room for the acceptance of (2)′ and (3)′ (or of (ii) and (iv)). But the intrinsic plausibility

of either pair of statements is the same, whatever their consequences; and so as an argument for the necessary existence of propositions, (1)–(4) should have no more cogency than the admittedly absurd argument (1)′–(4)′ for the necessary existence of individuals.

The original argument (1)–(4) may acquire a spurious cogency from the conflation of two opposing positions on modality. It is as if Plantinga starts off the argument as a classical modal logician, happy to accept the first premiss under a classical reading of possibility. He then continues the argument as a Priorian, accepting the second and third premisses because of the lack of any distinction between inner and outer truth. But such a conflation is untenable; as I have already made clear, the denial of the inner/ outer distinction brings with it, for the contingency theorist, the rejection of the classical concept of possibility.

One may detect in Plantinga's work, I think, an excessive reliance on the advantages to be gained from deductive argumentation. Occasionally the arguments that he propounds are invalid, as in the derivation of serious actualism from actualism *simpliciter*; and then there is no question of their having much value. But even when the arguments are valid, they sometimes do little to enhance the plausibility of their conclusions. The present argument for the necessary existence of propositions is a case in point; and so, I would argue, is his version of the ontological argument ([1974], ch. 10), though this is not the place to discuss it.

Plantinga's overreliance on deductive argumentation is perhaps symptomatic of a more general problem of such overreliance in contemporary analytic philosophy. The value of deductive argument in philosophical work is severely limited. I know of no long chains of reasoning, such as one finds in mathematics, that are both successful and helpful. Sometimes short arguments can lead to surprising conclusions; as in the Megarian derivation or the 'paradoxes' of strict implication. But such cases are the exception. Usually, deductive argumentation serves merely to articulate a previously held position. If philosophy is compared to stumbling in the dark, then deductive argumentation is like the groping one does with one's hands; it helps to fix where one is, but not to move forward.

6.   An Alternative Reduction

Let us return to the question of reduction. As we have seen, Plantinga's proposed reduction fails. The objects it uses, the world propositions and individual essences, already presuppose the objects to be eliminated, the possible individuals. It also has the further drawback, particularly severe for the undiscriminating actualist, that those propositions and essences will themselves be among the merely possible entities. How then is the challenge set out at the beginning of the chapter to be met?

I shall here outline an alternative solution to the problem, one that is not open to the previous difficulties and that, as I shall argue, is superior in other respects to Plantinga's proposal.

The central idea behind the new reduction[20] is that the possibilist claim 'some possible individual A's' ($\Sigma x A(x)$) be replaced by the modal claim 'possibly some (actual) object A's' ($\Diamond \exists x A(x)$). This idea works well in case $A(x)$ is a rigid statement, with unvarying conditions of application from world to world; but it breaks down in the other cases. To say some possible individual does not exist ($\Sigma x - Ex$) is not to say that possibly some (actual) individual does not exist ($\Diamond \exists x - Ex$).

This difficulty may be overcome by bringing in reference to the actual world. Instead of saying 'possibly some individual A's', one says 'the actual world is such that possibly some (actual) individual A's in that world'. This statement may, in its turn, be reduced in two further respects. First, the phrase 'the actual world' may be eliminated in accordance with Russell's theory of descriptions or, since there is exactly one actual world in each world, it may be eliminated in favour of either an existential or universal *actualist* quantifier for worlds, i.e. by one that, in each world, ranges over just that world. Second, the notion 'x A's in w' may be rendered as 'necessarily, if w exists then x A's'. Combining both of these changes then gives us the following rendering for 'some possible individual A's':

> There is an (actual) world such that possibly there is an (actual) individual for which necessarily if the world exists (is actual) then the individual A's ($\exists w \Diamond \exists x \Box (Ew \supset A(x))$).

A similar account can be given of possible worlds, but with reference to individuals replaced throughout by reference to worlds. Thus 'some possible world A's' ($\Sigma v A(v)$) becomes:

> There is an (actual) world v such that possibly there is an (actual) world w for which necessarily if w exists then v A's ($\exists w \Diamond \exists v \Box (Ew \supset A(v))$).

The ingredients for a full reduction of possibilist discourse are now at hand: simply, replace world-relative predications (such as x P's at w) by necessity statements ($\Box (Ew \supset Px)$), replace the actuality predicate for worlds with the existence predicate, leave identity alone, and eliminate the possibilist quantifiers for individuals and worlds in the manner prescribed.

There are two minor respects in which this reduction differs from Plantinga's. First, it retains quantification over worlds, though of the actualist, not possibilist, sort. In one version of the Plantinga reduction, worlds are traded in for propositions. However, our residual reference to worlds should

prove acceptable to the actualist. His objection is not to worlds as such, but to possible worlds; just as his objection is not to persons as such, but to possible persons. It is in his capacity as world-reducer, not actualist, that he will eliminate reference to the actual world; just as it is in his capacity as person-reducer that he might eliminate the reference to actual persons.

Secondly, we have not followed Plantinga in distinguishing between actual and existent worlds (see Loux [1979], 258). Possible worlds, it seems to me, stand in the same relation to the actual world as possible people stand to actual people: they are what might have been. But this difference of opinion need not hold us up. Should one prefer Plantinga's view, one can substitute the actuality-predicate on worlds for the existence-predicate.

It is clear that our reduction is not open to the objections that were levelled against Plantinga's; there is no question of its presupposing possible individuals or of its using any kind of possible entity. But even with these objections aside, our reduction has a marked advantage in two other respects: the economy of its ontology and of its assumptions. This is of dialectical significance, since it makes the reduction less vulnerable to criticism. But it is also of theoretical significance, since it shows that the additional elements are not necessary to a reduction as such.

The difference between the present and the previous advantage can be brought out as follows. Suppose we take Plantinga's reduction and, treating them as possibilist, eliminate the quantifiers over propositions and properties in accordance with our own reduction: essentially, when Plantinga says 'for some world-propositions (essence)' we shall say 'possibly for some world-propositions (essence)'. Then our previous objections fail. The difficulties that remain are what constitute the additional advantage of our reduction.

Take ontology first. Our own reduction is extremely modest in its ontological underpinnings. Apart from the original ontology of the modal language, it includes merely an actualist ontology of worlds (or their surrogates). It extracts from the possibilist language, with its possibilist ontologies of worlds and individuals, exactly what is actualistically acceptable. The Plantinga reduction, on the other hand, posits as much by way of ontology; for it must quantify directly over possible worlds (or their surrogates) and it must employ actualist quantifiers over individuals to explain both world-relative predications and the notion of individual essence. But his reduction must posit much more; for it must quantify over properties and also over *all* the possible worlds (or their surrogates). Our own reduction, it needs to be emphasized, requires no such commitment to abstract or intensional entities.

Now take the question of assumptions (in so far as this is separable from ontology). Once the grounds for the correctness of our reduction are examined, they are seen to consist in these two assumptions:

Existence Necessarily an (actual) world exists ($\Box \exists w Ew$);

Completeness Necessarily, if a statement *can* hold in a world then it *must* hold in that world ($\Box \forall x \Box \forall w (\Diamond (Ew \wedge A(x)) \supset \Box (Ew \supset A(x)))$) and similarly for when $A(x)$ contains more variables).

The Plantinga reduction must make corresponding assumptions about worlds (or their surrogates). But it must also assume that necessarily each individual has an individual essence (this is on top of the necessary existence of properties and propositions). This assumption has no counterpart in our own reduction. But it is also controversial, even granted an ontology of properties. If the essences are taken to include identity properties, then it must be allowed that there are genuinely *de re* properties. On the other hand, if the essences are taken to be purely qualitative, then their existence requires the Discernibility Doctrine and is only plausible under a platonic stance in which the class of relations is closed under such strong operations as infinitary conjunction and infinitary quantification.

We see that, in addition to the more basic defect of circularity, the Plantinga-style reduction is needlessly uneconomical in its ontology and its assumptions. Perhaps the only remaining appeal of the approach rests on some sort of commitment to proxy reduction, reduction that proceeds via the introduction of proxies or surrogates. It might be thought that all reductions should proceed in this way, with statements about the disputed entities giving way to statements about their proxies.

Plantinga's reduction fits into this mould, and, as I suggested, can perhaps best be motivated as an attempt to find suitable proxies for possible worlds and possible individuals. My own reduction, on the other hand, cannot be seen in this way. Talk of possibles is a way of talking about actuals, not just in the sense that is common to all reductions, but also in the sense that talk of possibles becomes a distinctively modal manner of talking about actuals. If it is asked, 'With what entities do you identify the possible worlds and individuals?', no sensible answer can be given. The reduction is one that trades in, not object for object, but object for mode.

There is, however, no good reason why all reductions should proceed via proxies. Some do; some do not. We are perhaps overimpressed by the examples from the logicist reduction of mathematics. We then attempt to assimilate all reductions to this model, while ignoring those reductions, such as the elimination of quantification over pairs in terms of pairs of quantifiers over individuals, that do not conform. But there is nothing in the nature of reduction to require that it proceed via proxies. Once this is appreciated, any remaining appeal in Plantinga's approach would seem to disappear.

6

The Problem of Possibilia

1. Introduction

Are there, in addition to the various actual objects that make up the world, various possible objects? Are there merely possible people, for example, or merely possible electrons, or even merely possible kinds?

We certainly talk as if there were such things. Given a particular sperm and egg, I may wonder whether that particular child which would result from their union would have blue eyes. But if the sperm and egg are never in fact brought together, then there is no actual object that my thought is about. (Cf. Gupta [1980], 20 n. 15.) Or again, in the semantics for modal logic we presuppose an ontology of possibilia twice over.[1] For first, we countenance various possible worlds, in addition to the actual world; and second, each of these worlds is taken to be endowed with its own domain of objects. These will be the actual objects of the world in question, but they need not be actual *simpliciter*, i.e. actual objects of *our* world.

What are we to make of such discourse? There are four options: (1) the discourse is taken to be unintelligible; (2) it is taken to be intelligible but non-factual, i.e. as not in the business of stating facts; (3) it is taken to be factual but reducible to discourse involving no reference to possibilia; (4) it is taken to be both factual and irreducible.[2] These options range from a full-blooded form of actualism at one extreme to a full-blooded form of possibilism at the other. The two intermediate positions are possibilist in that they accept the intelligibility of possibilist discourse but actualist in that they attempt to dispense with its prima facie commitment to possibilia. All four positions have found advocates in the literature. Quine, in his less irenic moments, favours option (1); Forbes ([1985], 94) advocates option (2), at least for certain parts of possibilist discourse; many philosophers, including Adams (1974) and myself, opt for (3); while Lewis [1986] and Stalnaker [1976] have endorsed versions of (4) that differ in how full-blooded they take the possible objects to be.

---

[1] See Kripke [1963] for a standard exposition of the semantics.

[2] See Fine [2001] for a general discussion of what these various options amount to.

My focus in this chapter is on the third option. I wish to see to what extent reference to possibilia might be understood in other terms. Can we regard talk of possibilia as a mere *façon de parler*, perhaps somewhat in the same manner as talk of the average man or of infinitesimals?[3] I shall not be concerned to argue directly against any of the other options. However, any argument for the viability of (3) is indirectly an argument against their plausibility. For (4), especially in its more extreme forms, offends against what Russell has called our 'robust sense of reality', (1) offends against our even more robust sense of what is intelligible, while (2) offends against our somewhat less robust sense of what is factual. It is therefore preferable to go with the third option, if we possibly can.

2.   Problems with Proxy Reduction

The most obvious way to make sense of possibilist discourse is in terms of surrogates or proxies. With each possible x is associated another entity x', acceptable to the actualist, and any statement $\phi(a, b, \dots)$ about the possibles $a, b, \dots$) is then understood in terms of a corresponding statement $\phi'(a', b', \dots)$ about the associated entities $a', b', \dots$ As a model for such a reduction, we may take the logicist-style reduction of numbers to sets: each number is associated with a 'representative' set, and a statement about numbers is then understood in terms of a corresponding statement about the associated sets.[4]

But what is the relationship between a possible object and its surrogate? For which entities are the possibilia traded in? The simplest view on the matter is that the relationship is one of identity; each entity is traded in for itself. But such a 'reduction', if it may be called that, is *always* available to us. And so how can it serve to alleviate ontological qualms in any particular case? The answer is that the significance of such a reduction must lie in the way the entities are described. We have a domain of entities that is characterized in problematic terms. It is then shown how each entity from this domain is identical to an entity from a domain that is characterized in relatively unproblematic terms; and doubts about the entities, *qua* members of the problematic domain, are thereby laid to rest. A physicalist's doubts about the ontological status of mental events, for example, might be put to rest in this way if he comes to believe that every mental event is in fact a physical event.

---

[3] As should be clear from Fine [2001] the viability of any reduction will also depend upon its success in accounting for our understanding of modal discourse and our knowledge of modal truth. See Peacocke [2002] for a broader discussion along these lines.

[4] For more on the general approach, see Quine [1964, 1969].

Is a similar kind of view available to the actualist? Can he maintain that possibilia are really just *Y*s, for some actualistically acceptable description *Y* (i.e. for some description that makes no reference to merely possible objects)? After all, the possible winners of a race consist of the actual losers. So could not something similar be true in the case of possibilia? Could not every possible *X* be identical to an actual *Y*, for some actualistically acceptable description *Y*?

It seems to me that no view of this sort can be correct. Suppose, to fix our ideas, that it is maintained that every (merely) possible person is identical to an actual property—one perhaps that specifies its 'essence'. Consider now a possible person. Then it is possibly a person. But no property is possibly a person and so no possible person is identical to a property: for there is a possibility for the one, namely that of being a person, which is not a possibility for the other.

A similar difficulty besets many other identifications of this sort that have been proposed. Possible states of affairs, for example, have often been taken to be propositions. But this cannot be correct, since any possible state of affairs is possibly a state of affairs but no proposition is possibly a state of affairs. Or again, Stalnaker ([1976], 230) and Plantinga ([1974], 44) have suggested that we might think of a possible world as a way the world might have been. But a possible world is possibly the world, just as a possible person is possibly a person, yet no way the world might have been is possibly the world, just as no way I might have been is possibly me. Thus it is not just that the actual world is not a way things might be, as emphasized by Stalnaker ([1976], 228) and van Inwagen ([1980], 407); no possible world is such a way either.

Whatever the merits of reduction via identity in other contexts, it is of no avail here. If there is to be a proxy reduction, it had better be achieved by means of proxies that are distinct from the possibilia themselves.

But again, an obvious solution suggests itself. For why not 'identify' each possible world with a proposition that is true in that world alone (or, if we wish to pick out a particular proposition, with the conjunction of all propositions that are true in the world)? And why not identify each possible object with a property that is necessarily borne by that object alone (or with the conjunction of all properties that are necessarily borne by the individual)? Each possible, be it world or object, is in effect identified with a description by which it might be specified.[5]

---

[5] A view of this sort was originally proposed by Prior (Prior and Fine [1977], ch. 2), though only for the case of worlds. Essentially the same account was later given by Adams ([1974], 204). The extension to possible individuals was proposed by Fine in Prior and Fine [1977] and possibly by Plantinga [1976] (though not if his disclaimers in Plantinga [1985], 330–2 are to be heeded).

The main difficulty with this proposal is that there can be no assurance, from an actualist point of view, that distinct possible objects or worlds can be identified with distinct surrogates. Let us provide a simple illustration of the difficulty. Suppose there is some radioactive material in the actual world w_0 that just happens not to emit any particles from a certain time on but that *might* have emitted two particles of the same type at that time. These two particles, call them α and β, are presumably merely possible; they are not identical to any actual particles. And it is plausible to suppose that there is no actualistically acceptable means by which they might be distinguished. Of course, there is a possible world w_1 in which α is distinguished by one trajectory and β another. But if there is such a world, then there is presumably another world w_2 just like it in which the trajectories are interchanged. For what is so special about α as opposed to β that it is destined to have the one trajectory rather than the other?[6] Thus we will be as unable to distinguish between the worlds as we are to distinguish between the particles themselves.

If we pretend that w_1 and w_2 and the actual world w_0 are the only worlds that there are, then we might depict the scenario as in Fig. 6.1.
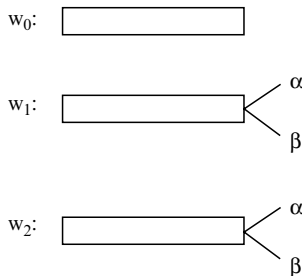
Here, in this miniature 'pluriverse', the worlds w_1 and w_2 are actualistically indiscernible, as are the particles α and β. Given that there is no actualistically acceptable means by which the particles or worlds might be distinguished, they will be associated with the very same surrogates, since any actualistically acceptable means of associating them with distinct surrogates would provide us with an actualistically acceptable means of distinguishing between the particles or worlds themselves; and given that this is so, the reduction must fail, since it will not even be capable of representing the fact that the particles or worlds are distinct.

Another kind of problem case arises from the possibility of there being indiscernible individuals *within* a world. Imagine a universe of eternal recurrence (with respect to both past and future) in which a new messiah

---

[6] We might even suppose that there were convincing scientific reasons for allowing both possibilities in determining the probability of emission.

appears in every epoch. There are then infinitely many possible messiahs; and presumably each of them will be actualistically indiscernible from the others.

A third kind of case arises from the possibility of there being indiscernible natural properties or kinds. (See Bricker [1987], 349–53; Lewis [1986], 158–65; McMichael [1983], for examples of this sort.) There are two sub-cases here, just as in the case of individuals, depending upon whether the indiscernibilities are *intra*world or *inter*world. Pure cases of interworld indiscernibility might always be disputed on the grounds that the identity of a kind, in these cases, is to be tied to role (as on the views of Swoyer [1982]; Shoemaker [1980, 1998]). Thus given that the kinds are indiscernible in their respective worlds, their roles will be the same and hence the kinds themselves must be the same. However, intraworld cases are not so readily disposed of. Suppose, for example, that there are two fundamental kinds of matter in the universe, positive and negative, governed by such laws as: like matter attracts; unlike matter repels. The two kinds of matter would then have completely symmetric roles and so as long as they are 'alien' kinds, not of this world, there would again appear to be no actualistically acceptable way in which they might be distinguished.

There are two main responses to these arguments. One is to dispute the possibilities upon which they are based. It has sometimes been denied, for example, that there can be worlds that are qualitatively, or actualistically, alike and yet differ merely in the identity of the individuals that they contain (e.g. by Lewis [1986], §4.4; Adams [1981]); and, under such views, there would only be one possibility for α and β depicted by w_1 and w_2 in the picture above, not two. But there is something unsatisfactory about making the reduction dependent upon such views—both because they are controversial and because we wish to explain what sense might be given to possibilist discourse by someone who did not accept them. It would be preferable, if at all feasible, to provide a reduction that was free from any substantive assumption about what was or was not possible.

The second response to the cases is to accept the putative possibilities and yet deny that they involve genuine actualist indiscernibilities. Despite our claims to the contrary, it will be maintained that the particles or the messiahs or the kinds of matter can be actualistically distinguished after all. For let x be any given possible object. Then associated with this object will be a certain identity property, the property of being identical to x. But in contrast to the object x itself, this property—like all properties—will exist necessarily. It will therefore be an actual object; and so we may use it, in an actualistically acceptable way, to distinguish x from all other objects (see Plantinga [1976]). (Of course, when x itself is a property or the like, we may proceed directly, by this line of reasoning, to the conclusion that it necessarily exists.)

One way of dealing with this response is to deny the claims of necessary existence upon which it depends. The property of being identical to Socrates, it might be countered, can only exist when Socrates exists; and the kind *positive matter* can only exist in a world in which there is positive matter.[7] But there is, I believe, a more fundamental objection to be made. Let us suppose that an actualist comes to the view that (necessarily) properties necessarily exist. Should the properties that he previously took to be problematic because they were merely possible now be regarded as unproblematic? I think not. Rather, they should still be taken to be problematic, though for reasons that no longer turn on their being merely possible.

For a more fundamental way to understand the actualist's position is that he objects to the idea that general possibilities might be the source of a distinctive ontology of objects that instantiate those possibilities. Consider the possibility that there is a talking donkey ($\Diamond \exists x Px$). The possibilist will claim that it follows from this possibility that there really is an object, possible if not actual, that instantiates it; there is an object, that is to say, that is possibly a talking donkey ($\exists x \Diamond Px$). The actualist will deny that there need be any such object (except as a mere *façon de parler*) and, in general, he will be suspicious of any object whose existence would appear to depend upon its being the instantiator in this way of a general possibility.

But the identity properties of merely possible objects and the alien kinds are just of this sort. It is only because of the possibility of there being an identity property for such-and-such a possible object and it is only because of the possibility of there being a kind which plays such-and-such a role that we are led to believe that there *are* such properties or kinds. Without the belief in the general possibilities, we would have no reason to believe that there were such things. On this understanding of what lies behind the actualist's position, then, he will remain suspicious of these properties and kinds on account of their possibilist origins, even though he accepts that they exist. He will think of them, like other problematic existents, of standing in need of analysis in terms of existents of another sort.[8]

3.   The Possibility of Proxy Reduction

As a result of these difficulties, many philosophers have given up on the idea of proxy reduction; and, indeed, the difficulties in the particular reduction proposed above might appear to extend to any reduction whatever. For consider again our miniature pluriverse with its three worlds w_0, w_1, w_2

---

[7] Fine in Prior and Fine ([1977], §4) and McMichael ([1983], 60–1) develop objections along these lines.

[8] A related objection is made in Fine ([1985] §2) and an altogether different objection to the necessary existence of alien properties is developed by Lewis ([1986], 160–1).

and its two particles α and β; and suppose that a represents, or goes proxy for, α. Then, as we have seen, it must also represent β. For a must be an actual object (or, at least, actualistically acceptable); and so, if it failed to represent β, we could distinguish between α and β in an actualistically acceptable manner, since α would have the property of being represented by a while β would not. This therefore suggests that it will in general be impossible to obtain a unique proxy for each possible individual and that any acceptable form of proxy reduction must therefore fail.

Uniqueness of proxies is not, however, necessary for a proxy reduction to succeed.[9] We may reduce three-dimensional Euclidean geometry to real analysis by identifying each point with a triple of real numbers. But the identification is far from unique. Indeed, any given point might be associated with any given triple. But the ambiguity will not matter as long as it does not result in any ambiguity in truth-value of the sentences to be reduced. This therefore suggests that we may let a represent α and b represent β under one scheme of representation as long as we are also prepared to allow that a represents β and b represents α under another. The previous difficulty then disappears since, given the symmetric nature of the representations (which cannot themselves be actualistically distinguished), we will be left with no way to distinguish between α and β.[10]

A problem remains, however. For a similar story should be told about w_1 and w_2. There will be two proxies, say w and v, that indifferently represent w_1 and w_2 or w_2 and w_1. Suppose now that we pick on a particular scheme of representation, say that in which a represents α, b represents β, w represents w_1, and v represents w_2. Then how are we to determine which paths for a and b are to be assigned in w? Whatever we say, the paths assigned in v must be the reverse. But there seems to be no basis for taking the paths to go one way rather than the other. Thus even when we pick on a particular scheme of representation, there appear to be irresolvable indeterminacies in how it is to be applied.

In order to solve this further difficulty, we must somehow 'co-ordinate' the representation of individuals and worlds. Let me indicate one way in which this might be done. (The less technically minded reader may skip the details.)[11] Let us suppose that we use the distinct actual entities $w_1, w_2, \ldots$ as proxies for the possible worlds and the distinct actual entities $i_1, i_2, \ldots$ as proxies for the possible individuals, both actual and merely possible. We assume that each actual individual i_k goes proxy for itself. Thus it is the actual entities that are not individuals that will go proxy for what we

---

[9] Contrary to what the criticisms in Lewis ([1986], 158, 163–4) might appear to suggest.

[10] Curiously, similar difficulties arise in understanding Cantor's account of cardinal numbers as sets of units (Fine [1998a]).

[11] The basic idea behind the method is presented in Fine in Prior and Fine ([1977], 148), and a related approach has been developed by Sider ([2002], §5).

take to be the merely possible individuals. Co-ordination may now be achieved by means of a *proxy pluriverse*. This consists of three items: the class W of world-proxies; the class I of individual-proxies; and a class of proxy relationships. Each proxy relationship is of the form $\langle w, R, i_1, i_2, \ldots, i_n \rangle$, where w is a proxy world, R is an (actual) *n*-adic relation, and $i_1, i_2, \ldots, i_n$ are proxy individuals. Intuitively, a proxy relationship indicates that the relation R holds of the possible individuals represented by $i_1, i_2, \ldots, i_n$ in the possible world represented by w. Thus a proxy pluriverse represents how the pluriverse might be; it provides an explicit tabulation or model, via the proxies, of the relationships that hold of the possible individuals in each of the worlds.[12]

A proxy pluriverse will not in general be 'realistic'; it will not represent the way the pluriverse really is. How then are such proxy pluriverses to be singled out? In order to answer this question, let us suppose that we are given a list (or well-ordering) $i_1, i_2, \ldots$ of all the proxy individuals. We may then define in a natural way what it is for the proxy world w of the proxy pluriverse to be *realized by* a corresponding list of individuals $x_1, x_2, \ldots$ For this requires that R hold of $x_{k_1}, x_{k_2}, \ldots, x_{k_n}$ just in case $\langle w, R, i_{k_1}, i_{k_2}, \ldots, i_{k_n} \rangle$ is a proxy relationship of the proxy pluriverse. Thus a proxy world will be realized by an assignment of individuals to proxy individuals if it correctly represents the relations that hold among those individuals. A proxy pluriverse may now be said to be *realistic* (given a list of its proxy individuals) if possibly there is an x_1, possibly there is an $x_2, \ldots$ such that:

(1) each $x_k = l_k$, when l_k is an actual individual (taken to go proxy for itself);

(2) x_j and x_k are distinct for $j \neq k$;

(3) necessarily any individual is identical to x_1 or x_2 or...;

(4) each proxy world is possibly realized by $x_1, x_2, \ldots$;

(5) it is necessarily the case that some proxy world is realized by $x_1, x_2, \ldots$

Clauses (1)–(3) say that $x_1, x_2, \ldots$ are pairwise distinct and together constitute the domain of possibilia; clause (4) says that each of the proxy worlds represents a genuine possibility (under the given assignment of individuals to proxy individuals); and clause (5) says that the proxy worlds exhaust the genuine possibilities.[13]

Given a realistic proxy pluriverse, we may then quantify over the proxy worlds and the proxy individuals as if they were the possible worlds and the

---

[12] We shall suppose that distinct proxy worlds enter into different relationships—so that if $w \neq v$ then there is a relation R and proxy individuals $i_1, i_2, \ldots i_n$ such that $\langle w, R, i_1, i_2, \ldots, i_n \rangle$ is a proxy relationship within the proxy pluriverse while $\langle v, R, i_1, i_2, \ldots, i_n \rangle$ is not, or vice versa.

[13] A similar modal description of the pluriverse is given in Prior and Fine ([1977], 147).

possible individuals of the real pluriverse. Thus instead of saying that R holds of a pair of possible individuals in a given possible world, we may say that $\langle w, R, i, j \rangle$ is a proxy relationship within the given proxy pluriverse. There will of course be many realistic pluriverses (and many ways of ordering their proxy individuals). But the ambiguity will not matter, since different realistic pluriverses are isomorphic and hence will yield the same truth-value for any given possibilist claim.

The resulting reduction is highly inelegant. It requires enormous expressive resources in order to capture a relatively modest extension in expressive power. For whether a given proxy pluriverse is realistic depends upon the truth of the infinitary proposition given by the clauses (1)–(5) above. And so, in stating any given reduction, we must either possess the means to express this infinitary proposition, in which case the language of the reduction must itself be infinitary, or we must possess the means to refer to this proposition (or to a corresponding sentence), in which case the language of the reduction must be capable of describing the structure and semantics of an infinitary language or ontology of propositions.

But there is a more serious problem. For how can we be sure that there *is* a realistic proxy pluriverse? The problem is essentially one of cardinality. For in order for a proxy pluriverse to be realistic there must possibly be an x_1, possibly be an $x_2, \ldots$ such that $x_1, x_2, \ldots$ are all the possible individuals that there are. There must therefore be as many variables 'x_1', 'x_2', $\ldots$ — or operators 'possibly an x_1', 'possibly an x_2', $\ldots$ — as there are possible objects. But suppose there are $\mathbf{c}$ such operators, for some cardinal number $\mathbf{c}$. It is then arguable that there could be a greater, infinite number $\mathbf{d}$ of possibilia. For there could be a possible world that contained $\mathbf{d}$ 'parallel' universes, each with its own particles; and since there are presumably only finitely many actual particles (and since, necessarily, each particle is necessarily a particle), at least $\mathbf{d}$ of these particles from the parallel universes will be non-actual.

There are perhaps ways in which this latter problem might be solved.[14] But a general form of the cardinality worry remains. For if a proxy reduction is to succeed, there must be a one–one correspondence between the possible individuals and worlds of the pluriverse, on the one side, and the objects of the actual world on the other (or perhaps we should say, more cautiously, between the possible individuals and worlds of the pluriverse and the objects of some *possible* world, since one might carry out the reduction from the

---

[14] One solution, suggested in Prior and Fine ([1977], 148), is to use so-called 'quasi-classes' to set up a one–one correspondence between the possibilia and the actualia (a great gain in elegance and simplicity is thereby also achieved). Quasi-classes are the possibilist counterpart of plural quantification (in the sense of Boolos [1984]) and were introduced, along with the general idea of plural quantification, in Prior and Fine ([1977], 146–7).

perspective of some possible world, viewed as actual, rather than from the perspective of the actual world itself).

But is such an assumption reasonable? Will there be a world within the pluriverse of the same 'size' as the pluriverse itself? This is a difficult question (and of some interest in itself). But I am inclined to think the answer is no. For there is a puzzle whose solution appears to require that we give up the assumption.[15] I shall state the puzzle for the case of 'communicating egos', though there are other forms it might take.

We imagine ourselves attempting to ascertain how many possible Cartesian egos there are. Now even if there are no actual Cartesian egos, there could be one. That is:

(1) There is at least one possible ego.

It is also plausible that:

(2) Given any possible world containing one or more egos, there is a possible world in which those egos exist and in which, for any subclass of those egos, there is an ego which is in telepathic communication with just those of the given egos that are members of the subclass.

Finally, we may wish to maintain that:

(3) Given any class of possible egos, there is some possible world in which they all exist.

Although each of these assumptions is individually plausible, together they are inconsistent. For from (3) (letting the class be the class of all possible egos), it follows that:

(4) There is a possible world (call it Descartes's world) in which all possible egos exist.

From (1), it follows that

(5) Descartes's world contains some egos.

And from (2), it follows that:

(6) Given any possible world which contains some egos, there is a possible world which contains more egos,

since in the world with telepathic communication there will be more communicating egos than egos with which they communicate. But (4) and (6) are

---

incompatible with one another, since there can be no possible world which contains more egos than the class of them all.

What are we to say? Which of the assumptions (1)–(3) should be given up? It is natural to suppose that it should be (3). But we would like this principle for the most part to be true. And if we ask what is it about the class of all possible egos that prevents them from all existing, the only acceptable answer would appear to be that the class is too large. In other words, the domains of each possible world will be subject to a 'limitation of size'; and even though the pluriverse may be capable of exceeding this size, the worlds within the pluriverse will not be. Each such world will possess an 'actual' or 'actualizable' infinity of objects and be incapable of accommodating the 'potential' infinity of possible objects that belong to the pluriverse as a whole.[16] But if this is our motivation for rejecting the possible existence of all possible egos, then we are obliged to conclude that there are more possible egos than there are objects in any possible world, since it is only this that prevents them all from possibly existing.

If this is right, then the assumption that there could be as many actuals as possibles is untenable and the whole idea of a proxy reduction should be abandoned.[17] But even if it is not right and another solution to the puzzle is discovered, there is still something unsatisfactory, for the reasons already given, about having the adequacy of the reduction depend upon such substantive metaphysical views; and it would be desirable if some other way of reducing possibilist discourse could be found.

4. Reduction without Proxies

It is important to bear in mind that a reduction need not proceed via proxies. The mother of all reductions, Russell's theory of descriptions, cannot readily be regarded as one in which entity gives way to entity, and another example, more pertinent to our present concerns, is that in which quantification over pairs is replaced by quantification pairs. Instead of saying 'there is a pair x such that . . . ', one says 'there is an x_1 and an x_2 such that . . . '. Here there is no single entity that goes proxy for a pair.

Many philosophers seem to have followed Lewis ([1986], 141) in supposing that they must either go with *proxy* reduction ('ersatzism') or accept possible worlds realism. But this is a false dilemma. For as I have indicated in

---

[16] This is a distinction that may be easier for the actualist rather than for the possibilist to maintain. For the actualist may argue that just as there is no perspective (one transcending all ordinals) from which the class of all sets is given, so there is no perspective (one transcending all possible worlds) from which the class of all possibilia is given. (In this connection, see Menzel [1986*a*, *b*]; Grim [1986])

[17] There is a related problem over cardinality in representing Fregean abstracts as sets within the cumulative hierarchy (Fine [1998*b*] or [2002], 14).

previous work,[18] it is possible to provide a straightforward non-proxy reduction of possibilist discourse.

The basic idea is to take modality as primitive and to treat the possibilist quantifier 'there is a possible object x' as equivalent to 'possibly there is an object x'—where the second quantifier (in the scope of the possibility operator) is actualist, ranging in each world over the actual objects of that world. Thus to say that there is a possible object that is possibly a talking donkey is to say that possibly there is an object that is possibly a talking donkey.

Unfortunately, the above method does not work in all cases. To say that there is a possible object that is not actual is not to say that possibly there is an (actual) object that is not actual, since the latter claim is necessarily false while the former claim is presumably true. The method must therefore be modified.

The difficulty is that the possibility operator takes us to another world, whereas we wish to evaluate the statement governed by the possibilist quantifier in the original world. We therefore need some device to take us back to the original world. There are various ways in which this might be done, but let me here present just one. Back-reference is to be achieved, in the most direct and straightforward manner, by means of reference to the actual world. Thus to say that there is a possible object that is not actual will be to say that the actual world is such that it is possible that there is an object whose non-existence is compatible with that world being actual. And, in general, to say that some possible object ϕs is to say that the actual world is such that it is possible that there is an object whose ϕing is compatible with that world being actual.

The reduction of possible worlds is now merely the special case of the reduction of possible individuals in which the individuals are taken to be the worlds. Thus to say 'for some possible world' will be to say 'possibly for some (actual) world' in the simplest case; and back-reference can be achieved in the general case in the same way as before. (Thus worlds will now play a double role, as the objects of quantification and as the means for securing back-reference.)

Of course, we do not get rid of the world on this approach—merely *possible* worlds. But the problem for the actualist is not with the actual world, but with possible entities, whether they be worlds or of some other kind. If we also wish to get rid of the actual world and treat it as a special kind of fact, say, or proposition, then this is something that might be tacked onto the present reduction but is of no concern to the actualist as such.

---

[18] Beginning with Fine in Prior and Fine ([1977], 130–9). A comparison with the standard proxy reduction is made in Fine ([1985] 180–3) and some technical details can be found in Fine [1980, 1981, 1982].

The beauty of the method is that it does not require any addition to the ontology. Quantification over possibilia, be they worlds or individuals, is eliminated in favour of the corresponding quantification over actualia. There is a direct trade between the ontology of possibilia, on the one hand, and the ideology of modality, on the other. Moreover, the assumptions upon which the reduction depends are minimal. It need only be assumed that:

(1) necessarily there is a world; and
(2) necessarily, for any world and true proposition, the truth of the proposition is implied by the existence of the world.[19]

Once these assumptions are granted, the adequacy of the reduction is guaranteed.

The main difficulty with this approach is that it is not clear how it is to be extended to quantification over sets of possibles (Fine in Prior & Fine [1977], 145). We could try to understand such quantification as quantification over possible sets. But a possible set can only consist of compossibles, i.e. of objects that can possibly all exist, whereas we should also allow for quantification over all sets of non-compossible objects.

A uniform solution to this problem is available in the case of any proxy reduction, since a set of the objects from the class of objects to be reduced can always be identified with the set of their proxies; and it would be desirable if a uniform solution could also be obtained in the case of any non-proxy reduction. One possibility here is to treat quantification over sets as a certain form of plural quantification. To say that there is a set X is to say, in effect, that there are certain individuals $x_1, x_2, \ldots$; and to say that $x \in X$ is to say, in effect, that x is one of the individuals $x_1, x_2, \ldots$ Let us be a little more precise. (Again, the less technically minded reader may skip the details.) Suppose that we are somehow equipped with an understanding of a first-order language L_1 in which the quantifiers range over individuals; and let it be granted that our understanding extends, in principle, to sentences of infinitary length (we could equally well work with propositions rather than sentences). Suppose that we now introduce a quantifier $\exists X$ over sets of individuals; and consider any sentence ϕ of the resulting language. We wish to extend the truth-predicate to the resulting language, though without quantifying over sets. This may be done inductively on the logical complexity of the sentence to which the truth-predicate is applied. The clauses in the case of the truth-functional connectives and the quantifier $\exists x$ over individuals are straightforward. And so that leaves sentences of the form $\exists X\phi$. Intuitively, we wish to say that such a sentence is true iff an instance is true, but we have no straightforward way of saying what an instance is. What we

---

[19] If we wish to take care of questions concerning the identity of worlds, then it should also be assumed that there is necessarily at most one world.

may do instead is to find a first-order counterpart of an instance. This can be obtained in two steps. First we replace each free occurrence of the set-variable 'X' in ϕ by a term '$\{x_1, x_2, \ldots\}$' with a given number of distinct new variables 'x_1', 'x_2', ... (sets give way to individuals); and then we replace each atomic subformula '$x\varepsilon\{x_1, x_2, \ldots\}$' in the resulting formula by '$x = x_1 \lor x = x_2 \lor \ldots$' (membership gives way to identity), and similarly for all other atomic subformula involving $\{x_1, x_2, \ldots\}$.[20] Let the resulting sentence be 'ϕ'. Then an instance of $\exists X \phi$ may be taken to be a sentence of the form '$\exists x_1, x_2, \ldots, \phi$'.

We thereby obtain truth-conditions for a language L_2 with variables for both individuals and sets of individuals. The same general method can be extended to a language L_3 with quantifiers that range over sets of 'rank' ≤ 2, i.e. over sets whose members are either individuals or sets of individuals; and the construction may then be continued into the transfinite. We thereby obtain truth-conditions for a language L_α of arbitrary order α; and so, as long as we are able to identify the sets we wish to quantify over as those whose rank is less than a given ordinal α, we are in a position to account for quantification over such sets in terms of our understanding of the base language.

This reduction does not allow us to eliminate reference to sets altogether, since the definition of truth requires the full resources of set theory.[21] But the reduction does show how we may extend our understanding of quantification over sets of arbitrary rank to the ontology of any infinitary first-order language. And since our non-proxy reduction of possibilist discourse extends straightforwardly to the infinitary quantifier 'there are possible objects $x_1, x_2, \ldots,$' we are thereby able to account for higher-order quantification over sets of possible individuals, sets of such sets, and so on throughout the cumulative hierarchy.[22]

5. Fictionalism

We have argued against any proxy reduction of the possible to the actual and in favour of a certain form of non-proxy reduction. But are there any other acceptable forms of non-proxy reduction?

---

[20] Atomic formulae of the form $\{x_1, x_2, \ldots\} \in x$, $x \in y$ and $x = \{x_1, x_2, \ldots\}$ are replaced by $\bot$; and $X = Y$ is treated as definitionally equivalent to $\forall x(x \in X \equiv x \in Y)$. Special provision should be made for the null class.

[21] Indeed, it also requires that we be able to treat the domain of sets in the object language as a set within the metalanguage. But this set-theoretic 'ascent' is something which one might argue is always available to us.

[22] The idea behind this reduction derives from Gödel's reconstruction of Russell's no-class theory in Gödel [1944], 132).

One candidate is the modal fictionalism of Rosen [1990].[23] The possibilist wishes to assert:

(e) possibly there are talking donkeys iff there is a possible world in which donkeys talk.

And, in general, where ϕ is a modal claim and ϕ^* is its possibilist translation, the possibilist will maintain:

(E) ϕ iff ϕ^*.

But, given that he accepts the possibility of talking donkeys and other such modal claims, he is thereby committed to a plethora of possible worlds. The fictionalist, by contrast, will think of the possibilist's views of the pluriverse as constituting a fiction and will therefore replace (e) with:

(e′) possibly there are talking donkeys iff it is true according to the fictional account of the pluriverse that there is some possible world in which there are talking donkeys;

and, more generally, he will replace (E) with:

(E′) ϕ iff it is true in PW that ϕ^*,

where PW is the fictional account of the pluriverse. In this way, he can take advantage of the possible world semantics for modal discourse without committing himself to its ontology. In making the transition from ordinary modal claims to their possibilist translation, we enter a fictional realm of possible worlds and their inhabitants, according to the fictionalist, rather than one that is genuinely there.

The view, as stated, would appear to fall flat on its face. For on any account of the fiction PW that might reasonably be proposed, there will presumably be possibilist translations ϕ^* of modal claims ϕ whose truth-value is not settled within PW. Perhaps ϕ^* is the claim that there is a possible world in which there are more than $\aleph_{17}$ individuals. It is not then implausible to suppose that

(I) it is not true in PW that some possible world contains more than $\aleph_{17}$ individuals and it is not true in PW that every possible world contains at most $\aleph_{17}$ individuals.

But, from the modified equivalence (E′) above and the first part of (I), it follows that it is not possible that there are more than $\aleph_{17}$ individuals and,

---

[23] A related form of fictionalism, to which similar criticisms apply, is that of Armstrong [1989]. An altogether different approach, which I shall not discuss, is that of Forbes ([1985], 89–95). The view is critically examined in Cresswell ([1990], 47–62) and Chihara ([1998], ch. 4).

from (E′) and the second part of (I), it follows that it is not necessary that there are at most $\aleph_{17}$ individuals. And this is a contradiction.

In the face of this difficulty, Rosen ([1990], 341–3) has suggested that modal claims ϕ like the one above should be taken to be indeterminate, i.e. to be neither true nor false. But this is of no help in avoiding the contradiction unless principle (E′) is somehow modified. Presumably, the intent is that it should take the form:

(E″) it is true that ϕ iff it is true in PW that ϕ^*,

where 'it is true that' is an operator that converts an indeterminate statement into one that is false. But the scope of the view is now seriously compromised, for we lack any account of what it is in general for a modal statement ϕ to hold. Where ϕ is indeterminate, we would like there to be a possibilist or quasi-possibilist translation that is correspondingly indeterminate. But the fictionalist is unable to provide any such translation, since ϕ^* and 'In PW, ϕ^*' are both false. Thus the fictionalist is unable adequately to represent the question 'Is it possible that there are more than $\aleph_{17}$ individuals?' He can only provide a question to which the answer is no, whereas we want a question to which the answer is neither yes nor no.

Numerous other difficulties for the view have been raised (see Rosen [1990, 1993, 1995]; Brock [1993]; Noonan [1994]; Divers [1995]; Hale [1995]; Nolan and Hawthorne [1996]; Chihara [1998]; Sider [2002]). Three strike me as especially serious. First, the account depends upon a problematic notion of what it is to be true in a fiction. For can we understand this notion in the required way without already presupposing an understanding of modality? Second, it is not clear how to specify an adequate fiction PW, one that will deliver the right truth-values, without already presupposing the truth of the modal statements whose truth-conditions are in question. Third, the account does not adequately represent the content of modal claims even should it get their truth-value right. To make the controversial claim that things are necessarily spatio-temporally connected is not to claim that it is true in a fiction, in which every possible world is taken to be spatio-temporally connected, that every possible world is spatio-temporally connected, even should the claim be true. (To some extent, these difficulties are interdependent. We might solve the first difficulty, for example, by taking truth-in-a-fiction to be strict logical implication, but the second difficulty then becomes more acute.)

From our own point of view, Rosen's fictionalism involves a large element of overkill. For it attempts to get rid of the ordinary modal idioms in addition to the ontology of possible worlds and individuals. But suppose we are happy with the modal idioms and merely wish to rid ourselves of possibilia. A much more satisfactory form of fictionalism can then be maintained. For we can take the possible worlds semantics itself to

constitute a fiction. Thus among the basic postulates of the fiction will be the following:

(1) A statement is true iff it is true in the actual world;
(2) Possibly A is true in a world iff A is true in some world;
(3) Something ϕs is true in a world w iff some individual of w ϕs in w.

We also import all truths into the fiction as long as their quantifiers are restricted to what is actual.[24]

There are three major differences between our fictionalism and Rosen's. First, instead of telling a metaphysical story about the constitution of the pluriverse, as with Rosen's account, our fiction tells a semantical story about the connection of the pluriverse with the modal facts. Second, truth-in-a-fiction is not a new substantive notion for us; it is simply logical implication (in the strict sense). Third, the connection between modal and possibilist claims is reconceived. Instead of modifying the original equivalence (E) to (E′) (or to E″)), we modify it to:

(E‴) it is true in the fiction that (ϕ iff ϕ^*).

Thus the original equivalence (E) is itself taken to be assertible within the given fiction and reasoning can proceed within the fiction as if we were bona fide possibilists.

It is clear, in the light of these differences, that our account is not subject to the difficulties mentioned above. Since we do not insist upon (E′), the difficulty over indeterminacy does not arise. But should the actualist statement ϕ be true, there is no difficulty in showing that ϕ^* is true in the fiction. For (ϕ iff ϕ^*) will be true in the fiction by the semantical postulates, ϕ will be true in the fiction by importation, and so ϕ^* will be true in the fiction as a logical consequence. Thus (E′) will never fail when ϕ is either true or false; and there will be no unwanted gaps. Since the imported modal truths may be used in this way to deliver the correct possibilist consequences, there is no special difficulty in providing an adequate non-circular account of what the fiction is. Finally, there will be no difficulty over according the correct content to modal claims, since no attempt is made to ascribe a content to them. Our aim is simply to adopt a fictionalist simulacrum of possibilist discourse.[25]

The new form of fictionalism is analogous to if–then-ism in the philosophy of mathematics[26] and is not without its attractions. It is still subject to difficulties, however. For we have substantive views about the nature of

---

[24] This corresponds to Rosen's 'encyclopedia' ([1990], 335). We need the restriction to prevent the importation of something like 'everything is actual'.

[25] I might note that the objections made by Brock [1993] and Hale [1995] are also inapplicable to the present version of fictionalism.

[26] As characterized in Putnam ([1967a], § 3), for example.

possible worlds—we do not think of them as mere ciphers. We are inclined to think, for example, that no two worlds can be exactly alike or that what is true at a world cannot be different from what it is. These views should not, of course, be understood as being literally true of how things are for the fictionalist, since he does not believe in many worlds, but it should be possible for him to understand them as being true of how things are in the fiction. Thus he should take it to be true in the fiction that no two worlds are exactly alike or that what is true in a world cannot be different from what it is. However, under the most natural construal of what the fiction is, these various questions concerning the content of the fiction will not be settled one way or the other. The worlds serve merely as pegs upon which to hang the modal truths and nothing beyond their serving this structural role need be said about their nature. So the view will suffer from a problem of incompleteness after all, not with respect to ordinary modal claims but with respect to the superstructure of worlds within which they are embedded.

How might this incompleteness be repaired? There are two main options. The first is to add postulates to the fiction that explicitly describe the nature of the worlds. Thus there may be a postulate stipulating that no two worlds are exactly alike. But we then face a variant of the third of the objections listed above. For to claim, in the intended sense, that no two worlds are exactly alike is not to claim that this is true in a fiction in which it has been stipulated to hold. The other option is to have these various claims follow from actualist modal truths in much the same way that the existence of worlds with talking donkeys follows from the possibility that donkeys talk. Thus suppose we take it to be true that necessarily for any (actual) world w and necessarily for any distinct world v there is some elementary fact holding in v but not in w (or vice versa). Then the rest of the fiction might be so set up that, once this modal truth is imported into the fiction, the desired possibilist truth concerning the discernibility of distinct worlds will follow. But in this case, the fictionalism does not work, for, given that our actualist modal language already contains quantification over worlds, possibilist quantification over worlds and individuals will be uncontroversially definable in the manner of our own reduction. Thus fictionalism of the supra-modal sort is either inadequate or redundant.

This page intentionally left blank

III
Issues in Metaphysics

This page intentionally left blank

7
The Varieties of Necessity

Necessity abounds. There are the necessary truths of logic, mathematics, and metaphysics, the necessary connections among events in the natural world, the necessary or unconditional principles of ethics, and many other forms of necessary truth or connection. But how much diversity is there to this abundance? Are all necessary truths and connections reducible to a single common form of necessity? And if not, then what are the different ways in which a truth might be necessary or a necessary connection might hold?

It is the aim of this chapter to show that diversity prevails. I shall argue that there are three main forms of necessity—the metaphysical, the natural, and the normative—and that none of them is reducible to the others or to any other form of necessity. Thus, what it is for a necessity or possibility of any of these forms to obtain does not consist in the obtaining of some other form or forms of necessity or possibility.

Although the focus here falls squarely within the philosophy of modality, some of my arguments may be of broader interest. For certain currently fashionable views on scientific essentialism and ethical naturalism entail the collapse of forms of necessity that I would wish to keep distinct. Thus I have found it crucial to indicate what it is in these views that I take to be in error; and this has required consideration of questions from within the metaphysics of natural kinds and the epistemology of ethical belief.

1. Necessities

A proposition is necessary if it *must* be true and possible if it *might* be true. On the face of it, there are different ways in which a proposition might be necessary or possible. Suppose I ask, 'Is it possible to get from London to New York in under an hour?' Then I might answer 'No', meaning that it is impossible given the currently available means of transport; or I might answer 'Yes', meaning that it is scientifically possible. Or again, suppose I ask, 'Is it possible to get from the earth to the sun in under 2 hours?' Then

I might answer 'No', meaning that it is scientifically impossible; or I might answer 'Yes', meaning that it is logically possible.

Given that there are these different ways in which a proposition might be necessary, then how are they related? Is it possible to define, or otherwise explain, some in terms of others? And if it is, then which are the most basic?[1]

I suspect that many philosophers, in response to these questions, might be attracted to some version of modal monism. They would maintain that there was a single underlying modal notion in terms of which all others could be defined or understood. However, philosophers of this persuasion might well be tempted to adopt different views of what that underlying notion was. Many philosophers of the 'old school' would take it to be that of logical necessity in the narrow sense. This is the sense in which it is necessary that anything red is red, though not necessary that nothing red is green or that I am a person. The philosophers of the 'new school', on the other hand, would take the single underlying notion to be that of logical necessity in the broad sense, or what is sometimes called 'metaphysical' necessity. This is the sense of necessity that obtains in virtue of the identity of things (broadly conceived). Thus, in this sense it is necessary not only that anything red is red or that nothing is both red and green, but also that I am person or that 2 is a number.

Depending upon which notion of necessity one starts with, there are two main strategies for defining the other notions of necessity.[2] Suppose one starts with the narrow notion of logical necessity (or with some other suitably narrow notion). The main problem will then be to define the broader notions of necessity; and the obvious way to do this is by relativization. Consider the case of conceptual necessity—the necessity that holds in virtue of the identity of concepts. It will be necessary in this sense that nothing is both red and green, though not necessary that I am a person. Now let it be granted that there are some basic conceptual truths—perhaps given by the *definitions* of the various concepts—and that the class of such truths can be defined without appeal to any modal notions (besides logical necessity). We might then define a proposition Q to be a conceptual necessity if it follows from the definitions: that is, if the conditional, 'if P then Q', is logically necessary for some conjunction P of basic conceptual truths. The conceptually necessary truths, in other words, may be taken to be those that are logically necessary *relative to*, or *conditional upon*, the basic conceptual truths.

Suppose, on the other hand, that one starts with the broad metaphysical notion of necessity (or with some other suitably broad notion). The main

---

[1] I take necessity to be a feature of propositions, though nothing of any importance for my purposes will turn on this assumption.

[2] I do not wish to suggest that these two strategies represent the only possible ways of defining one notion of necessity in terms of others.

problem will then be to define the narrower notions of necessity; and the obvious way to do this is by restriction. Consider the case of mathematical necessity, the form of necessity that pertains to the truths of mathematics. We may then define a proposition to be mathematically necessary if it is necessary in the metaphysical sense and if, in addition, it is a mathematical truth—where this latter notion is presumably one that can be defined in non-modal terms. In this case, the new form of necessity is defined by means of a restriction that can be stated in non-modal terms (or, at least, without appeal to further modal notions).

I am inclined to think that the second of the two strategies *can* successfully be pursued. Given the notion of metaphysical necessity, the various narrower notions of necessity—be it logical, mathematical, conceptual, or the like—can each be defined by restriction.[3] Each of them can be regarded as a *species* of metaphysical necessity.

The feasibility of the first strategy, however, is open to serious doubt. It is, in the first place, not at all clear that metaphysical necessity can be defined in terms of logical necessity; for it is not clear that one can provide a non-modal characterization of some basic metaphysically necessary truths from which all other metaphysically necessary truths will be a logical consequence. But even if one sets this problem aside and allows the use of both logical and metaphysical necessity, there would appear to be concepts of necessity that are broader still, yet equally resistant to definition.

The two main concepts of this sort are the concepts of natural and normative necessity;[4] and it is my aim in the rest of the chapter to show how these concepts raise serious problems for the doctrine of modal monism. There are two main ways, in either case, in which the doctrine might be defended. It might be denied that either of the other concepts of necessity is genuinely broader than the metaphysical concept; natural and normative necessity should be regarded as *restricted* forms of metaphysical necessity. Or it might be maintained that the other concept is indeed broader, yet definable as a *relative* form of metaphysical (or logical) necessity. I have attempted to show, in each case, that neither line of defence can be made to work.

It is important to bear in mind some limitations in my approach. First, I have not directly addressed the question of whether there might be some other concepts of necessity that cannot be understood in terms of the three

---

[3] Some of my reasons for thinking this are outlined in Fine ([1994], 9–10), though much more needs to be said on the question.

[4] Another possible candidate is the concept of historical necessity, that form of necessity for which the past is 'closed' yet the future may be 'open'. The interesting question of whether the concept of natural necessity is merely a special case of this other concept is not one that I shall consider. Nor do I consider the epistemic, deontic, or tense-logical modalities, since I do not view them as constituting genuine forms of necessity.

upon which I have focused. One should think of the discussion as representing an 'end-game' in which the other candidate concepts of necessity have been removed from the board. Second, I have not considered all possible ways in which one of the remaining concepts of necessity might be defined or understood in terms of others. But if there are others, then I do not know what they might be. Finally, my concern throughout has been to arrive at the most basic modal concepts: that is, those that are not to be defined or understood in terms of other modal concepts. I have not directly considered the question of whether it might be possible to break out of the sphere of the modal and understand it in altogether different terms. Thus my conclusions have no direct bearing on the issue of modal realism—that is, on whether the modal facts are themselves most real—for I may merely have tracked modality down to its penultimate source, within the sphere of the modal, rather than to a possibly more ultimate source.[5]

2.   Natural Necessity: Subsumption

Natural necessity is the form of necessity that pertains to natural phenomena.[6] Suppose that one billiard-ball hits another. We are then inclined to think that it is no mere accident that the second billiard-ball moves. Given certain antecedent conditions and given the movement of the first ball, the second ball *must* move. And the 'must' here is the *must* of natural necessity.

The above elucidation of natural necessity does not presuppose that the notion has primary application, or even *any* application, to natural law. However, it is very plausible to suppose that if there are particular necessary connections of the above sort, then there are also general necessary connections of this sort. Thus, not only will it be necessary that this billiard-ball move in these particular circumstances, it is also necessary that any billiard-ball will move in relevantly similar circumstances.

What is the relationship between metaphysical and natural necessity? Is every natural necessity a metaphysical necessity? And is it therefore possible to regard the one form of necessity as a restricted form of the other?

The answer to these questions would appear to be a straightforward 'No'. For surely it is conceivable, and hence metaphysically possible, that the one ball should strike the other in the given circumstances without the other moving. And surely it is conceivable, and hence metaphysically possible, that many of the natural laws that govern our universe should fail to hold,

---

[5] All the same, my conclusions may provide some succour for the realist concerning modality. For realism about possible worlds will not be plausible given that there are different primitive notions of necessity; and a reductive form of anti-realism will not be plausible given that the modal does not supervene on the nonmodal.

[6] Some of the issues of this section are further discussed in a paper of Sidelle [2002].

that bodies should attract one another according to an inverse cube law, for example, rather than the inverse square law.[7]

However, ever since Kripke (1980), we have learnt to be suspicious of such considerations. For can we be sure that the hypothetical situation in which an inverse cube law is envisaged to hold is one in which the bodies genuinely have mass? Perhaps they have some property like mass, call it schmass, which conforms to an inverse cube law. And can we be sure that the hypothetical situation in which the second billiard-ball is envisaged not to move is one which genuinely contains the given billiard-balls rather than some schmassy counterparts?

In either of these cases, the proposed counterexample would fail; and if the same is true for any other counterexample that might be proposed, then the way would be clear towards maintaining that every natural necessity was a metaphysical necessity. Indeed, several philosophers have recently been attracted towards such a view[8] and it might also be thought to be especially congenial to my own way of thinking. For I take metaphysical necessities to be those that are rooted in the identity of 'things' (Fine [1994], 9); so natural necessities might then be taken to constitute the special case in which the things in question are the natural properties or kinds. Natural necessities would simply be the special case of those essentialist truths that arise from the identity of natural kinds.

However, it seems to me that the scope of these counter-considerations is severely limited, and that the restrictionist view remains highly problematic.[9] It may be conceded that we should exercise caution in judging a natural necessity to be metaphysically contingent—for what is taken in a given hypothetical situation to be a property or kind that figures in the natural necessity may be no such thing. But this, I believe, should merely lead us to adopt a more discriminating view as to which natural necessities are metaphysically contingent, rather than to give up the idea that there are any such necessities.

To see why this might be so, let us return to the putative counterexample to the metaphysical necessity of the inverse square law; and let us concede that the envisaged hypothetical situation involves schmass, rather than mass, and that the counterexample therefore fails. Still, that very same hypothetical situation may be used to provide a counterexample to the metaphysical necessity of a *different* natural necessity. For consider the proposition that there is no schmass (i.e. that there are no instances of schmass). Then this proposition should be taken to be a natural necessity. For our original

---

[7] For expository purposes, I take an oversimplified view of what the scientific laws are.

[8] They include Shoemaker ([1980], 244; 1998), Swoyer [1982], and Ellis [1999]. Kripke raises the issue ([1980], 99, 164) but without taking a stand.

[9] Further criticisms of the subsumptionist view are made in Sidelle [2002].

judgement was that the inverse square law was a natural necessity, though not a metaphysical necessity. Now that we see that the metaphysically possible worlds in which it was taken to fail are ones with schmass rather than mass (and given that our universe is taken to be completely governed by the Newtonian Laws), we should take it to be a natural necessity that there is no schmass. In either case, the 'fabric of the universe' is envisaged as excluding a certain sort of behaviour—whether this be the deviant behaviour of mass or the normal behaviour of schmass. Moreover, my opponent should concede it to be a metaphysical possibility that there is schmass, since it was through postulating schmass—or the like—that the original putative counterexample to the metaphysical necessity of the inverse square law was reinterpreted. So he should grant that the absence of schmass is a natural, though not a metaphysical, necessity.

Indeed, there is no reason in general why the sophisticated post-Kripkean should not agree with the naive pre-Kripkean as to which of the metaphysically possible worlds are naturally impossible. For whereas the pre-Kripkean will take such a world to be a natural impossibility because of the straightforward failure of a law, the post-Kripkean will take it to be a natural impossibility because of the instantiation of an alien property or kind. Thus even though sensitivity to the cross-world identity of natural properties or kinds may lead one to redescribe the hypothetical situations in which a natural law is taken to fail, it should not lead one to reject the natural impossibility of those situations.

It might be objected that there is not even a *putative* counterexample to the metaphysical necessity of the inverse square law. But such a view is too outlandish to deserve consideration;[10] and once we have the putative counterexample, then we have the basis, if I am right, for deriving an actual counterexample. It might also be objected that the proposition that there is no schmass remains true in the hypothetical situation in which the inverse square law is thought to fail, since it means that there is no *body* with schmass and, in the hypothetical situation, there are only schbodies, not bodies. Thus 'body' goes the way of 'mass'. But, if that is the objection, then let us formulate the proposition that there is no schmass with an absolutely unrestricted quantifier: there is nothing whatever with schmass. Or, alternatively, we might use the proposition that there are no schbodies (again with an unrestricted quantifier).

A more serious objection concerns the existence of the relevant properties or kinds. It might be thought that the properties or kinds that figure in natural law are *immanent* in the sense of only existing if instantiated. The

---

[10] Perhaps even for someone like Shoemaker ([1998], n. 11), who believes that nothing but H_2O could behave the way water ordinarily behaves, since what is at issue here is whether some alternative to water might behave in some other way.

kind *schmass* will therefore not exist. It might also be thought that a proposition exists only if the items it directly concerns exist. So, since the kind *schmass* does not exist, nor does the proposition that there is no schmass; and so we have no counterexample to the subsumption of natural under metaphysical necessity.

Whether this is so depends upon exactly what the subsumption thesis is taken to be. If it is the thesis:

> Every (actual) proposition is such that it is natural necessity only if it is metaphysical necessity,

then no counterexample has been given under the stated assumptions. But if it is the thesis:

> Necessarily$_M$ every proposition is such that necessarily$_M$ it is a natural necessity only if it is a metaphysical necessity,

or even the weaker thesis:

> Necessarily$_M$ every proposition is such that actually it is a natural necessity only if it is a metaphysical necessity,[11]

then there is a counterexample. For the proposition that there is no schmass exists in the hypothetical situation in which there is schmass, and this very proposition is a natural necessity in the actual world, though not a metaphysical necessity. Moreover, in standard formulations of modal logic, it is the stronger theses that are required if natural necessity is to be eliminable in favour of metaphysical necessity.

In any case, there are counterexamples that require no appeal to uninstantiated properties or kinds. Let P, Q, . . . be an exhaustive list of all the kinds (or all the fundamental kinds) that there actually are. Then presumably it will be a natural necessity that every object (or every fundamental object) is of one of the kinds P, Q, . . ., but it will not be a metaphysical necessity. Or again, suppose that determinism is true and holds of natural necessity: it is a natural necessity that every event has a cause. (Or, if we wish to avoid appeal to the notion of cause, we can say: it is a natural necessity that for any event *e* there is a preceding event *c* such that it is a natural necessity that *e* occurs if *c* occurs.) But surely it is a metaphysical possibility that determinism is false. It would be absurd for my opponent to maintain that the hypothetical situation in which determinism appears to fail is one that does not really involve events or time. Thus, given that it is a metaphysical possibility that determinism should hold, we have the metaphysical possibility of a natural necessity not being a metaphysical necessity.

---

[11] I use the subscripts 'M' (and 'N'), here and elsewhere, to indicate the kind of necessity in question.

We can even construct a counterexample on the basis of standard laws. Consider the inverse square law as an example. Now my opponent will maintain that this law is still true in the hypothetical situation in which there is schmass rather than mass, though vacuously. But surely he will concede that, even though the law is true in this hypothetical situation, it is not a *law*—or, at least, not a law that *prevails*—in that situation.[12] Indeed, if it were, then, by parity of reasoning, the inverse cube law for schmass would have to be a law that prevails in our world; and surely it is not. But now, whatever it takes to be a prevailing law, it seems clear that it is a natural necessity that the inverse square law is such a law. Not only is there no natural possibility of its failing to hold, there is no natural possibility of its failing to be a law that governs the universe. But then the proposition that the inverse square law is such a law is another actual counterexample to the thesis that every natural necessity is a metaphysical necessity.

The lesson to be learnt from these counterexamples is not that we should go back to our pre-Kripkean intuitions of metaphysical contingency, but that we should attempt to be more discriminating about which laws of nature are to be regarded as metaphysically contingent and which are not. There is an intuitive distinction to be drawn here. That electrons have negative charge, for example, strikes one as metaphysically necessary; it is partly definitive of what it is to be an electron that it should have negative charge. But that light has a maximum velocity or that energy is conserved strikes one as being at most naturally necessary. It is hard to see how it could be partly definitive of what it is to be light that it should have a given maximum velocity, or partly definitive of energy that it should be conserved.[13] It is equally a defect of the old view that saw all laws of nature as metaphysically contingent and of the new view that sees them all as metaphysically necessary that they fail to heed this distinction; rather than take a blanket view of the modal status of these laws, we should attempt to refine and systematize the intuitive discriminations that we are naturally inclined to make among them.[14]

---

[12] Cf. Shoemaker ([1980], 248): 'Nothing I have said precludes the possibility of there being worlds in which the causal laws are different from those that prevail in this world.'

[13] Lowe [2001] has also stressed the metaphysical contingency of the values born by the fundamental physical constants, and Chalmers ([1999], 13–14) has stressed the metaphysical contingency of the conservation laws.

[14] Thus I do not share Shoemaker's scepticism on this point ([1980], 249–51). It is not that we need a general criterion for saying when we have one kind of necessity as opposed to another but a clearer conception of what, in particular cases, might plausibly be taken to be relevant to the identity of a given natural property. Where he sees a problem, I see an interesting project. The present distinction is somewhat akin to the Kantian distinction between the 'pure' and 'empirical' parts of science, which was later taken up by some of the logical positivists (see Friedman [1994] for a general discussion).

Although I have emphasized the way in which natural necessities may outrun the metaphysical necessities, it seems to me that there is one respect in which this may not be true. For I am inclined to think that there are no distinctive *de re* natural necessities. Let us suppose that x and y are two particles, and that it is a natural necessity that they attract one another (assuming, of course, that they exist!). Then it is plausible to suppose that this should follow from (1) its being a metaphysical necessity that each of the particles is of the kind that it is and (2) its being a natural necessity that particles of this kind attract one another. Thus the *de re* natural necessity will reduce to a *de re* metaphysical necessity and a *de dicto* natural necessity; and it might be thought that something similar should be true of any *de re* natural necessity or, indeed, of any form of *de re* necessity whatever. All forms of *de re* necessity (and of essence) will be fundamentally metaphysical, even though some forms of *de dicto* necessity may not be.

3. Natural Necessity: Definition

Even if post-Kripkean sensitivity to the cross-world identity of natural kinds does not enable one to *subsume* natural necessity under metaphysical necessity, it might still appear to hold out the hope of defining it as a *relative* form of metaphysical necessity. For suppose we uphold the doctrine of immanent universals. We may then let the existence of natural properties or kinds be our guide to the natural possibilities for a given world, a possible world being a natural possibility relative to a given world if it contains only (or perhaps all and only) those natural kinds that exist in the world.[15] A world of schmass, for example, will not be a natural possibility, since the kind *schmass* does not actually exist; and, in general, any objects that behaved in a nomically irregular way within a given world would have to be of kinds that do not actually exist, and hence would belong to a world that was not a natural possibility. (And, of course, once given the naturally possible worlds, we can define the natural necessities as those that hold in every such world.)

Instead of presupposing the doctrine of immanent universals in formulating the definition, as is often done, we may appeal instead to what is taken to be required for a kind to exist. Thus we may say that a world is a natural possibility if it instantiates only those kinds that are actually instantiated, and thereby sidestep the issue of the conditions under which a universal exists. Nor is there any need to place such emphasis on instantiation as the condition for the existence of universals. Perhaps we can allow kinds to exist in the manner of Hume's missing shade through being suitably related to

[15] Clearly, the natural kinds should also be taken to include the various fundamental physical relations.

other kinds that exist, even though they are not themselves instantiated. If we free up the account in both these respects, then we are left with the general idea that the natural possibilities for a given world will turn upon the status and distribution of its natural properties and relations.

But accounts of this sort, it seems to me, are subject to a familiar form of objection. It is sometimes pointed out that two possible worlds might *merely* differ as to what is a natural necessity and that regularity-type views must therefore be mistaken, since they would be unable to distinguish between the two worlds. This objection will not work against the present view, since it might be argued that any difference in the natural laws would make a difference to the natural properties that exist in the two worlds. But a variant of the objection *can* be made to work.[16]

Consider, for example, a metaphysically possible world w_N that is Newtonian. Then bodies in this world will have mass, be subject to force, and so on (or have something similar to mass and be subject to something similar to force, since actual mass is not itself strictly Newtonian). By the same token, there will be a metaphysically possible world w_M which is Schmewtonian. The bodies in this world behave like bodies in w_N but are subject to the inverse cube law or some other variant of the Newtonian Laws. The bodies in this world will not have mass, according to our opponent, but they will have something similar to mass, say schmass; and likewise for force and the rest. Now surely it is a natural possibility in both w_N and w_M that there be no bodies; after all, there is nothing in the natural laws of either world that requires that there be anything to which they apply. So there is going to be an empty world v_N that is a natural possibility for w_N, and an empty world v_M that is a natural possibility for w_M. Since v_N is a natural possibility for w_N, it will verify all of the natural necessities of w_N; so, since it is a natural necessity in w_N that there is no schmass, it will be a natural necessity in v_N that there is no schmass.[17] Moreover, since the world w_M contains schmass, we may safely assume that it is a natural possibility in the empty world v_M that there be schmass; for it would be bizarre in the extreme to suppose that the non-existence of any bodies somehow precluded the possibility of there being schmass.[18] So the empty worlds v_M and v_N differ as to what is a natural possibility. But it is hard to see how there can be any difference in the status of their natural properties; for the natural

---

[16] Carroll ([1994], §3.1) advances a similar line of objection, though without attempting to take care of the rejoinder that the two worlds might differ in their natural properties. Similar objections to Humean accounts of objective chance have also been considered in the literature.

[17] I have assumed that natural necessity is subject to the S4 axiom, $\Box A \to \Box \Box A$. But even without the benefit of this assumption, it would be odd to suppose that, in w_N, the non-existence of bodies somehow required the possibility of there being schmass.

[18] Alternatively, we could appeal to the assumption that natural necessity was subject to the S5 axiom, $A \to \Box \Diamond A$, though nothing so strong is required in this particular case.

properties that exist in the two worlds and their pattern of instantiation are just the same.[19]

A similar counterexample (though not subject to worries over empty space-time) runs as follows. Consider a metaphysically possible world w_D for which mind–body dualism is true. The world w_D may not consist of mental and physical events as we conceive them, but it will then consist of related kinds of events—the mental$_D$ and the physical$_D$, say. Let us suppose that epiphenomenalism is also true in w_D, so that the mental$_D$ and the physical$_D$ events of w_D are each subject to their own laws, but with no nomological interaction between them. By the same token, there should be a metaphysically possible epiphenomenal world w_E in which the physicalistic events are subject to essentially the same laws as in w_D, but the mentalistic events to somewhat different laws. It is reasonable to assume, or at least to allow, that the physicalistic events of w_D and w_E are of the same kind, even though the mentalistic events are not.

Now surely it is a natural possibility in both w_D and w_E that, under given physical conditions, there be nothing mentalistic in the world. Thus there will be a mind-free world v_D that is a natural possibility for w_D, and a physically similar mind-free world v_E that is a natural possibility for w_E. But then by the same line of reasoning as before, w_D and w_E will differ on what is a natural possibility (for the mentalistic part of the world), even though there is no difference in the 'status' or distribution of their natural properties.

Of course, if these counterexamples are correct, then they tell not only against the property-based definitions, but also against any other account that would make the natural possibilities supervene, as a matter of metaphysical necessity, upon the non-nomic facts.

There is, however, another, more radical objection to be made. So far I have argued that any definition of natural in terms of metaphysical necessity will be extensionally incorrect—there will be a difference, or at least a possible difference, in the propositions that fall under the definiendum and those that fall under the definiens. But it might be argued that even if we had an extensionally correct and non-circular account of natural necessity, it still would not be likely to provide an adequate definition.

We may illustrate the nature of the difficulty with the doctrine of logical fatalism. Suppose one holds, for whatever reason, that every truth is necessary. Then:

(*) for every proposition p, p is necessary iff it is true;

---

[19] One might maintain that the kind mass exists in v_N but not in v_M, but that is presumably only because the instantiation of mass is a natural possibility in the one but not the other, and so the concept of natural possibility is already presupposed.

and since this proposition is itself true, it follows:

(**) necessarily, for every proposition p, p is necessary iff it is true.[20]

But even the logical fatalist will not accept (**) as a correct *definition* of necessity, despite the presence of necessary coincidence and the absence of circularity, since it will be important for him to maintain that the necessity of a proposition does not *consist* in its being true. It so happens, if I may put it this way, that every true proposition is necessary; but the proposition's being true is not that in which its necessity consists.

Another, though somewhat more problematic, case is provided by the standard definition of logical necessity (narrowly conceived) in terms of invariance. For let it be granted, if only for the sake of argument, that:

Necessarily$_M$, a proposition is logically necessary iff its truth is preserved under any substitution for its non-logical constituents.

Still, it might be maintained that such invariance is not what it is for a proposition to be logically necessary. After all, the proposition that B.C. is not an angel remains true under any substitution for the constituent B.C. but is not, on that account, a necessary truth. So why should it be any different in the logical case? What we have *at best*, on this view, is a definition of logical *truth*, rather than of logical *necessity*.

One might even argue against my proposed definition of metaphysical necessity in terms of essentialist truth along similar lines (Fine [1994], 9). I wish to claim:

Necessarily$_M$, a proposition is metaphysically necessary iff it is true in virtue of the identity of some (possible) objects.

But it might be argued that what we have on the right-hand side is merely an account of the source of the proposition's truth and not of its modal status. Essentialist truth is no more capable than logical truth of conveying modal import.[21]

A similar problem, I suspect, is bound to arise for any proposed definition of natural necessity in terms of metaphysical necessity. For it will usually be possible to see such a definition as a case of relativization. Certain propositions will be picked out by means of a suitable description, call it 'being a law'; and a proposition is then taken to be a natural necessity iff it is entailed

---

[20] I here ignore the difficulties over including merely possible propositions within the scope of the definition.

[21] Another illustration of the distinction is provided by Quine's arguments against analyticity, which have as their principal target a certain kind of truth, rather than a peculiarly modal status. Almog [1991] draws a similar distinction between a 'primal', or constitutive, truth and its modal import.

by the propositions that satisfy the description—that is, by the laws.[22] But such an account is subject to the obvious objection that it does not provide an adequate account of the natural necessity of the 'laws' themselves. For where the proposition P is a law, its being a natural necessity, according to the definition, will consist in: (1) its being entailed by the various 'laws', including P itself; and (2) its being a law. But (1), which is merely a matter of self-entailment, can hardly contribute to the given proposition's being a natural necessity; and it will be hard to see, in any given case of (2), how the defining feature of a 'law' might constitute an adequate account of the *necessity* of the given proposition. Consider the definition proposed above by way of illustration. This may be put in the form:

> a proposition is a natural necessity iff it is entailed by the proposition that $K_1, K_2, \ldots$ are the only kinds that there are,

where $K_1, K_2, \ldots$ is an inventory of all the kinds that there are. The 'law' here is the proposition that $K_1, K_2, \ldots$ are the only kinds that there are, and its being a 'law' essentially consists in its being true. But we are inclined to think that, in so far as it is a natural necessity that there are no other kinds, it is because there is something in the nature of the world that prevents there being other kinds; and the mere fact that there *are* no other kinds can hardly be taken to constitute an adequate account of what this force, or form of necessity, might be.

The general problem is that a definition of natural necessity as a form of relative necessity will tend to make the necessity of the propositions with respect to which the necessity is relative a trivial or insubstantial matter; yet we are inclined to think that the necessity attaching to the laws and the like is not of this trivial sort. Any true proposition whatever can be seen as necessary under the adoption of a suitable definition of relative necessity. Any proposition that I truly believe, for example, will be necessary relative to the conjunction of my true beliefs, and any proposition concerning the future will be necessary relative to the conjunction of all future truths. The problem therefore is to explain why the necessity that issues from the definition of natural necessity is not of this cheap and trivial sort; and I doubt, in the case of any otherwise reasonable definition that might be proposed, that this can be done.

One might wish to press the objection further and claim that no definition stated entirely in terms of metaphysical necessity could capture the peculiarly modal force of truths that are naturally necessary yet metaphysically contingent. Just as it has been supposed that there is a conceptual barrier

---

[22] I have supposed that the laws are picked out by a description that is external to the entailment, but one might also provide an analysis of the form $L \rightarrow P$, as long as one is prepared to resort to double indexing, as in van Fraassen [1977].

between normative and non-normative concepts, so one might think that there is a conceptual barrier, not merely between modal and non-modal concepts, but also between different 'grades' of modality. But even though I would wish to endorse this more general claim, there is no need to appeal to it in arguing against the plausibility of particular accounts of what this peculiar 'modal force' might be.

I conclude that there appears to be no reasonable way of understanding natural necessity as a restricted or relative form of metaphysical necessity.[23]

4.   Normative Necessity: Naturalism

There is a familiar distinction between accidental and non-accidental generalizations within the natural sphere, but what is not so often appreciated is that a similar distinction can be drawn within the moral sphere. This may be illustrated by the claim that every war is wrong. For this might be meant in the sense that every war, in the circumstances that actually prevail, is wrong; or it might be meant in the sense that every war, in whatever circumstances might prevail, is wrong. In the latter case, the claim is taken to be necessary—to hold unconditionally, or in all possible circumstances; while in the former, the claim is not taken to be necessary, but merely to hold conditionally upon the circumstances that actually obtain.

The distinction between accidental and necessary generalizations in nature is often drawn in terms of the ability to sustain counterfactuals. A necessary generalization that all Fs are Gs will sustain the counterfactual 'if this were to be an F it would be a G', while the corresponding accidental generalization will not. The distinction may be drawn on a similar basis in the moral case. For the *de facto* pacifist need not commit himself to the view that if there were a war of such-and-such a hypothetical sort, then it would be wrong, though the more radical pacifist will be so committed. Indeed, it is perhaps only in so far as moral judgements bear this counterfactual force that they can be of any real help as a guide to action; for even if we do not do something, we still wish to know whether it would have been better if we *had*.

The sense of necessity in which the radical pacifist wishes to maintain that it is necessary that any war is wrong I propose to call *normative*. I am inclined to think, as the term 'normative' suggests, that the same kind of necessity has application to other normative domains; but I shall bracket this question in what follows and simply focus on the moral case.

---

[23] Among recent theorists, Armstrong ([1983], 92–3) and Fales ([1993], 140) have been attracted by the view that some form of natural necessity or necessary connection might be primitive.

It is in this sense of necessity that the moral supervenes on the natural, and, indeed, such cases provide the least contentious examples of normative necessity. Suppose that D is a complete description of the world in naturalistic terms. Then we will be inclined to make certain moral judgements about the world so described—that such-and-such a consequence was unfortunate or such-and-such an action wrong. But in so far as we are prepared to make such judgements, we will also be prepared to say that it was no accident that they are true. In those particular circumstances, the consequences *had* to be unfortunate, the action *had* to be wrong.

It is perhaps only because moral truths may hold with this kind of necessity that it is appropriate to talk of ourselves as being subject to moral *law*. For just as we are inclined to think that if one billiard-ball hits another in given circumstances, then the other *must* move, so we are inclined to think that if I make a promise to someone in given circumstances, then I *must* keep the promise. And here the 'must' is not merely the 'must' of obligation. I am obliged to keep the promise, but that I am so obliged is something that is *required* by my having made the promise in the first place. The obligation is itself something that falls under the rubric of necessity.

How should normative necessity be understood? Is it a species of natural or metaphysical necessity? Or somehow definable in terms of these other forms of necessity?

It seems bizarre to suppose that normative necessity is a species of natural necessity. Indeed, it is commonly held that there are no natural necessities that essentially involve normative concepts. But from this we would hardly wish to conclude that there are no non-trivial normative necessities.

Whether normative necessity is a species of metaphysical necessity is more contentious. One reason for thinking that it is derived from the traditional doctrine of naturalism, according to which any moral property will be coextensive, as a matter of conceptual necessity, with some natural property.[24] In order to see how the argument from the one to the other might go, let us suppose that a given proposition *P*, say that lying is wrong, is a normative necessity. This may be symbolized as follows:

(1) $\boxed{n}\, P\,(\mathrm{W})$,

with the predicate 'W' for 'wrong' made explicit. Given naturalism, it is a conceptual necessity that wrongness is coextensive with a certain natural property *N*:

(2) $\boxed{c}\, (x)\,(\mathrm{W}x \leftrightarrow \mathrm{N}x)$.

[24] The traditional form of naturalism is to be distinguished from the more contemporary form, in which all that is required is that the moral 'supervene' on the non-moral. Under certain assumptions, which need not be subject to doubt in the present context, the two will be equivalent.

Let us use the notion of normative necessity in an inclusive sense so as to include all of the conceptual necessities. Or, to put the matter differently, we shall not allow something to be a normative possibility unless it is also a conceptual possibility. It then follows from (2) that wrongness and the natural property N are coextensive as a matter of normative necessity:

(3) $\boxed{\text{n}}$ (x) (Wx ↔ Nx).

Since this is so, one may be substituted for the other in (1), and we obtain that it is a normative necessity that lying has the naturalistic property N:

(4) $\boxed{\text{n}}$ P(N).

But that lying has the property N is a purely naturalistic proposition, and so, given that it is a normative necessity, it must also be a conceptual necessity; for normative necessity merely serves to restrict the *connection* between the naturalistic and the normative possibilities, it does not serve to restrict the naturalistic possibilities themselves. So:

(5) $\boxed{\text{c}}$ P(N).

But again, given, by (2), that W and N are coextensive as a matter of conceptual necessity, one may be substituted for the other in (5); and we obtain that it is a conceptual necessity that lying is wrong:

(6) $\boxed{\text{c}}$ P(W).

In this way, any normative necessity can be shown to be a conceptual necessity (and hence also to be a metaphysical necessity under the traditional view).

The argument rests on two general assumptions. The first, which we may call *Inclusion*, is that every conceptual necessity is a normative necessity. Or, in schematic form:

Inc. $\boxed{\text{c}}$ A → $\boxed{\text{n}}$ A.

The second, which we may call *Conservativity*, is that every naturalistic normative necessity is a conceptual necessity:

Cons. $\boxed{\text{n}}$ A → $\boxed{\text{c}}$ A, for A naturalistic.

If these assumptions are themselves taken to hold of conceptual necessity, then it may be shown to be a conceptual necessity that any given normative necessity is a conceptual necessity.

Ever since Moore, however, most moral philosophers have taken naturalism in its traditional form to be an instance of the 'naturalistic fallacy'; and if, as I think, they are right, then the present route to subsuming the normative notion of necessity under the conceptual notion will be blocked. This is not the place to attempt a vindication of this objection to naturalism,

but let me make a few comments on how I think it might best be understood. Moore thought that if there were an analysis of an ethical property in terms of a naturalistic property, then it would no longer be an open question whether things with the naturalistic property had the ethical property. But, as has often been pointed out, it may not be obvious that a correct analysis is indeed correct, so such an analysis could still leave open the question of the connection between analysandum and analysans.

Perhaps a more satisfactory way to formulate the objection is as follows.[25] If there is a correct analysis of *good*, say, as what promotes pleasure over pain, then something's being good must consist *in nothing more* than its promoting pleasure over pain.[26] But we have a strong intuition that it *does* consist in something more. Here we are not relying on the purported *epistemic* status of a correct analysis, as is Moore, but on its *metaphysical* consequences.

This argument, moreover, can be strengthened. For suppose one merely takes it to be a conceptual necessity that something is good if it promotes pleasure over pain. Now, if this is true, then presumably it must also be true that something is good in virtue of promoting pleasure over pain. Indeed, it is only because something is good in virtue of promoting pleasure over pain that there is the conceptual connection between the one and the other. But now what is this in-virtue-of relationship that accounts for the conceptual connection? The only possible answer, it seems, is that it is the relationship of one thing *consisting in no more than* some other; for this would appear to be the only in-virtue-of relationship capable of sustaining a conceptual connection. But if this is right, then the argument can also be taken to apply to statements of conceptual implication, and not merely to analyses.

5.   Normative Necessity: Neo-naturalism

Many philosophers have recently been willing to grant that normative necessity is not a form of conceptual necessity, but have been tempted, all the same, by the view that it is a form of metaphysical necessity.[27] This alternative view does not appear to have been based upon any serious

---

[25] A rather different way to obtain a version of the open question argument is to 'reverse' the argument given above. For suppose we reject the conclusion of the argument: i.e. take it to be a conceptual possibility that something holds of normative necessity though not of conceptual necessity. Then, granted the conceptual necessity of the assumptions (Inc.) and (Cons.), it follows that (2), the doctrine of naturalism, will be false. This also strikes me as being a powerful objection to naturalism.

[26] The notion of *consists in*, which I appeal to at various places in the chapter, is discussed at greater length in Fine [2001].

[27] Their number includes Dreier ([1992], 15), Klagge ([1984], 378), McFetridge ([1985], 251–2), Shoemaker ([1985], 441), and Zangwill [1995].

consideration of the matter. It is observed that there is a necessary connection between the naturalistic and the normative features of a given situation, and it is simply assumed, given that the connection is not conceptual, that it must be metaphysical. These philosophers sometimes appeal to the fact that the connection holds in all possible worlds, but it is only if these worlds are themselves taken to be metaphysically possible that the metaphysical necessity of the connection would thereby be established. No insight into the status of the necessary connection is to be gained in this, or any other, case by an appeal to possible worlds.

If metaphysical necessity is taken to be that form of necessity that derives from the nature of things, then it is prima facie highly implausible that the necessary connection between the naturalistic and normative features of a given situation should be taken to be metaphysical. For there would appear to be nothing in the identity of the naturalistic or normative features that demands that they be connected in the way they are. It is no part of what it is to be pain that it should be bad, and no part of what it is to be bad that it should include pain. There is a striking, intuitive difference between the connection between being water and being composed of H_2O, on the one hand, and the connection between being a pain and bad, on the other. For the identities of the respective features require that the connection holds in the one case, though not the other. I might also note that my previous argument against normative necessity as a form of conceptual necessity would appear to work equally well against its being a form of metaphysical necessity, since it is hard to see how a metaphysically necessary connection between the naturalistic and the normative could hold without the latter simply consisting in the former.

There is, however, a way in which this line of reasoning might be resisted. For it might be maintained that the normative features may have a 'hidden' nature, and that, once it becomes clear what this is, it will be apparent how these features may be connected, as a matter of metaphysical necessity, with appropriate naturalistic features. The elaboration of such a view is to be found in the new 'metaphysical' version of naturalism. This differs from the old 'conceptual' version of naturalism in two main respects. First, the naturalistic property in terms of which 'good', or what have you, is analysed is a high-level 'functional' property, rather than a low-level 'criterial' property. Second, the extension of the term 'good' is taken to be 'fixed' by means of such a property. Thus an analysis of 'good' (one that is meant to reveal our understanding of the term) may be put in the following general form:

(*) for any x, x is good iff x has the property that actually fits the good-making role.

We might suppose, for example, that for a property to fit the good-making role is for it to be what is valued under ideal conditions of valuation, and

that the property that actually fits this role is the property of promoting a balance of pleasure over pain. It would then be a metaphysical necessity that:

(**) for any x, x is good iff x promotes a balance of pleasure over pain.

However, in contrast to traditional versions of naturalism, (**) would be a posteriori, since it would be an a posteriori matter that the property of promoting a balance of pleasure over pain is what is valued under ideal conditions of valuation.[28]

An immediate consequence of this view is that it enables one to see normative necessity as a straightforward case of metaphysical necessity. For the normative necessities are merely those metaphysical necessities that arise from looking at the naturalistic content of the ethical predicates, without regard to how that content might have been fixed. Thus (**) above, the paradigm normative necessity from which all others follow, will also be a metaphysical necessity. The view is also able to preserve the distinction between the substantive criteria for goodness and what it is to be good; and it is not subject to our previous argument for the collapse of normative and conceptual necessity, for it is readily shown that the Inclusion assumption should be given up. Indeed, if it were to hold—that is, if each conceptual necessity were taken to be a normative necessity—then the Conservativity assumption would fail.[29]

However, I believe that the view is still open to serious objection.[30] Even though it may be capable of yielding a better account of the metaphysics of morality than the traditional naturalist account, it is still not capable of yielding a satisfactory account of its epistemology. For it is unable properly to respect the non-empirical character of ethical belief.

In explaining what this is, there are two main problems that need to be addressed. The first is that of which ethical judgements should be taken to be non-empirical. The second is: in what should their non-empirical character be taken to consist. It clearly will not do to say that *all* ethical judgements are non-empirical; for that Joey did something wrong yesterday is clearly an empirical judgement, since its truth rests upon the empirical naturalistic fact that he did one thing rather than another. We might attempt to pre-empt the relevance of empirical judgements in this way by making the ethical judgements in question conditional upon a complete description of

---

[28] Wiggins ([1987], 206), Lewis ([1989], 132), M. Smith ([1994], 190–2), and Jackson ([1998], 143) are among those who have been tempted by a view of this sort.

[29] Indeed, necessity will collapse to truth. For suppose that S is the case. Then it is a conceptual necessity that the actual truth-value of S is the truth-value of S, and a metaphysical necessity that the actual truth-value of S is True. So if the two necessities could be combined, it would be necessary that the truth-value of S is True, and hence necessary that S.

[30] As will become clear, the objection is also likely to apply to versions of naturalism that insist on functionality without also insisting on rigidity.

the facts. Thus an ethical judgement will now take the form: if this is how things are naturalistically (there follows a complete naturalistic description of how things might be), then this is how things are morally (there follows a partial account of how things should be). However, such a judgement might be true, not because the consequent is true, but because the antecedent is false, and in this case there would be no reason to expect it to be non-empirical. In order to get round this problem, I propose that we consider instead the normative necessity of the above judgement. Thus the judgement is now of the form: it is (normatively) necessary that if this is how things are, then this would be good (say); or, to put it in counterfactual terms, if this is how things *were*, then this *would* be good.[31] Call such judgements *world-bound normative conditionals*. Then one formulation of the claim that ethics is non-empirical is that every world-bound normative conditional should be non-empirical.

We must now say in what the non-empirical character of an ethical judgement is to consist. One could adopt here a traditional characterization of the non-empirical, or a priori, as what can be known independently of experience. But this makes the claim that world-bound judgements are non-empirical highly problematic, both because it is not altogether clear that we can arrive at *knowledge* of such judgements, and also because it is not clear what epistemic role should be assigned to moral imagination and the like. I myself would not shrink from ascribing a strong form of aprioricity to ethical judgements. But in order not to prejudge the issue, I shall take the non-empirical character of ethics to consist in its conformity to what one might call *the criterion of perceptual independence*.

Let us use the term 'inner experience' to refer to experience that the subject does not take to be a case of veridical perception (we might also add the condition that the experience actually not be a case of veridical perception); and let us use the term 'outer experience' for any other kind of experience. The *criterion* for a judgement to be non-empirical is then:

> The reasons one can have for making the judgement (or its negation) on the basis of inner experience are as good as any reasons one can have;

and the corresponding *principle of perceptual independence* for ethics is that:

> The reasons one can have for judging a world-bound normative conditional to be true (or false) on the basis of inner experience are as good as any reasons one can have.[32]

---

[31] It may not be necessary to give a complete description of a world, but merely one that is qualitatively complete. And, in this case, what we may have is not strictly speaking a conditional, but a universal claim of the form $\forall x(Fx \supset Gx)$, where F is naturalistic and G normative. I shall not concern myself with such niceties in the text.

[32] It might be argued that if every world-bound normative conditional is non-empirical, then so is any normative necessity whatever. For let $D_1, D_2, \ldots$ be an exhaustive list of all world

The principle does not commit one to having some special kind of a priori access to ethical truth, since it is perfectly compatible with inner experience providing us with an empirical basis for our ethical beliefs. Thus it allows one to steer a middle course between embracing a strong form of empiricism for ethical inquiry, on the one hand, and a strong form of apriorism on the other. Nor does the principle require one to deny that outer experience could be helpful in coming to the conditional judgements. It merely requires that, in such cases, the experience be dispensable—that whatever probative value it might have could always in principle be matched by the probative value of some inner experience. Thus someone who took himself to be a brain in a vat (and perhaps is a brain in a vat) will be in as good an epistemic position to arrive at the judgements in question as a creature firmly ensconced in the real world.

Before applying the principle, it will help to clarify it further. In the first place, it should be clear that the principle is intended to apply to an ideal cognizer who is capable of grasping a complete description of a world. This idealization is merely a device to factor out irrelevant empirical consider-ations and, in actual applications of the principle, could probably be wea-kened. For all that we require is that the antecedent of the conditional should incorporate anything that might be empirically relevant to the truth of consequent. We might be in a position, for example, to say that there is something bad about a situation involving intense suffering, whatever other circumstances might prevail.

Second, even to grasp the concepts involved in the conditional judgement might require us to have certain outer experience.[33] Thus it should be taken for granted that the cognizer has had whatever outer experience is required in order to grasp the judgements in question. This means that our reformu-lation of the principle in terms of a brain in a vat may not be altogether accurate, since we must assume that the brain in the vat (perhaps in a previous 'embodied' existence) has had whatever outer experiences are required in order to grasp the relevant concepts.

---

descriptions, and suppose A is a normative necessity. Then: (i) $\Box(D_1 \bigvee D_2 \bigvee \ldots)$; and (ii) $\Box(D_i \supset A)$ for each world description D_i. Now $\Box$ A logically follows from (i) and (ii), and so, given that (i) and (ii) are non-empirical, it is plausible to suppose that $\Box$ A is also non-empirical. However, the assumption that (i) is non-empirical is not entirely unproblematic. I am inclined to think it holds, since I take normative necessity to coincide with conceptual necessity in its application to naturalistic statements. But, on a view in which it is taken to coincide with metaphysical necessity, the assumption might be subject to doubt.

[33] I merely make this concession for the sake of argument. My own view is that the most basic ethical principles can be formulated in terms of concepts whose possession does not require any contact with the external world. Instead of talking of *human beings*, for example, which appears to require such contact, we can talk of *moral agents* or *conscious beings*, which does not.

Third, there are some reasons that are parasitic upon other reasons, in the sense that their cogency rests wholly upon the cogency of those other reasons. Testimony is a clear case of what I have in mind; since the cogency of testimonial evidence rests upon the cogency of the reasons available to the person from whom the testimony is drawn. Now it is possible in cases of this sort that the parasitic reasons may have greater probative value than the reasons upon which they depend. I may have good reason to believe some testimony, for example, even though the person providing the testimony has no good reason for saying what he does. Let us call a reason of this sort *unsustainable*. The principle should then be restricted to sustainable reasons, since it may not be possible to match an unsustainable reason, obtained on the basis of outer experience, with a reason obtained on the basis of inner experience.[34]

With these clarifications in place, I hope it is clear that the principle is indeed plausible. For how might outer experience provide reasons for forming an ethical judgement? One way is for it to inform me of the circumstances in which the moral concepts are to be applied. But this is irrelevant in the present case, since the relevant circumstances are already completely specified in the antecedent of the conditional. Another way is for it to make vivid to me *how* the concepts are to be applied in any given circumstance. It may be through seeing one person torture another, for example, that I learn to appreciate how awful torture is. But in so far as outer experience teaches me this lesson (without also informing me of the relevant circumstances), then it would appear to be irrelevant that the experience is, or is taken to be, veridical. Suppose, after what I take to be an experience of seeing one person torture another, I learn that the experience was not veridical. Does this make the moral lesson I take away from the experience any less worthy of consideration? Surely not. Surely what I learn from the experience is something that I could have learnt from a virtual form of the experience or even from a highly developed moral imagination.[35] Finally, a trusted moral authority might inform me that such-and-such an ethical judgement was correct. But in that case, given that my reason for trusting the authority is sustainable, I can simply take myself to have whatever reason the authority might have. If it is constituted by inner experience, all well and good; if it is not, then it can ('by induction') be replaced by whatever inner experience might serve in its place. It seems to me that these are essentially the only kinds of case that can arise; so, granted the adequacy of my responses, the principle is secure.

---

[34] It was an objection of Tony Martin's that made clear to me the need for a qualification of this sort.

[35] It is partly for this reason that fiction can function so effectively as a substitute for experience in the development of moral sensibility.

Let us return to the neo-naturalist. There will be for him world-bound normative conditional truths that are a posteriori. For whether a property *P* fits the good-making role is presumably a contingent (and a posteriori) matter. Indeed, if it were not, then the use of 'actually' in the formulation (*) of the position would be unnecessary, and the view would collapse into a version of the traditional form of naturalism. Consider now a world-bound normative conditional, such as: necessarily, if this is how things are, then this is good. Then, in general, to determine whether this is true will require determining what property actually satisfies the good-making role.

But the neo-naturalist now faces an intolerable dilemma. For consider his account of the good-making role. There are two possibilities: (1) it is egocentric in the sense of being indexed to the speaker; (2) it is not egocentric—that is, it is either indexed, though not simply to the speaker, or not indexed at all. The first case is illustrated by 'the property *I* would value (under ideal circumstances)', and the second two cases by 'the property *we* would value' and by 'the property *everyone* would value'. Now in any plausible version of the first option, there will be no genuine possibility of moral disagreement; for each of us, in talking about what is good, will essentially be talking about ourselves. We might attempt to secure the possibility of moral disagreement by adopting the other option. But the non-empirical character of ethical judgement will then be lost. For whether a world-bound conditional holds will in general depend upon what in fact fits the good-making role; and this, in turn, will depend upon how things are 'outside' myself. But, in that case, it is hard to see why having a window on the world (or taking myself to have such a window) would not put me in a better position to determine whether the conditional holds. Thus it appears that our neo-naturalist must either deny that there are genuine moral disagreements or must give up on the non-empirical character of ethical belief.

One might attempt to finesse this difficulty by taking the good-making role to be a question of our ideal valuational dispositions and yet taking it to be an a priori matter that we all have the same ideal dispositions. Thus it will be sufficient to ascertain my own dispositions (on the basis of inner experience) in order to ascertain them all. But the problem with this intermediate position is in seeing how it might be contingent that I have the ideal dispositions that I do and yet a priori that you have the same dispositions as me. How can my other possible self be so different from you?

There is another version of neo-naturalism that appears to avoid these difficulties. It holds, in common with the previous version, that it is an a posteriori metaphysical necessity that goodness is (or coincides with) such-and-such a naturalistic property, but it denies that any specific good-making role is part of our understanding of the term 'good'. The reference of the term, on this view, is taken to be determined 'empirically' rather than 'conceptually'. It is given by causal or other such links between our use of

the term and the real world—in much the same way, so it has been supposed, as the reference of natural kind terms, such as 'electron' or 'water'.[36]

This approach avoids the previous dilemma, since it no longer provides us with any descriptive content for the good-making role by which the dilemma might be stated. But it is still subject to a serious epistemological problem. For what is this mechanism for fixing reference, that both allows for genuine disagreement on matters of morality and yet respects the non-empirical character of moral belief, meant to be? How can the reference of the terms hook up to the real world, yet our justification for believing a substantive body of ethical truths not require any access (or any substantive access) to that world?

On this point, the much-vaunted analogy with natural kinds is of little help, and actually stands in the way of seeing what the mechanism might be. For our beliefs concerning natural kinds are not in general independent of perceptual experience. If we were to learn that most of our perceptual experience was non-veridical, then little would be left of our knowledge of natural kinds. The brain-in-the-vat is at a severe epistemic disadvantage in coming to any form of scientific knowledge; and if there really were an analogy between our understanding of scientific and of ethical terms, then one would expect him to be at an equal disadvantage in the effort to acquire moral wisdom. It is for this reason that the continuity in moral and scientific inquiry so much stressed by writers such as Boyd ([1988]; 123–4) and Railton ([1986]; 138) appears entirely misplaced. A much better analogy is with our understanding of mathematical terms, for which the idea of a hook-up with the real world is far less plausible.

I conclude that naturalism, in either its traditional or contemporary versions, is unable to rescue the doctrine that normative necessity is a species of conceptual or metaphysical necessity. There remains the possibility, of course, that normative necessity might somehow be definable in terms of another form of necessity; and two proposals along these lines may briefly be considered. One is that a normative necessity should be taken to be a normative (or moral) proposition that is true in all possible circumstances, where the circumstances are given in entirely naturalistic terms, so that there is no danger that the form of possibility by which they are qualified is moral. The error in this suggestion is that if the circumstances are taken to be naturalistic, then the idea of a moral proposition's being true *in* such a circumstance—that is, of there being a necessary connection between the circumstance and the truth of the proposition—presupposes the very notion of necessity in question.

The second proposal is that normative necessity be taken to be a form of relative necessity. We specify the moral laws without appeal to the notion of

---

[36] Sturgeon [1984], Boyd [1988], Brink [1989], and Railton [1986] adopt a view of this sort.

normative necessity, and then define a normative necessity to be whatever is entailed by the moral laws. But it is not altogether clear how we might define the moral laws without appeal to the notion of normative necessity, and the view is subject, in any case, to the difficulty that it trivializes the form of necessity possessed by the moral laws themselves.

The notion of normative necessity would therefore appear to constitute yet another basic form of necessity.[37]

6. Modal Pluralism

There remains another possibility for defining the notions of metaphysical, natural, and normative necessity. For perhaps one or other of these notions can be defined as the restriction of a more comprehensive notion of necessity. Indeed, if each could be defined as the restriction of the most comprehensive notion of necessity, then modal monism could be saved.

Perhaps the most plausible suggestion of this sort is that metaphysical necessity be defined as a restriction of the (inclusive) notion of natural necessity. In this case, there is arguably no difficulty in stating the relevant restriction in non-modal terms. For the metaphysical necessities can be taken to be those natural necessities that are essential truths. This definition is reminiscent of the earlier proposal that metaphysical necessity be defined as essentialist truth, but it does not suffer from the same difficulty over modal force, since modal force is now included in the requirement that the essentialist truth should be a natural necessity.

There is, however, a related difficulty. For we do not thereby appear to capture the *relevant* modal force. There appears to be an intuitive difference to the kind of necessity attaching to metaphysical and natural necessities (granted that some natural necessities are not metaphysical). The former is somehow 'harder' or 'stricter' than the latter.[38] If we were to suppose that a God were capable of breaking necessary connections, then it would take more of a God to break a connection that was metaphysically necessary than one that was naturally necessary. It would be harder, for example, to break the connection between the truth of P & Q and the truth of P than the connection between cause and effect. It is also because of this difference in strictness that it is so much more plausible to think of the natural necessities as already including the metaphysical necessities than it is to think of the metaphysical necessities as already including the natural necessities.

---

[37] As far as I know, Moore ([1922]; 275) was the first to suggest that there might be a distinctive form of normative necessity, in his marvellous paper 'The Conception of Intrinsic Value'.

[38] It seems to be something like this that Kripke ([1975]; 99) has in mind when he talks of necessity 'in the highest degree'.

It is difficult to say in more precise terms what this difference comes to. But one way to bring it out is in terms of the consequences of a proposition failing to be necessary. A proposition may fail to be metaphysically necessary even though it is naturally necessary. Perhaps it is a natural necessity that *e* causes *f*, though not a metaphysical necessity. Now we are inclined to think in such a case that there exists a genuine possibility of the proposition's being false. On the other hand, if a proposition were a metaphysical necessity, though not a natural necessity (in the narrow sense), then there would be no genuine possibility of its being false, since the 'hardness' of the metaphysical necessity would stand in the way.

I am inclined to think that the objections become more compelling when we consider the possibility that natural and normative or metaphysical and normative necessities might both be restrictions of a more comprehensive notion of necessity. For the character of the necessities seems even more strikingly different in these cases, and, in addition, there are difficulties in seeing how the relevant kind of restriction might be defined. It will not do to say, for example, that a normative necessity is a comprehensive necessity that essentially involves moral (or other normative) components. For under certain strange theological views, it may be a natural necessity, and hence a comprehensive necessity, that what in fact is the actual world is the best of all possible worlds. This is a comprehensive necessity that essentially involves a normative component, yet it is not naturally taken to be a normative necessity, for normative necessity is biased not towards things going well, morally speaking, but merely towards things going in the appropriate moral manner, good or bad, given how things are.

I conclude that there are three distinct sources of necessity—the identity of things, the natural order, and the normative order—and that each gives rise to its own peculiar form of necessity. Neither form of necessity can be subsumed, defined, or otherwise understood by reference to any other forms of necessity; and any other form of necessity, if my survey is complete, can be understood by reference to them. I have no a priori commitment to there being these three forms; but I must admit to finding some satisfaction in the thought that the three main areas of human inquiry—metaphysics, science, and ethics—should each give rise to their own form of necessity.

There has been a tendency in recent discussions of modality to focus on the notion of metaphysical necessity, just as earlier there had been a tendency to focus on the narrow notion of logical necessity. But it needs to be remembered that there are other forms of necessity, not intelligible in terms of these, that are equally important for philosophy, and equally worthy of study. Philosophers like to think of themselves as having found the key to the universe. But where there are many locks, it should be recognized that we may have need of many keys.

8
Tense and Reality

There is a common form of problem, to be found in many areas of philosophy, concerning the relationship between our perspective on reality and reality itself. We make statements (or form judgements) about how things are from a given standpoint or perspective. We make the statement 'it is raining' from the standpoint of the present time, for example, or the statement 'it is here' from the standpoint of where we are, or the statement 'I am glad' from the standpoint of a subject. In each of these cases, the statement has a certain 'aspect' or perspectival character in virtue of which its truth is capable of varying from one standpoint to another. Thus the statement 'it is raining' is tensed, the statement 'it is here' is 'spatiocentric' and the statement 'I am glad' is first-personal. The problem we then face is to determine whether this aspect is a feature of the reality that is described or merely a feature of the statement by which it is described. Is reality itself somehow tensed or spatiocentric or first-personal or is it merely that we describe a tenseless or spatially uncentred or impersonal reality from a tensed or spatiocentric or first-personal point of view?

My broad aims in this chapter are to get clearer on what the issue is and to make some suggestions as to how it might be resolved. The two will be intimately connected, since the suggestions I make concerning how the issue might be resolved will be largely shaped by the distinctive way in which I think the issue should be conceived. My focus will principally be on the case of tense, although I shall also devote some attention to the first-personal case and to how the different cases of aspect might or might not compare.

My central claim, in regard to the question of clarification, is that essential appeal must be made to the concept of reality in saying what the issue is. Although the concept of reality or of 'the world' is often invoked in discussions of the topic, I suspect that its use is not usually regarded as essential or even as desirable. My view, on the other hand, is that the issue cannot be properly stated without making explicit use of the concept.

Indeed, it is my view that there is not a single concept of reality to which all sides can adhere in stating their respective positions and that the issue is to a large extent about which concept of reality should be adopted. In distinguishing the different positions, we shall need to make three key

distinctions in the concept. These are between how things are and how things are in reality ('mere' versus metaphysical reality), between how things are in reality *simpliciter* and how things are in reality from a certain standpoint (absolute versus relative reality), and between reality being 'of a piece' and its being fragmented. Each of these concepts of reality will then give rise to its own characteristic species of realism.

My central claim, in regard to the question of resolution, is that there is room for a third view, between anti-realism and the standard form of realism. Realism has commonly been thought to involve a combination of two views: that reality is aspectual (tensed, first-personal, etc.) and that there is a privileged standpoint (the present, the self, etc.) from which the aspectual character of reality may be discerned. The two naturally go together, since given that there are tensed facts, then one naturally supposes that there must be a privileged standpoint, the present, from which they obtain. But it seems to me that one can hold the first of these views without holding the second. Thus it may be allowed that there are tensed facts (or the like) but denied that the present time is in any way privileged. Although the resulting view is somewhat unfamiliar, I argue that it is much better able to withstand the many objections that have been levelled against the standard forms of realism.

Non-standard realism itself comes in two different versions. Under the first, we give up the idea that reality is absolute. Reality is relative to a standpoint; and for different standpoints there will be different realities. Under the second, we give up the idea that reality is of a piece. Reality will divide into fragments, no two of which can be regarded as belonging to a single coherent whole. In what follows, I am mainly concerned to argue for the merits of the non-standard view; but I shall also provide reasons for preferring the fragmentalist version of that view to the relativist version.

The arguments, in both cases, are conditional in form. I argue that, if one is going to be a realist about tense, then one should be a non-standard realist and that, if one is going to be non-standard realist, then one should be a fragmentalist rather than a relativist. I do not directly address the question of whether one should be a realist. But there is one important respect in which the considerations of this chapter may bear on this question. For many philosophers have found realism about tense to be intuitively very plausible but have despaired of saying what the view is or even of making it coherent. Thus simply showing the view to be coherent removes what, for these philosophers, is one of the principal obstacles to believing it to be true.

The chapter is long and it may be helpful to give a general overview of its contents. It is in four main parts. The first begins by criticizing various standard formulations of the issue (§1) and then argues that the difficulties in formulations are to be resolved by introducing a distinctively metaphysical concept of reality (§2). This concept, and its variants, will play a central role in the discussion to follow.

I turn, in the second part (§§3–5), to a consideration of McTaggart's argument against the reality of time. Although this argument has been much discussed in the literature, I am of the opinion that its full force and value have not been properly appreciated and that this can only be done once considerations of reality are brought explicitly into play. It is this argument, or at least our formulation of it, that will provide us with the principal tool for classifying and investigating the different forms of aspectual realism. I first lay out a simple version of the argument (§3), then present a more sophisticated version that is closer to McTaggart's own argument (§4), and finally consider the different responses one might make to it (§5).

The third part (§§6–10) constitutes the bulk of the chapter and attempts to provide a sustained argument in favour of the non-standard position on realism that emerges from the second part. It will first be useful, for purposes of comparison, to consider the realist response in connection with other forms of aspect (§6). We then consider what I regard as the three main arguments for adopting a non-standard form of tense-theoretic realism: the argument from passage or the 'flow' of time (§7); the argument from truth and its connection with the 'facts' (§§8, 9); and the argument from special relativity and its denial of an absolute notion of simultaneity (§10).

The fourth and final part (§§11–13) discusses three topics that naturally arise from our discussion of non-standard realism. I first argue that there are reasons to be a fragmentalist rather than a relativist, to think of reality as not genuinely being 'of a piece' (§11). I then consider how a plausible form of first-personal realism might be developed, one that takes seriously our subjective perspective on the world (§12). Whether or not this position is ultimately to be adopted, it provides a way of making sense of the view that there is an 'empirical' self that stands inside the world and a 'metaphysical' self that stands outside of the world. It also provides a much more illuminating comparison with the tense-theoretic case than the more usual modal analogy. I conclude with some general remarks on the nature of the debate (§13). It is argued the debate is as much about the *concept* of reality as about the *constitution* of reality and that it is only by getting clear on what we might mean by reality that we can come to a cogent view as to whether tense, or some other form of aspect, is real.

Given the voluminous literature on the subject, it would be difficult to say anything entirely new; and much of what I write will indeed make contact with the work of others. But what I have hoped to achieve, even when I have gone over familiar ground, is the development of a systematic framework within which the issues might be discussed. Anyone familiar with the literature will be aware of its elusive character. The content of the different positions and the cogency of the arguments for them is often far from clear (the wild divergence in the interpretation and assessment of McTaggart's argument being an obvious case in point). I do not want to claim that what

was once unclear or unconvincing now becomes clear and convincing. But at least it should now be clearer where the lack of clarity or cogency may lie and what must be done if further clarity or cogency is to be achieved.

1.   The Entailment Test

I begin by considering some of the ways in which other philosophers have attempted to clarify the issue of aspectual realism. The inadequacies in their accounts will help us appreciate the need for an alternative approach.

One common way to present the issue is in terms of what is required for a complete description of reality. Suppose we provide a complete tenseless description of reality; we say what happens when, and in what order, but without any appeal or orientation towards the present time. We may then ask: is the description complete? Or is it a further fact, not implicit in the description itself, that I am *currently* sitting, for example? The realist about tense will claim that there is a further fact, while his opponent, the anti-realist, will deny this (and similarly for other cases of aspectual realism).

Let us remark—though this will not be essential to our subsequent discussion—that, even if these answers are indeed the ones that would be given, they do not fully account for the difference in the two positions. If we give a positive answer, then it is clear why we should think that reality is tensed but, if we give a negative answer, then it is not at all clear why we should think that reality is tenseless. For the claim that one can give a complete description of the world in tenseless terms does not, in itself, rule out the possibility that one can also give a complete description of the world in tensed terms. Indeed, this latter view is quite a plausible one for the anti-realist to adopt since, by his own lights, he can get at all the tenseless facts by saying what is happening at the *present* time and what is happening at any specified interval before or after the *present* time. Thus his reason for thinking that reality is tenseless is paralleled by an equally good reason for thinking that reality is not tensed. And so why does he accept the one conclusion and yet reject the other? Clearly, there is more to his position than the mere unidirectional claim of completeness.[1]

But there are more serious difficulties. I take it that reference to 'reality' or to a further 'fact'—natural as it may be—is intended to be merely incidental to the formulation; for to 'describe reality' is merely to say what is the case, and to specify a further 'fact' is merely to specify something else that is the case. So at the heart of the formulation, once it is stripped of inessentials, is

---

[1] Similarly, we might note, for other realist issues. That one can give a complete description of the world in physical terms does not in itself guarantee the truth of physicalism since it might also be possible to give a complete description of the world in psychological terms.

the following question: is there a true tensed statement, such as the statement 'I am sitting', that is not entailed by any true tenseless statements?

If we are to answer this question, then we must know what is meant by 'entailed'. There are two main possibilities. With any statement—such as 'I am sitting'—that is made at a given time t may be associated both a *content* and a *character*. The content is the specific context-sensitive information conveyed, which in the given case we might take to be the tenseless proposition that I am sitting at t. The character, on the other hand, is the context-free manner in which the content is conveyed, which in the given case we might identify with the tensed proposition that I am sitting. Thus character is independent of context while content is dependent both on character and on context.[2] Entailment can then be with respect to content, or specific information conveyed, or with respect to character. We can be asking 'is the content of the tensed statement entailed by the content of the tenseless statements?' or 'is the character of the one entailed by the character of the others?'. And presumably, what we are asking in the second case is whether the characters are such as to guarantee that the tensed statement will be true at any time at which the tenseless statements are true.

In the first case, we get the anti-realist's answer of 'yes' and, in the second, the realist's answer of 'no', since the truth-value of 'I am sitting' may vary from time to time as the truth-value of the tenseless statements remains the same. But then which of these answers should be taken as our guide to the metaphysical question? We had in mind a notion of entailment with metaphysical import: the absence of an entailment was meant to indicate that reality was tensed; and its existence was meant to indicate that reality was tenseless. But why think that one or other of these notions of entailment captures, or corresponds to, the notion we had in mind?

Indeed, it may be argued that neither notion can *generally* be taken to correspond to the intended metaphysical notion. For each will deliver a uniform verdict in the cases of interest to us. If entailment is by way of content, then it will be denied that there is a further fact in the temporal, spatiocentric, and first-personal cases while, if entailment is by way of character, then it will be allowed that there is a further fact in each of these cases. From a metaphysical viewpoint, however, we may well wish to take a differential stand on the issue. We may wish to say that there is a further fact in the tensed and first-personal cases, for example, but not in the spatiocentric case or a further fact in the tensed case but not in the first-personal or spatiocentric cases. We therefore need a further criterion as to

---

[2] If we wish, we might follow Kaplan [1989] in identifying the content with the set of possible worlds in which the proposition is true and the character with the function that takes each time into the content that the statement has at that time.

which notion of entailment should be our guide in any given case if the test is to be of any help.

Similar difficulties would appear to beset other attempts to elucidate the issue. One common approach is in terms of the distinction between relative and absolute properties. The question, it has been said, is whether the property of *sitting* is an absolute or relative feature of an individual, one that can be understood to hold *simpliciter* or only relative to a time. But when I ask whether this property is absolute or relative, then what am I talking about? Is it something on the side of Kaplanesque content, whose 'completion' is meant to give a tenseless proposition, such as *Socrates is sitting at t*, or something on the side of character, whose 'completion' is meant to give a tensed proposition, such as *Socrates is sitting*? If the former, then the property is relative; and if the latter, it is absolute. But it is not clear, in either case, why our taking the referent one way or the other should be relevant to the metaphysical question; and nor is it clear how the answer to the metaphysical question might sensibly be taken to vary from case to case.

Another common approach is in terms of indexicality. The question, it has been said, is whether such terms as 'here' or 'now' or 'I' are indexical? But what is meant by 'indexical'? It is presumably some kind of relativity to context. But the use of all these expressions is in a clear sense relative to context and there is also a clear sense in which their use—what one might call their 'disengaged' use—is not relative to context. So how does the issue of indexicality get any metaphysical bite or enable one to differentiate between the different possible cases of aspect?

Or again, one might appeal to some neutral notion of content or proposition, one that is not in itself committed to the content being either tensed or tenseless. The question, then, is whether the utterance of a tensed statement expresses a tensed or a tenseless proposition. But what is this neutral notion and why should it not be possible for the realist or the anti-realist to go either way on the question depending upon how it is understood? It seems that in so far as the realist or the anti-realist feels obliged to go his own way, it can only be because he has somehow already understood the notion with the required metaphysical import.

We appear to face a quite general difficulty. When we attempt to frame the realist issue in the usual terms—by reference to propositions or properties, say, or content and context—then it appears that either there will be no relevant difference between cases that we would like to be able to distinguish or that the terms will be understood in such a way as to presuppose the very issue in question. Thus the usual formulations appear to be inadequate; and this suggests that some fundamentally new approach to understanding the problem is called for.

2. The Reality Test

What I would like to suggest is that references to 'reality' or to 'fact' should be taken seriously in the standard formulations of these issues. It is, of course, common to use such terms as 'real' or 'fact' in the informal presentation of the issue. Thus one might ask whether tense is real or whether there are any tensed facts. But it is usually supposed that the use of these terms is incidental to the formulation and that a rigorous statement of the issues should be found elsewhere. Our view, on the contrary, is that it is only by reference to some conception of 'fact' or 'reality' that the issue can be properly understood.

There is a familiar objection to this way of thinking, which perhaps explains why it has not been pursued. The realist about tense wants to say that my currently sitting is a fact or belongs to reality, while his opponent wants to deny this. But I *am* currently sitting; so my currently sitting is a fact; and, since reality consists of all the facts, that fact belongs to reality. The metaphysical issue is thereby trivialized; it simply becomes a question of whether we are prepared to accept a tensed statement.

In order to meet this objection, we must distinguish between *mere* reality, or how things are, and *metaphysical* reality, or how things really are. Whatever is really the case (belongs to metaphysical reality) may, with some plausibility, be taken to be the case (belong to mere reality). But the converse will not in general hold; and so there is the possibility of the concept of reality doing some genuine work in the formulation of the issue. I might accept that I am sitting and even accept that it is a fact that I am sitting, for example, but not accept that this fact is constitutive of how things really are.

But what is this concept of reality?[3] I doubt that it is possible to define the concept in other terms but the general idea behind its application is that, in a representation of reality, there may be features of the representation that do not faithfully reflect what is represented. There are three principal ways in which this may happen. One is ontological; the representation might not faithfully represent *what* there is; it might depict there being nations, for example, when all there is in reality are its citizens. Another is ideological; the representation might not faithfully represent *how* things are; it might depict physical objects as having colours, for example, when in reality they have only certain primary qualities. The third is factive; the representation might represent *that* things are so, when it is not even in the business of stating how things really are; it might depict there being moral facts, for example, when all there is in reality is the expression of certain attitudes.

---

[3] See Fine [2001] for further discussion of my general views on the topic of realism.

Of these three ways in which a representation may fail to be faithful to reality, the first two are not strictly relevant to our present concerns. It is true that if we take a particular tensed fact, such as that I am sitting, then it may fail to belong to reality in either of the first two ways. It might be denied, for example, that there really are any *people* or that they really have the property of *sitting*. But these reasons for disputing the reality of the fact are incidental to the issue at hand, for our interest is in the tense or aspect of the statement rather than with its specific ontological or ideological content; we simply wish to know whether the tense or aspect of the statement might be an impediment to its faithfully representing the facts. Thus it is only the third kind of failure that is relevant—and in a very particular way.

These remarks should at least point in the direction in which we wish to understand the concept. But we shall also attempt to use the concept in a reasonably disciplined way; and, to this end, it will be convenient to suppose that we have an 'official' idiom for making reality claims. My preference, though this may not be the only option, is to take there to be a primitive sentential operator, call it $\mathscr{R}$, whose intended reading is, 'in reality, it is the case that'. Reality claims may then be formed by affixing this operator to an appropriate sentence S. Thus someone who wished to subscribe to the reality of tense might well endorse the claim that in reality I am sitting ($\mathscr{R}$S), while someone who wished to deny the reality of tense would endorse its negation ($-\mathscr{R}$S).

The more formal-minded reader can imagine that all reality claims are made within the official idiom. However, for ease of expression, it will often be helpful to speak more loosely; and we shall find it helpful, in particular, to talk in terms of a 'container' model of reality. Instead of saying 'in reality, I am sitting', we shall say that 'reality contains—or is constituted by or is composed of—the fact that I am sitting'. Such talk involves a double reification: to reality as a 'container'; and to the facts as to what is contained. But the reference to reality or to the facts as entities in themselves is, strictly speaking, inessential and might always be avoided by reverting to the official idiom.

We are now in a position to provide a very simple statement of the realist issue. Let us take for granted that we have the notion of a tensed (or some other kind of aspectual) fact. The realist issue is then the question of whether any tensed (or aspectual) facts are constitutive of reality or, more precisely, whether for any tensed (or aspectual) statement S it is constitutive of reality that S. It is of course essential here, if this formulation of the issue is to be properly understood, that the reality of the fact that I am sitting, say, should not be taken to consist in anything like the reality of the fact that I am sitting at t, where t is the time at which the assertion of reality is made. It is the reality of something intrinsically tensed that is in question.

Modest as this proposal might appear to be, it helps to bring some conceptual order to the topic. It suggests, in the first place, that the issue is

to be clarified by appeal to a distinctively metaphysical conception of reality, one that embodies a distinction between what is really the case and what is merely the case. This suggestion does not in itself exclude the possibility of further clarification, but it indicates that further clarification is to be achieved, if at all, through a better understanding of the concept of reality. And this is, indeed, the route we shall take; it is by appealing to various refinements of the concept of reality that we shall attempt to elucidate the different kinds of realist position that might be held.

The present proposal also helps us better to appreciate the defects in the previous proposals and how they might be rectified. Consider first the formulation in terms of entailment. The issue before us, under this formulation, was whether every tensed truth is entailed by all the tenseless truths. We wanted the notion of entailment to be metaphysically relevant but neither entailment with regard to character nor with regard to content seemed well suited to this purpose. We do better by invoking the concept of reality. Let the *consequences* of a class of statement be its entailments with regard either to character or to content; call a consequence of a class of true statements a consequence *for reality* if it is part of how things really are; and let the *metaphysical import* of a class of statements be the class of its consequences for reality. The relevant notion of entailment is then containment of metaphysical import; and the question of interest to us is whether some tensed truth has a metaphysical import not included in the metaphysical import of the tenseless truths. Thus we do not choose between the two kinds of entailment; we allow both and then use the concept of reality to filter out those of the consequences that are metaphysically relevant.

We might in a similar way make sense of the formulation in terms of relative and absolute properties. The question was whether *sitting* should be regarded as relative or absolute; and this, in its turn, was a question of what is required for its 'completion'. But in defining the relevant notion of completion, it will not do to appeal to the ordinary notion of fact; we must appeal to those facts that are a part of reality (or, at least, have consequences for reality). So again, it is through invoking the metaphysical concept of reality that we obtain a more adequate account. And once we see this, then it is evident that nothing is gained by adopting these other formulations and that one might just as well go for a direct formulation in terms of what is real.

There are, of course, philosophers for whom none of these explanations would be acceptable. They would reject the whole idea of a metaphysical conception of reality and of anything that might be explained with its help. This is not the place to consider the general question of how we are to make sense of realist claims, but there are two considerations that make a sceptical view seem especially implausible in the present case. In the first place, we do not require a full commitment to the metaphysical concept of reality. As I mentioned before, all that is strictly required is its application in

respect of tense or 'aspect'; and so someone who had general misgivings about the concept might well be happy with its application in this particular case.[4]

But secondly, and more significantly, it is only the most hard-nosed philosopher who would deny the intelligibility of the issue in the cases at hand. Surely, given all the tenseless facts or all the spatial facts, there is a significant metaphysical question as to whether these are all the facts that there are. But it is hard to see how one might make sense of this question without bringing in the concept of reality. For the question concerns the fit, or lack of fit, between the tensed character of our representations and the character of reality itself. But fit is not simply a matter of truth. Both sides to the debate can agree to the truth of particular tensed statements (such as the statement that I am sitting). Fit in some deeper sense is involved; and the metaphysical concept of reality simply provides a way of codifying its presence. It therefore seems that we must either accept the metaphysical concept of reality or deny the intelligibility of the issue.

3.   Simple McTaggart

This concludes our discussion, in the first part, of the proper formulation of the reality of tense. We turn in the next part (§§3–5) to a discussion of the McTaggartian argument against the reality of tense. This is also, of course, an argument against the reality of time should the reality of time be taken to require the reality of tense, but this is an aspect of McTaggart's original argument that will not concern us. Many philosophers dismiss McTaggart's argument as a mere sophism. This is not our view. I believe that the argument has a great deal of cogency; and it is through articulating the assumptions of the argument and seeing how they lead to a contradiction (§§3–4) that we are able to discern the different ways in which one might defend the reality of tense (§5).[5]

Although McTaggart's argument is to the conclusion that tense (or time) is unreal, the concept of reality plays no explicit role in the assumptions upon which the argument depends, at least, as these are usually stated. To many this would be no great surprise, since the reference to reality would simply be regarded as a rhetorical flourish. But to us, it is a serious flaw; and we shall try to be explicit as we can about how the concept of reality is being used.

Our version of the argument will rest upon four assumptions, the aim of the argument being to show that they are in conflict. I shall begin with a

---

[4] He might be a semi-quietist in the sense of Fine [2001].

[5] Other accounts of the argument include Baldwin [1999], Broad [1938], Christensen [1974], Dummett [1960], Dyke [2002], Horwich [1989], Lowe [1992], Mellor [1998], Shorter [1984], and Thomson [2001]. I have made no attempt to compare these accounts or my account with theirs.

simple version of the argument in the present section and then present a more sophisticated version in the following section, one which is closer, in certain key respects, to McTaggart's original argument.

The simple argument has four assumptions, the aim of the argument being to show that they lead to a contradiction. They are:

Realism Reality is constituted (at least, in part) by tensed facts.

Neutrality No time is privileged, the tensed facts that constitute reality are not oriented towards one time as opposed to another.

Absolutism The constitution of reality is an absolute matter, i.e. not relative to a time or other form of temporal standpoint.

Coherence Reality is not contradictory, it is not constituted by facts with incompatible content.

Neutrality, as stated, is a little vague but we shall always be concerned with its implications for what one might call the 'orientation' of reality. What this means, in the present case, is that there should be no privileged time t for which the totality of tensed facts constituting reality are ones that obtain at t. It follows, in particular, that the present time t should not be such a time; the totality of facts constituting reality should not be ones that presently obtain.

Absolutism is somewhat different from the other assumptions, since it is an assumption about how the concept of constitution that figures in those other assumptions is to be understood. If this assumption is given up—if, that is to say, the concept of constitution is taken to be relative, then the formulation of the other assumptions must be appropriately modified. Coherence, for example, will now say that reality is not constituted by incompatible facts *at* a given time (or standpoint).

For the purposes of the argument, the absolute notion of constitution that figures in the other assumptions can be taken to be either tensed or tenseless. Thus in saying that a given fact constitutes reality, one can either be speaking about the *present* constitution of reality or about its *eternal* composition. It is natural to suppose that tensed facts constitute reality in a tensed fashion and that tenseless facts constitute it in a tenseless fashion but there is no reason, in principle, why the tense-theoretic status of the fact and of the form of constitution should not come apart.

It is important, if the assumptions of the argument are to have their intended import, that the notion of constitution be properly understood. Suppose, for example, that someone were to take reality to be constituted by a tensed fact f (say, the fact that I am sitting) just in case, for some time t, it is constituted by the fact f-at-t (the fact that I am sitting at t). The assumption of Realism would then hold but not in its intended sense, since reality's being constituted by a tensed fact would amount to no more than its being constituted by a corresponding tenseless fact. Or again, suppose someone

were to understand how reality might be constituted in a relative manner by taking reality to be constituted at time t by the tensed fact f (say, the fact that I am sitting) just in case it is constituted by the fact f-at-t (the fact that I am sitting at t). The assumption of Absoluteness would then fail but not in the intended way, since the underlying notion of constitution, in terms of which the relative notion was understood, is itself absolute.

   We may avoid difficulties of this sort by requiring that the relevant notion of constitution be *basic*, i.e. that it should not be one that is understood in terms of some more basic notion of constitution. The deviant forms of realism or relativism or the like will not then arise. The derivation of a contradiction can now be given. It follows from Realism that reality is constituted by some tensed fact. There will therefore be some time t at which this fact obtains. Now Neutrality states that reality is not oriented towards one time as opposed to another. So reality will presumably be constituted by similar sorts of tensed facts that obtain at other times (given that there are other times!). We wish to show that it then follows that reality will be constituted by incompatible facts. Now there is no *logical* guarantee that the facts constituting reality that obtain at t will be incompatible with the facts constituting reality that obtain at other times, since reality might be so boring that the same tensed facts hold at every single time. However, any reasonable view of how temporal reality might be constituted should allow for its being reasonably variegated over time; and presumably it will be then be constituted by incompatible facts, i.e. by facts with incompatible contents. If, for example, it allows for the present fact that I am sitting, then it should also allow for the subsequent fact that I am standing. By Absolutism reality is absolutely constituted by these facts; and this is then contrary to Coherence.

4. Sophisticated McTaggart

If the realist admits that there is a basic notion of constitution, then he should be willing to assert the assumption of Realism for that notion; and the other three assumptions are also reasonable. The previous simple version of the argument will therefore gain a foothold against the realist's position. But what if the realist is unwilling to admit that there is a basic notion of constitution? He is clear that there is a notion of constitution for which he wishes to assert Realism; and he may admit that there are other notions of constitution. But he is unwilling to make judgements as to which of these notions is most basic. It is not then so clear that there should be a notion of constitution for which all four assumptions hold since, as we have seen, there are various derived notions of constitution for which various of the assumptions do not hold.

   The more sophisticated version of the argument is designed to get round this dialectical difficulty. Instead of insisting that the notion of constitution that he uses should be basic, it insists that it should meet certain explanatory

demands. The argument uses four assumptions that are analogues of the original assumptions. However, in stating these assumptions we shall use the term 'composition' in place of 'constitution' to signal that the relevant notion may not be basic.

The analagous assumptions (I've kept the old labels) are:

Realism Reality is composed of tensed facts.

Neutrality No time is privileged, the facts that compose reality are not oriented towards one time as opposed to another.

Absolutism The composition of reality is not irreducibly relative, i.e. its relative composition by the facts must be explained in terms of its absolute composition by the facts.

Coherence Reality is not irreducibly incoherent, i.e. its composition by incompatible facts must be explained in terms of its composition by compatible facts.

The dialectical force of Absolutism is this. Suppose that the realist asserts that reality is composed of different facts at different times. Then he must explain how this is possible in terms of the absolute composition of reality. In other words, he must provide an explanation of relative composition in terms of absolute composition which then accounts for how reality might be composed, in the way that it is, by different facts at different times. Similarly for Coherence. Suppose the realist asserts that reality is composed of incompatible facts. Then he must explain how this is possible in terms of a coherent notion of composition, one that does not allow incompatible facts. The apparent incompatibility must disappear on a deeper view of how reality is composed.

Thus in defending his original claim of Realism, the realist may be forced to make use of other notions of composition, ones that may be absolute or coherent when the original notion is relative or incoherent. The assumptions of Neutrality, Absolutism, and Coherence are also meant to apply to these other notions of composition. If, for example, the realist uses an absolute though incoherent notion of composition in defending a relative notion, then he must show how the resulting incoherence can be 'removed'.

The Realism assumption can be taken to apply just to the realist's original notion of composition. But we must then impose a further requirement on the explanations of composition that might result from the other assumptions. I call this the requirement of 'No Collapse'. It states that, in explaining one notion of composition in terms of another, the realist's position should not collapse into an anti-realist position. In other words, reality's being composed of certain tensed facts should not be taken simply to be a matter, according to the explanation, of its being composed of tenseless facts. Suppose, for example, that the realist were to provide the following explanation of relative composition; for reality to be composed of a tensed fact f at a time t is for it to

be composed of the fact f-at-t. This would then be clearly in violation of No Collapse; and it is also clear, when there is a violation, that the realist's position is realist in name only. Given No Collapse, it will follow that any of the subsequent notions of composition used by the realist should also conform to Realism since, if they did not, then collapse would be unavoidable.

The argument from these new assumptions can now be stated. Suppose the realist asserts his position using some notion of composition. It can be tensed or tenseless, relative or absolute, coherent or incoherent. However, we know from the original argument that it cannot conform to all four assumptions. Since it is required to conform to Neutrality, it must either be relative or incoherent. Assume that it is relative (the argument being similar in the other case). He is then required by Absolutism to account for the relativity in terms of an absolute notion of composition. This absolute notion will conform to Realism, by No Collapse, and also to Neutrality; and so by the original argument, it must be in violation of Coherence. The realist is therefore required by Coherence to account for the incoherence in terms of a coherent notion of composition. As before, this coherent notion will conform to Realism, by No Collapse, and also to Neutrality; and thus, by the original argument, it must be in violation of Absolutism. And so the argument will continue.

Thus any purported explanation of the relativity or incoherence will result in an infinite regress in which the relativity or incoherence constantly re-appears. But this means that no purported explanation of the relativity or incoherence can succeed since, in any such explanation, we will ultimately have to appeal to the very feature that we were trying to explain away. Thus Absolutism and Coherence cannot be satisfied, compatibly with the other assumptions, after all.

An analogy may make the point clear. Suppose someone were both a physicalist and a nominalist; he thought that the mental could be explained in terms of the physical and the abstract in terms of the concrete. But suppose now that any explanation of the mental in terms of the physical required the use of the abstract and that any explanation of the abstract in terms of the concrete required the use of the mental. His position would not then be sustainable. There would only be the appearance of explaining the mental in terms of the physical or the abstract in terms of the concrete. For since the physical presupposed the abstract and the abstract presupposed the mental, the purported explanation of the mental in terms of the physical would be circular; and similarly for the purported explanation of the abstract in terms of the concrete.[6]

---

[6] One might also modify Neutrality in the same way in which we have modified Absoluteness and Coherence and require that any bias in the composition of reality should have an explanation in neutral terms. A somewhat similar form of the argument could then be made to go through.

Our two arguments are clearly McTaggartian in spirit. However, they differ in certain crucial ways from McTaggart's own version of the argument (in [1908] and [1927]). It will be recalled that McTaggart thinks that the realist about tense is required to hold that any given event is past, present, and future. Our own construal of the incompatibility is more abstract: we do not presuppose an ontology of events; and nor do we suppose that the incompatibility lies in the determination of something as past, present, and future. Indeed, for our purposes, the simple example of my sitting and my standing, without any explicit reference to events as the subject of the statement or to tenses as their predicate, is sufficient to make the point.

More significantly, our arguments do not begin by supposing that there is a prima facie contradiction in the realist's position from which he must somehow extricate himself. Many commentators have questioned whether there is a prima facie difficulty here at all. It is as if one were to tell a free man that he was imprisoned. It would then *look* as if there were no possibility of escape, since there is no relevant change he could make to his condition. But, of course, the correct conclusion for him to draw is that he was not imprisoned in the first place! Our argument, by contrast, attempts to *demonstrate* a contradiction. It is therefore not to the point to show that the contradiction is only apparent. All one can properly do, by way of response, is to impugn the reasoning by which the contradiction is derived or challenge one of the assumptions upon which it rests.

The reasoning in the second version of our argument corresponds in a loose way to McTaggart's. For McTaggart has his protagonist attempting to evade the prima facie contradiction by relativizing the claims he makes, which then results, once he adopts a neutral standpoint, in his having to accept further seemingly contradictory claims. Our realist is forced to oscillate in a similar way between an unacceptable form of relativity and an unacceptable form of incoherence. But whereas there is some question as to whether the resulting regress is vicious in the case of McTaggart's argument, there is no real doubt in the case of our own argument. For our regress is a regress in explanation, which exposes the circularity that must exist in any proposed explanation of the relativity or incoherence.[7]

Finally, we might note that it is crucial to the formulation of the premises of our argument that explicit appeal be made to a metaphysical concept of reality. Suppose, for example, that we were to drop the reference to the concept of reality in the formulation of Realism and Neutrality. Realism would then become the claim that there are tensed facts, which is not something that can be sensibly denied, while Neutrality would become the

---

[7] One possible difference in the arguments is that we make no appeal to embedded tense. However, the role of embedded tenses in McTaggart's formulation of the argument has been a matter of dispute. See Taylor [1997] for a discussion.

claim that the tensed facts are not oriented towards a particular time, which is not something that can be sensibly affirmed. It is through using the metaphysical concept of reality that we can convert these trivial truths and falsehoods into something with genuine metaphysical bite. And I suspect that it has been the failure to recognize a distinctive metaphysical concept of reality of this sort that has primarily stood in the way of finding a satisfactory formulation of McTaggart's argument.

5.   Responses to the Argument

I can think of only one objection with any degree of plausibility to the reasoning of the argument. I claimed that if reality was composed of a tensed fact that obtained at one time and was also composed of tensed facts that obtained at other times, then it was plausible that some of these facts would be incompatible with one another. But one might adopt a Broad (in fact, narrow!) view concerning the constitution of reality: the only facts constituting reality concern what is-or-has-been (cf. Broad [1959], §I.2). These facts will then grow over time (thus if I am-or-have-been sitting then it will always be true that I am-or-have-been sitting); and so any two of them will be compatible.

But it is not clear that this will work, even if we go along with the underlying metaphysics. For we may want reality at any given time to be constituted not merely by the particular facts that are-or-have-been but by there being no other particular facts that are-or-have-been; and the absence of any further particular facts at one time will then be incompatible with their presence at another time. Thus the fact that there are-or-have-been no other particular facts is not one that will continue to obtain.[8] I might note that there is a further problem should time continue indefinitely into the future, since the compatibility of *all* of the particular facts (not just two of them) would then require the existence of a time beyond all time.

We therefore appear justified in accepting the reasoning of the argument and the only question is which of the four assumptions should be rejected. The original intent behind the argument was that Realism should be given up but, now that the other assumptions have been brought into play, we may consider whether one of these might reasonably be rejected in its place. There are three possibilities in all.

The first (standard realism) is to retain Realism but reject Neutrality. It will be maintained that there is a privileged time, one to which the facts comprehending reality are oriented; and this privileged time will, of course, be the present. Thus on this view, there is an absolute notion of constitution,

---

[8] This difficulty is not considered by Broad ([1959], 79–84) in his discussion of the matter.

but it is tensed; and the tensed facts that constitute reality are those that presently obtain.

Some anti-realists would object to the idea of a tensed constitution and it will be worth considering what their reasons might be before continuing with our review. To fix our ideas, let us suppose that the realist takes the fact that I am sitting to be a fact that currently constitutes reality (any other example would do). Let us use '*f*' for the fact that I am sitting and '*f*-at-t' for the fact that I am sitting at t (or some corresponding tenseless fact). The anti-realist objection can then be seen to rest on the following two equivalences:

> *First Equivalence*: that *f* currently constitutes reality is equivalent to the present time t being such that *f* constitutes reality at t.
> *Second Equivalence*: that *f* constitutes reality at *t* is equivalent to *f*-at-*t* tenselessly constituting reality.

From these two equivalences, we may infer:

> *Concluding Equivalence*: that *f* currently constitutes reality is equivalent to the present time t being such that *f*-at-*t* tenselessly constitutes reality.

There are then two ways in which the Concluding Equivalence might be used to mount an objection to the realist's position. Under the first, it is claimed that the realist will reject the right-hand side of the equivalence. He is therefore obliged to reject the left-hand side as well, thereby contradicting his own position. (The objection in this form only requires the left-to-right implications of each equivalence under a material reading of the conditional.) Under the second, it is claimed that, given the Concluding Equivalence, the truth of its left-hand side can amount to no more than the truth of its right-hand side and so, in asserting the left-hand side, the realist will not have staked out a distinctive position. His view will simply have collapsed into that of his opponent. (The objection in this form only requires the right-to-left implications of each equivalence, under a reading of 'implies' as 'at least amounting to'.)

The first equivalence is beyond reproach (under either reading). The second equivalence might be justified as follows. That a fact constitutes reality is not, properly speaking, a relative matter; it is not something that holds relative to a time (or relative to a 'standpoint' of some other sort). We must therefore make sense of the relative constitution of reality posited on the left of the equivalence in terms of an absolute conception of reality. Yet what could *f*'s constituting reality at t amount to unless it is that some tenseless counterpart of *f*, such as *f*-at-t, should constitute reality?

Let it be granted that some explanation of relative constitution in terms of absolute constitution is called for. It is then critical to the defence of the realist's position that he come up with some alternative account of relative constitution. But this can be done. He may explain relative constitution in

terms of his own favoured absolute notion of tensed constitution by taking *f* to constitute reality at t if *f* currently constitutes reality *when* t is present. Thus if L is the tense-logical operator 'always', then *f* will be taken to constitute reality at t if L(t is present ⊃ *f* constitutes reality), where the 'constitutes' on the right is absolute and tensed. Both sides of the equivalence will then be questionable for the realist. For he can accept that *f* constitutes reality at t, i.e. that *f* constitutes reality whenever t is present, without thereby accepting that *f*-at-t constitutes reality; and in accepting that *f*-at-t constitutes reality, he is not thereby committed to *f*'s constituting reality at t in the intended sense, i.e. to *f*'s constituting reality whenever t is present.

This line of defence is perhaps even more convincing in the modal case. Let *f* now be the fact that snow is white. The modal analogue of the second equivalence takes the following form:

> That *f* (the fact that snow is white) constitutes reality at the world w is equivalent to *f*-at-w (the fact that snow is white at w) constituting reality.

But the left- and right-hand sides of the equivalence will have a completely different status for the actualist. The left-hand side concerns the *hypothetical* constitution of reality. Reality might be different from what it is; and in considering the constitution of reality at a world, the actualist is considering how reality might have been had that world obtained. The right-hand side, on the other hand, concerns the *actual* or *categorical* constitution of reality. It is to the effect that reality is actually or categorically constituted by certain world-relative facts. And similarly for the realist about tense. That *f* constitutes reality at a given time is a hypothetical claim and that *f*-at-t constitutes reality is a categorical claim; and there is no difficulty in seeing how we might have the one without the other.

Let us continue with our review. The current response has been the standard realist response to the argument and I suspect that many philosophers may have thought that it constituted the only reasonable form of realism about tense (I myself was once of their number!). The thought—and it is a very natural one—is that as realists about tense we should think of tensed facts as obtaining absolutely; and this then serves to distinguish the current temporal standpoint as the one from which the facts that absolutely obtain will obtain. But our version of the McTaggart argument opens up two other ways in which the Realist assumption might be preserved.[9]

The first (external relativism) is to retain Neutrality but reject Absolutism. It will be maintained that the constitution of reality is relative. Reality is

---

[9] Percival ([1991], 94–5) has a distinction between 'dynamism' and 'realism' about contexts, which corresponds loosely to our distinction between standard and non-standard realism, but he takes the two positions, at bottom, to be the same.

indeed composed of the tensed fact that I am sitting and also the tensed fact that I am standing, but it is composed of these facts through being constituted by them at different times. Of course, both the anti-realist and the standard realist can admit a sense in which the constitution of reality might be relative. The anti-realist can say that for reality to be constituted by a tensed fact f at time t is for it to be constituted by the fact f-at-t, while the standard realist can say that for reality to be constituted by a tensed fact f at t is for it to be constituted by f whenever t is present. However, in both of these cases, the relative notion of constitution is explained in terms of an absolute notion. The neutral realist not only rejects these explanations; he rejects any demand to explain the relative notion of constitution in terms of an absolute notion. For him, reality is *irreducibly* relative.[10]

Since the anti-realist, the standard realist, and the current non-standard realist can all meaningfully take reality to be relative to a time, it can sometimes be difficult to keep the views apart. But each of them is operating with a very different conception of what the relative reality is and how it relates to the time to which it is relative. For the anti-realist, reality at a time is what one might call a *facet* of reality; and what properly belongs to reality is not the facet itself but the fact that it is instantiated at the given time. For the standard realist, reality at a time (other than the present) is a *hypothetical* reality; what properly belongs to reality is not the hypothetical facts constituting this reality but the fact that they would be the facts were this reality to obtain. For the non-standard realist, by contrast, reality at another time is an *alternative* reality. It is neither a facet of the one true reality nor a hypothetical determination of the one true reality, but another reality on an equal footing with the current reality; and the facts belonging to such a reality are full-fledged facts, sharing neither in the incomplete status of a facet nor in the insubstantial character of a hypothetical fact.

It is also crucial to a proper understanding of this position that one sharply distinguish between what one might call the 'internal' and 'external' forms of relativity. There is a sense for the anti-realist in which tense is relative. Tensed facts constitute reality only in so far as they assume a relative form; and what properly belongs to reality is not the fact that I am sitting but the fact that I am sitting at a given time (or some other fact of this sort). For the

---

[10] A view of this sort has occasionally been aired in the literature. Thus Dummett [1960] considers rejecting the idea that 'the description of what is really there, as it really is, must be independent of any particular point of view', and Horwich [1989] (in § 5 of 'The Moving Now') considers the possibility that 'the facts "e is past", and so on, might themselves obtain only relative to a temporal perspective'. Some superficially similar views are expressed in Bigelow [1991], Schlesinger [1991, 1993, 1995], and Tooley [1997]. But it is hard to see any of these authors as subscribing to a form of external relativism as it is understood here. They appear either to be standard realists with an odd modal interpretation of the tenses or anti-realists with an odd view as to how tensed features might be relativized.

neutral realist, by contrast, tense is not relative in this sense, but absolute. The tensed facts themselves belong to reality, and do not get to belong to reality through being relativized to a time. However, their *belonging* to reality is a relative matter; and so the tensed content of reality can be taken to vary from one moment to another even if the tensed contents are not themselves relativized. For the external relativist, there are different realities at different times and there is no saying how reality is without presupposing a temporal standpoint from which the description is given. For the internal relativist, on the other hand, there is a single reality, which does not vary over time, but in saying how this reality is we must say how things are at each time, thereby relativizing the descriptions themselves to one time as opposed to another.

Philosophers have often debated whether tensed features are absolute or relative and neither view has seemed completely satisfactory. But we can say both! For we countenance two kinds of relativity, one internal to the facts and the other external to the facts; and, in regard to the first, tensed features will be absolute while, in regard to the second, they will be relative. We might say that they are *relatively absolute*, thereby doing justice to both sides of the debate.

When philosophers have considered a view of this sort, they have tended to think of it as collapsing into triviality or some version of the anti-realist position. But this is because they have failed to heed the distinction between the internal and external forms of relativism. For them, it is as if the realist has wanted to assert that each time or time-slice of the world is present but subject to the qualification that this should only hold at the time in question; and put this way, it is hard to see how it might amount to anything more than a triviality to which even the anti-realist could agree. But the proper formulation of the intended claim is that reality is constituted, at each time t, by the fact that t is present.[11] This is quite different. The facts that belong to reality are genuinely tensed though their belonging to reality is a relative matter whereas, for the anti-realist, the facts that belong to reality are tenseless while their belonging is an absolute matter.

The final response to the argument (fragmentalism) is to retain Neutrality and Absolutism but reject Coherence. One naturally assumes that in a correct account of reality all apparent contradictions will be ironed out. If something is both hot and cold, it must be because one part is hot and the other cold, or because it is hot and cold at different times, or because being hot is somehow compatible with being cold. But on the present view, this fundamental assumption is given up. It is taken to lie in the character of

---

[11] For the purposes of the present discussion, we should ignore any doubts that the realist might have about times or time-slices being among the basic constituents of reality.

reality that certain apparently contradictory aspects of it cannot be explained away. Reality may be *irredeemably* incoherent.[12]

Under such a view, reality will be fragmentary. Certain of the facts constituting reality will 'cohere' and some will not. Any fact is plausibly taken to belong to a 'fragment' or maximally coherent collection of facts; and so reality will divide up into a number of different but possibly over-lapping fragments.

These fragments will correspond to the external standpoints of the relativist. But their status is quite different; for the fragmentalist will understand the standpoints of the relativist in terms of the fragments rather than the other way round. Thus for a fact to belong to reality *at* a standpoint, according to him, is for the standpoint to be a fragment of reality to which the fact belongs. Similarly, the relativist may acknowledge the über-reality of the fragmentalist. But for him, it will be explained in terms of the relative realities rather than the other way round. For a fact to belong to the über-reality, according to him, is for it to be a fact that belongs to a reality at a time. Thus both are concerned, in their own ways, to deny the existence of a single coherent reality. But the relativist denies that it is single, while the fragmentalist denies that it is coherent.

One might naturally explain the fragmentalist's notion of coherence in terms of conjunction; for two facts that belong to reality to cohere is for their conjunction to belong to reality. But natural as this explanation may be, there is reason to think that it may not be fundamental. For one thing, one might want to explain the obtaining of a conjunctive fact in terms of the obtaining of its conjuncts. Thus conjunctive facts will disappear from reality on this view in favour of their conjuncts. But also, the obtaining of a conjunctive fact may presuppose substantive questions of identity. In the first-personal case, for example, the conjunctive fact of feeling and seeing something will obtain when the *same person* feels and sees something; and one might want to explain in more basic terms what it is for the same person to be involved in these two experiences rather than just building it into the 'metaphysics' of conjunction. For this reason, one might want to take the notion of coherence as fundamental in addition to the notion of reality.[13] One would then expect there to be various substantive 'rules of coherence' concerning the conditions under which a set of facts would be coherent and the way in which the coherence of one set of facts might constrain the coherence of another set. For example, in the classic tense-logical case, one would want that if *f* were to cohere (i.e. to be simultaneous) with *g* and *g*

---

[12] It is this view, or something akin to it, that is dismissed in Moore ([1997], 48–50).

[13] Coherence is a relation that will hold among one or more facts. Given a primitive relation of coherence, one might take a fact to belong to reality when it is self-coherent. This gives a coherence theory of *fact* or *truth*, though not in the usual sense!

were to cohere with *h* then *f* would cohere with *h* (though, within a relativistic setting, this rule would have to be dropped).

It is very tempting to want to explain away the contradiction that the fragmentalist claims to find in the facts. The facts themselves, one wants to say, cannot be incompatible; and so lying behind the 'facts' that the fragmentalist takes to be incompatible must be facts that are compatible and that are what are really in question. Thus it is not the fact that I am sitting and the fact that I am standing that belong to reality but the fact that I am sitting at t and the fact that I am standing at t′; and if not that, then something else of that same sort. But the fragmentalist, like all realists about tense, is animated by a robust sense of the inviolably tensed character of the facts. In relativizing them, they are destroyed. Combine this sense of what the facts are like with an egalitarian and undifferentiated view of what they are and her position is forced upon one.

Although there is a sense in which the fragmentalist takes reality to be contradictory, her position should not be seen as an invitation to accept contradictions. Even if reality contains both the fact that I am sitting and the fact that I am standing, it will not be correct for me simultaneously to assert both that I am sitting and that I am standing. For any such assertion will only relate to those aspects of reality that 'cohere' with the existence of the given assertion; and so, it will only be correct for me to assert that I am sitting if, at the time of the assertion, I am sitting. And nor should her position be seen as a general invitation to accept other, more radical, forms of pluralism. To establish an acceptable form of fragmentalism (or of external relativism) one must show both that our judgements are relative in the relevant respect and that what most plausibly accounts for the relativity in the judgement is a plurality or relativity in the reality with which it deals. In the cases of interest to us, the relativity in the judgements is clear and what is not so clear is what might account for it. But in the more radical cases—in which it has been supposed that there might be different cultural or social or conceptual realities, for example—it is not even clear that we have the required relativity in judgement.

Each of the non-standard positions is committed, in its own way, to a distinction between a single comprehensive über-reality and a plurality of more particular realities. However, it is very hard to say what this distinction comes to; and there is a constant temptation to try to understand it in more intelligible, yet ultimately inappropriate, terms. Thus one might think of the particular realities as providing a 'perspective' on the über reality.[14] But this is doubly inappropriate. For there need be nothing 'behind' the facts of which they are a perspective; the facts come as they are and there are no underlying facts of which they might provide a different 'take' or version.

---

[14] Dummett [1978] and Horwich [1989] slip into this way of talking.

And there need not be anything in 'front of' the facts that has the perspective. The division of über-reality into particular realities is indifferent to the presence of an observer; and in so far as different observers have a different perspective on reality at different times, it is because of a difference in the reality upon which they have a perspective, not the other way round. Or again, one might think of über-reality as a manifold of possible or potential realities. But there is no possibility or potentiality without actuality; and so, on this view, one of the realities is distinguished as actual, whereas the view is that all are equally real.

One might say that über-reality 'manifests itself' in the form of the particular realities, that it becomes 'alive' or 'vivid' through the particular realities obtaining. Each particular reality presents itself as the whole of reality. It creates the illusion, if you like, that there are no further facts, even though there are many such realities and each is equally real. But it should be acknowledged that these remarks merely gesture in the direction of a certain idea and that, if we have here a viable conception of a pluralistic universe, then none of the usual models for making sense of it will apply.[15]

We conclude with a flow-chart (Fig. 8.1) representing the different positions and the way in which they arose: The beauty and power of
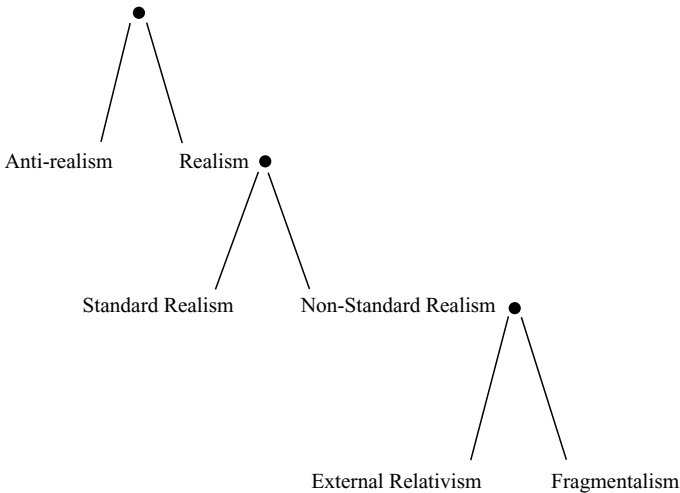
[15] We should distinguish the equal reality of the particular realities from the equal reality of Lewisian worlds (Lewis [1986]). The facts from within one world, for him, are perfectly compatible with the facts from within any other world. The Lewisian semantics is misleading in this regard, since it may be taken to deliver the verdict that 'Nixon lied' is true in the actual world w and yet false in some possible world w′. But what makes the sentence true in w is that the actual referent N of 'Nixon' lied and what makes it false in w′ is that some counterpart N′ of

McTaggart's argument, as I conceive it, is that it forces us to adopt one of these positions. It does not categorically establish that tense is unreal. But it does show that belief in the reality of tense requires either that one privilege the present or that one takes reality to be relative or that one conceives of it as fragmentary.

6.   Four Types of Realism

This concludes our discussion of McTaggart's argument and its philosophical significance. We turn now, in the third part of the chapter, to the defence of realism. Although standard realism has been the most commonly accepted form of realism, I do not believe it to be the most plausible; and, in the following three sections, I want to consider some of the reasons for favouring a non-standard form of realism.

But first it will be helpful to compare the issue of realism in regard to time with some analogous cases. What we have, in the temporal case, is a certain aspectual feature of judgements, their 'tense'; and an associated form of relativity. The question then arises as to the relationship between these judgements and the reality they describe. Is the tense of these judgements also a feature of the reality? Or is it only the relativized form of these judgements that can properly be said to describe how things are? A similar question arises in other cases in which there is an 'aspectual' feature of judgements and an associated form of relativity. Three main cases of this sort have been considered: the modal, the spatial, and the first-personal. In the modal case, the aspect is that of being contingent (or 'worldly') and the relativity is to a world; in the spatial case, the aspect is that of being spatiocentric (as with 'here' or 'there') and the relativity is to a location; and in the first-personal case, the aspect is that of being first-personal (as with 'I' or 'you') and the relativity is to a person.[16]

In each of these other cases we can construct an analogue of our McTaggart-style argument and a similar range of responses can then be distinguished. In the modal case, for example, the assumption of realism takes the form:

N did not lie; and there is no incompatibility in the facts (N's lying, N''s not lying) that constitute those two worlds. We should also distinguish our view from Goodman's [1978] despite some superficial similarities in formulation. Goodman's pluralism is animated by a deep hostility to the intelligibility of talking about how the world is in itself; ours, on the other hand, is animated by a desire to understand how the world is in itself. Thus for us the question is what kind of Goodmanian 'world-versions' might remain once we exclude those that are not faithful to how the world is in itself.

[16] These are not, of course, the only cases; and I might note that it was questions of ontological relativity, and not of indexicality as it is usually conceived, that prompted the present line of investigation.

> *Worldly Realism* Reality is composed of worldly facts,

and the principle of neutrality takes the form:

> *Worldly Neutrality* No possible world is privileged, i.e. the facts that
> compose or constitute reality are not oriented towards one possible
> world as opposed to another.

The standard realist will claim that there is a privileged world, namely the
actual world, while the non-standard realist will treat all worlds on an
ontological par (but still hold to the reality of worldly facts). Similarly, in
the first-personal case, the assumption of realism takes the form:

> *First-Personal Realism* Reality is composed of first-personal facts,

while the principle of neutrality takes the form:

> *First-Personal Neutrality* No person is privileged, i.e. the facts that
> compose or constitute reality are not oriented towards one person
> as opposed to another.

The standard realist will claim that there is a privileged person, namely
himself, while the non-standard realist will treat all people on a par (but
still hold to the reality of first-personal facts).

   If we ask in the modal case whether we should be a non-standard realist
(and adopt the principle of neutrality), then the answer has seemed to most
philosophers to be a clear 'No'. It has seemed evident that, of all the possible
worlds, the actual world is privileged; it is the standpoint of reality, as it
were, and the facts that constitute reality are those that obtain in this world.
On the other hand, if we ask, in the first-personal case whether we should be
a non-standard realist (given that we are going to be first-personal realists in
the first place), then the answer to most philosophers has seemed to be a
clear 'Yes'. It has seemed metaphysically preposterous that, of all the people
there are, I am somehow privileged—that my standpoint is *the* standpoint of
reality and that no one else can properly be regarded as a source of first-
personal facts.

   The case of time is perplexing in a way that these other cases are not. On
the one hand, there is not the same wide metaphysical gulf between the
present and other times as there is between the actual world and other
possible worlds. What goes on in the present and at other times is somehow
part of the same all-encompassing reality in a way in which what goes on in
the actual world and in other possible worlds is not. On the other hand,
there is not the same metaphysical equality between the present and the past
or future as there is between different minds. What goes on in the future, or
in the past, does not seem real to the same extent or in the same way as what
goes on in the present. Thus the past and future appear to have some kind of

intermediate status—neither clearly existing outside reality, like other possible worlds, nor clearly being part of reality, like other minds.

The dialectical situation in the case of time and in the other two cases is therefore quite different. There has been a tendency to treat the cases of time and modality in an analogous manner, taking what is plausible in the one case to be equally plausible in the other. One extreme consequence of this tendency has been the willingness, on the part of some, to embrace an anti-realist view about the worldly (as in Lewis's form of modal realism) given the plausibility of the corresponding anti-realist view on tense. But there may have been some unfortunate consequences in the other direction as well. For neutral realism about the worldly is as implausible as the anti-realist view; and the implausibility of the position in the modal case may have led philosophers, sympathetic to a realist view, to overlook its plausibility in the tense-logical case. If philosophers had only taken the case of first-personal case realism more seriously as a model, they might have been spared this particular form of metaphysical myopia.

Despite the lack of any clear intuitive evidence in the tense-theoretic case, I believe that there are at least three sets of considerations that tell against the standard view and in favour of some sort of neutral realism; and so let me discuss each of these in the following three sections.

7. The Argument from Passage

The first set of considerations arises from the original motivation for the realist view. There has been thought to be a significant metaphysical difference between space and time. Time flows; there is something that one might call the passage of time, the movement of time from one moment to the next, which has no counterpart in the case of space.

Now it is very difficult to say more clearly what this distinctive feature of time is. But a thought commonly had by realists is that it essentially consists in the successive possession of the property of being PRESENT or NOW; for time to pass from one moment to the next is for a property of presentness to pass from one moment to the next. Of course, this property of being present cannot be the relativized property of being present at the very time in question if the proposed explanation is to do any work; and so it is only by supposing that there is an absolute property of presentness, genuinely had by things, that realists can adequately account in this way for the passage of time.

But although the standard realist can grant that there is such a property, his metaphysics makes it entirely unsuited to accounting for the passage of time. For what can he say about the present that might bear upon the question? He can maintain that something is present, not merely in the sense in which the anti-realist might concede, but also in the sense in

which this fact, or facts from which it follows, are themselves constitutive of reality. Thus presentness for him is a genuinely absolute and objective feature of things. He can also maintain that the present time t (and any other present thing) is present.

However, this will not serve his purpose; for the passage of time requires that the moments of time be *successively* present and this appears to require more than the presentness of a single moment of time.[17] The realist at this point might appeal to the fact that any particular future time t^+ *will be* present and that any particular past time t^- *was* present. However, the future presentness of t^+ amounts to no more than t being present and t^+ being later than t and, similarly, the past presentness of t^- amounts to no more than t being present and t^- being earlier than t. But then how can the passage of time be seen to rest on the fact that a given time is present and that various other times are either earlier or later than that time?

The point can be strengthened. For the fact that time flows is a tenseless fact about time; it is not one that holds at one time rather than another. But the proposed explanation is tensed; it states of each of a number of times that it is *present* or was *past* or will be *future*. But if we try to convert it into a tenseless explanation, we end up with a triviality. We must say something like: it is always the case that some time is present, that all earlier times were present and all later times will be present. And this is something that even the anti-realist can accept.

The standard realist faces a general difficulty. For suppose we ask: given a complete tenseless description of reality, then what does he need to add to the description to render it complete by his own lights? The answer is that he need add nothing beyond the fact that a given time t is present, since everything else of tense-theoretic interest will follow from this fact and the tenseless facts. But then how could this solitary 'dynamic' fact, in addition to the static facts that the anti-realist is willing to accept, be sufficient to account for the passage of time? We naturally read more into the realist's tense-logical pronouncements than they actually convey. But his conception of temporal reality, once it is seen for what it is, is as static or block-like as the anti-realist's, the only difference lying in the fact that his block has a privileged centre. Even if presentness is allowed to shed its light upon the world, there is nothing in his metaphysics to prevent that light being 'frozen' on a particular moment of time.

The two forms of non-standard realism are not subject to these difficulties since they do not single out any one time as *the* present. For the external relativist, each time is objectively present at that time: at each time t, reality

---

[17] As Gödel [1949] puts it, 'the existence of an objective lapse of time, however, means (or, at least, is equivalent to the fact) that reality consists of an infinity of layers of the "now" which come into existence successively'.

is constituted by the absolute fact that t is present (or it is constituted by facts from which this follows). And for the fragmentalist, each time t is objectively present *simpliciter*—i.e. reality is constituted by the absolute fact that t is present (or by facts from which this follows). Here there is no significant relativity, even of an external sort, to the time in question. But in either case, presentness, in so far as it is a genuine feature of reality, applies equally to all times. Presentness is not frozen on a particular moment of time and the light it sheds is spread equitably throughout all time.

Of course, this feature, by itself, does not account for the passage of time. Consider the analogous first-personal case. Here the non-standard, or neutral, realist will suppose that no person is privileged—me-ness applies across the board to everyone. But that is hardly enough to secure a moving me! So clearly, something more than the equitable distribution of presentness is required to account for the passage of time. But at least, on the current view, there is no obvious impediment to accounting for the passage of time in terms of a successive now. We have assembled all the relevant NOWs, so to speak, even if there remains some question as to why the relationship between them should be taken to constitute a genuine form of succession.

8. The Argument from Truth

The second argument against the realist position arises from the need to account for the connection between language and reality or, more generally, between thought and reality.[18] Anyone who has a view as to what is real is under an obligation to explain how what is real accounts for what is true. The facts must be adequate to account for the truths and also largely essential—there should not, in general, be facts whose presence in reality is irrelevant to accounting for what is true. The present argument is to the effect that the realist is unable to provide a reasonable account of the connection between the truth of tensed utterances and the tensed character of reality. If, as I attempt to show in the next section, it is only effective against the standard realist, then it provides another indirect argument in favour of the non-standard view.

The argument make use of three critical notions—the notions of truth, content (or *stating*), and verification. We take *truth* to be a property of utterances, *stating* to be a relation between an utterance and a proposition,

---

[18] Considerations of this sort go back to Evans ([1985], ch. 12). The arguments I present bear a close resemblance to those stated in Mellor [1986 and 1998]), though I have been much more explicit about the assumptions upon which the arguments depend and how, in particular, they are meant to relate to a realist view. I am mystified as to why Mellor should regard them as versions of McTaggart's argument, which has nothing to do with the connection between language and reality. The reader will have noticed that I have been careful to eschew all reference to language in my own presentation of McTaggart's argument.

and *verification* to be a relation between a fact, or some facts, and a proposition. Talk of propositions and facts as objects in their own right is not strictly necessary and we might always adopt a sentential mode of expression in its place. Thus instead of saying that an utterance states the proposition that P, we might say that it states-that P, treating 'states-that' as a sentential operator; and similarly for verification.[19]

There are somewhat different ways in which these various notions might be understood. But it will be critical for what follows that they be understood in such a way as to be relevant to the connection between truth and reality. The truth of an utterance depends upon two factors: one, lying on the side of the language, is a matter of what the utterance states; the other, lying on the side of the world, is a matter of the facts. The various notions should therefore submit to the following constraint:

> *Link* An utterance is true if and only if what it states is verified by the facts (in reality).

Since the facts of interest to us are those that belong to reality, we may take it to be built into the notion of verification that only such facts are capable of verifying what an utterance states. The qualification 'in reality' in the formulation of Link is then redundant.

For the purposes of the argument, we should imagine that I make two utterances U1 and U2 of the sentence 'I am sitting'—one now while I am sitting and the other earlier while I was standing. We make two innocuous assumptions concerning their truth-value:

> *Truth-Value1* U1 is true;
> *Truth-Value2* U2 was false (i.e. not true).

We also make two assumptions concerning their content:

> *Content1* U1 states that I am sitting;
> *Content2* U2 stated that I am sitting.

It is important that these last two assumptions be taken to relate to what one might call a 'disengaged' use of the expression 'I am sitting'. Thus what U1 should be taken to state is the *tensed* proposition that I am sitting, one that does not itself encode any temporal information concerning the time of utterance; and similarly for U2. However, it is not important to the argument that each utterance be taken to state the particular tensed proposition that I am sitting. This is certainly the most natural choice, but the argument would work equally well with another tensed proposition in its place as long as it was the same for both utterances.

---

[19] The framework I adopt is close to that in Percival [1989 and 2002], although my treatment of 'facts' is somewhat different.

Finally, we make two general assumptions concerning the stability of truth-value and content:

> *Truth-Value Stability* If an utterance is true (false), then it is always true (false).
> *Content Stability* If an utterance states that P, then it always states that P.

These two assumptions should be taken to hold not merely at their time of utterance but at any time whatever.

We now show how these various assumptions lead to contradiction. By Truth1, U1 is true. By Content1, U1 states that I am sitting. So by the left-to-right direction of Link, there are facts—say f_1, f_2, . . . —that verify that I am sitting. By Content2, U2 stated that I am sitting. So by Content Stability, U2 states that I am sitting. Since the facts f_1, f_2 . . . verify that I am sitting, it follows by the right-to-left direction of Link that U2 is true. But by Truth-value2, U2 was false and so, by Truth-value Stability, U2 *is* false, i.e. not true. A contradiction.[20]

It is perhaps a weakness of the argument that it relies on the assumption that the content of the two utterances of the given type is the same. One might have the view, for example, that an utterance of 'I am sitting' at t states the conjunctive proposition that I am sitting at t and t is the present time. In this case, the content would be tensed yet different for different utterances of the same sentence-type.; and so the assumption would be false.[21] It would be odd if this were the only way in which the argument could be resisted but, all the same, it is worth noting that there is an alternative, somewhat more complicated, version of the argument for which the assumption is not required.

We now imagine that I make only a single current utterance U1 of the sentence 'I am sitting'. We assume Truth-value1, Content1, Truth Stability, and Content Stability, as before, and the left-to-right direction of Link. We make, in addition, the following three assumptions:

> *Fact* It is not always the case that I am sitting;
> *Factuality* If some facts verify P then those facts obtain;
> *Conditionality* If some facts verify P and those facts obtain then P.

---

[20] The argument can also be stated without appeal to temporal locutions, such as 'always states' or 'always true', the intelligibility of whose application in the given context might be doubted. Instead of saying that U1 is always true, for example, we may make a metalinguistic ascent and say that an utterance of 'U1 is true' will be true whenever it is made.

[21] This is what Percival [1989] calls a 'mixed' indexical view; both the content of the utterance and the truth-value of the content are allowed to vary over time (see also Percival [1991], 96). But he is thinking of a view, like the one described in Dummett ([1973], 382–400), in which the content is taken to be tenseless, whereas I have in mind a relatively innocuous case in which the content is taken to be tensed.

(A more natural reading of Factuality and Conditionality may be obtained by substituting 'putative fact' for 'fact', since this then leaves it open whether what verifies a proposition is something that obtains.) From these last two assumptions, we may infer: if some facts verify P then P. This is somewhat different from the right-to-left direction of Link, since we have P as the consequent, not the truth of an utterance that states that P.

The modified version of the argument now goes as follows. By Content1, U1 states that I am sitting; and so by Content Stability, U1 always states that I am sitting. By Truth1, U1 is true; and so by Truth Stability, U1 is always true. By the left- to-right direction of Link, it is always the case that some facts verify that I am sitting. But then by Factuality and Conditionality, it is always the case that I am sitting—contrary to Fact.[22]

How might the realist respond to these arguments? Consideration of this question is complicated by the possibility that there are different uses of the key notions and that some of the assumptions might hold for some of these notions and fail for others. There are two significant considerations that serve to fix how these notions might be understood. One, which has already been mentioned, is that the use of these notions should conform with Link. Indeed, it is hard to see how else the realist might provide an account of the connection between truth and reality. The other consideration derives from the realist's particular views about the nature of temporal reality. For presumably she does not simply hold that reality is tensed. She also believes that the truth of tensed utterances *requires* that reality be tensed, since otherwise the tensed character of reality will be divorced in a completely inexplicable manner from our ability to make tensed assertions. Thus we may also demand that the realist's use of the various notions also be in conformity with:

> *Relevance* A tensed utterance is only verified with the help of tensed facts,[23]

where the facts may be taken to verify an utterance if they verify what it states.

We may say that the use of the notions of truth, content, and verification is *(metaphysically) relevant* if it is in conformity with Relevance and Link. I have argued that the realist should be willing to abide by a metaphysically relevant use of these notions and the question we therefore face is whether

---

[22] Percival ([2002], 103–4) discerns the two forms of argument in Mellor's version of McTaggart. His own solution to the difficulties is to deny Truth-Value Stability on the grounds that present utterances will not exist in the past or the future and hence will not then be capable of being true or false. This strikes me as extreme and it would be preferable if we could find a solution that was compatible with more plausible ontological views.

[23] An exception should perhaps be made for trivial tensed utterances such as 'if I am sitting then I am sitting'.

there is a metaphysically relevant use of these notions that satisfies the other assumptions of the argument.

About some of these other assumptions, there can be no doubt. Anyone, realist or not, should accept the assumptions concerning Truth-Value and Fact. It is also hard to see how Conditionality or Content Stability might reasonably be doubted. Whatever our understanding of 'verify', surely it must be at least as strong as the material conditional; if some facts $F_1, F_2, \ldots$ verify P, then it is not the case that $F_1, F_2, \ldots$ obtain but P does not. Of course, Content Stability might be disputed on the grounds that even if, at the current time t, a current utterance of 'I am sitting' states the proposition that I am sitting, understood as the proposition that I am sitting at t, still, at a later time t^+ it will not state the proposition that I am sitting, now understood as the proposition that I am sitting at t^+. However, for the realist, the content of a tensed utterance is a *tensed* proposition and the use of the clause 'I am sitting' is consequently 'disengaged' from the time of utterance or the time of evaluation; and given that this is so, there then appears to be no good reason to dispute the assumption.

Factuality is very plausible as it stands, but we may also argue for it on the basis of two further assumptions:

> *Reality1* If some facts verify a proposition then they belong to reality.
> *Reality2* Whatever facts belong to reality obtain.

The first follows from the realist's conception of verification, which is meant to serve as a link between truth and reality. The second follows from his conception of reality; the (putative) facts which belong to reality are among the facts that obtain.[24]

One might argue for Content1—or for the claim that the tensed utterance U1 states a tensed proposition, even if not the proposition that I am sitting— as follows. Suppose that U1 does not state a tensed proposition. Then it states a tenseless proposition (granted that it states a proposition at all!). Since U1 is true, it follows by the left-to-right direction of Link that some fact or facts verify the proposition. But since the proposition is tenseless, presumably the fact or facts will also be tenseless—contrary to Relevance. Given Content1, Content2 is then also plausible in the case in which the very same content is in play. But even without Content2, we can rely on the second, more complicated, version of the argument. It is perhaps worth emphasizing that it is the metaphysical requirement of Relevance (in conjunction with Link) that forces upon us the relevant notion of content. There are no purely semantical considerations that might tell us what the relevant notion of content should be.

---

[24] It is the counterpart of the reality-operator of the T-axiom, $\Box A \supset A$, in modal logic.

Only one assumption remains: the Stability of Truth-Value. Now it might be thought that rejection of this assumption provides us with an easy way of dealing with the argument. For suppose we so understand truth that the truth of an utterance is taken to be equivalent to the truth of the corresponding type. The truth of an utterance will then vary with how things are at the time it is being considered for truth; and so Stability will no longer hold. Moreover, this notion of truth will be metaphysically relevant in the sense explained above as long as we understand the notion of verification in a correspondingly tensed way. A tensed utterance is currently true, for example, if the tensed proposition that it states is currently verified by the facts.

However, the mere existence of a metaphysically relevant notion of truth for which Stability fails does not exclude there being a metaphysically relevant notion for which it holds. And it seems to me that there is such a notion. We might put the issue in the following way. The realist wishes to maintain that the truth of a tensed utterance requires in general that reality be tensed, for a suitable notion of truth. We may take it that the way the truth of a tensed utterance requires reality to be tensed is through the mechanism of Link: the tensed utterance states a proposition that is verified only by the tensed facts. The question is therefore whether there is a stable notion of truth for which it is reasonable for the realist to maintain that this requirement should be met. Is there a stable sense of 'truth' for which the truth of a tensed utterance requires that reality be tensed?

There are two related considerations that strongly suggest that there is. The first involves appeal to our ordinary notion of truth. Our ordinary notion of truth, as applied to utterances, appears to be stable. Suppose I utter the words 'I am sitting' while sitting; and suppose that a few minutes later I stand up. Someone may then ask 'is that utterance KF made five minutes ago true?'. The correct answer is surely 'Yes', despite the fact that I am now standing.

We should not here be misled by the fact that is also correct and also more idiomatic to say that the utterance *was* true. For the use of the past tense here is plausibly taken to relate to the past existence of the subject rather than to the past application of the predicate. It is similar with a sentence such as 'his heart attack was sudden'. What makes the use of the past tense appropriate, indeed mandatory, is that the heart attack occurred in the past, but the property that is attributed to the heart attack, that of being sudden, is not sensibly taken to vary over time; the property of being sudden is one that applies to an event either tenselessly or whatever the time.

Contrast this with a sentence, such as 'KF was sitting', in which the past tense relates to the application of the predicate. It is not just (or not even) KF that is taken to be in the past but also his sitting. In cases of the latter sort, we can sensibly qualify with 'once' or 'no longer'. Thus we can say that 'KF was once sitting' or that 'KF is no longer sitting'. But not with cases of the former

sort. We cannot say 'the heart attack was once sudden' or 'it is no longer sudden'. And similarly with 'KF's utterance was true'. We cannot sensibly say that it was once true or that it is no longer true.

Now it seems to me that in so far as we have any inclination to claim that the truth of a tensed utterance requires that reality be tensed, it is in the ordinary sense of 'true'. I utter the words 'I am sitting' while sitting and that utterance is in the ordinary sense true. And surely it is in that very sense of 'true' that the realist wishes to claim that the truth of the utterance requires that reality be tensed. Indeed, it would be odd if the ordinary notion of truth did not provide a link between tensed language and the realist's conception of temporal reality and if it was only in some artificial custom-made sense of the term that he could explain what the connection was.

However, the issue need not be made to turn in this way upon what we take our ordinary notion of truth to be. It will be agreed that the truth of a *current* tensed utterance requires that reality be tensed. Now there is a sense of 'true' (whether the ordinary sense or no) in which it will still be correct to say a moment later in time that the utterance is true. But surely we are unable to discern any metaphysical difference between the truth of the utterance at the one time and at the other time. In so far as we are inclined to say now that the truth of a current utterance requires that reality be tensed, then surely we are equally inclined to say a moment later that its truth requires that reality be tensed. We have no sense of the metaphysical ground for the truth of the utterance shifting under our feet, as it were, as we go from the one time to the other.

What these considerations bring out is the way in which we are willing to adopt an eternal perspective of what the truth of a tensed utterance might require of reality. The requirement is the same whether we consider the truth of the utterance at one time or at another; and to the extent that this is so, it is impossible for the realist to evade the argument by appeal to the instability of truth.

It is worth noting that the corresponding arguments do not work in the modal case (as has often been pointed out). We are under no inclination to take the utterances of contingent truths to be necessarily true and nor does there appear to be a metaphysically relevant notion of truth, one relating the truth of an utterance to the worldly facts, that is stable across worlds. Thus the corresponding arguments in the modal case can simply be met by rejecting the Stability of Truth-Value. However, the corresponding arguments *do* work in the first-personal case. Take a true utterance of 'I am hungry' (or, ignoring tense, 'I am hungry at time t'). Then it would be correct for *you* to say that the utterance was true, regardless of whether you were hungry; and this notion of truth is surely relevant to how the truth of the utterance relates to the facts. In this respect, the better analogy is again with the metaphysics of the first-person rather than with the metaphysics of modality.

9.   The Non-standard Response to the Argument from Truth

The above arguments are naturally taken to favour the anti-realist position. For the anti-realist does not believe in tensed facts and so does not believe that they are required for the verification of tensed utterances. There is therefore no principle such as Relevance that might force him to accept a sense of the key terms of the argument under which all its assumptions will hold.

However, it seems to me that the non-standard realist is also able to resist the argument. Unlike the anti-realist, she will accept that the content of a tensed utterance (in a metaphysically relevant sense) is a tensed proposition. But she will reject the Link principle by which the connection between truth and reality is to be mediated. This is because she does not believe that there is a single coherent reality to which the truth of an utterance is to be referred. Rather there are many such realities; and in accounting for the truth of an utterance one must specify the reality upon which it is taken to bear. This, naturally enough, is taken to be the reality at the time of the utterance made. Thus in place of Link, we have:

> *Relative Link* An utterance is true if and only if what it states is verified by the facts that obtain at the time of utterance.

Given Relative Link in place of Link, the contradiction can no longer be derived.

Although the modified version of Link might appear innocuous, it represents a radical departure from the way language is usually taken to connect with reality. For what is taken to verify what the utterance states is not a tenseless fact but a tensed fact whose provenance varies with the time of utterance. Thus it is no longer supposed that there is a single reality and a single set of facts in virtue of which an utterance is true. Rather, the reality and the appropriate set of facts will vary with the utterance.

The usual model might be depicted as in Fig. 8.2. The utterance determines a proposition; and reality determines whether the proposition is true,
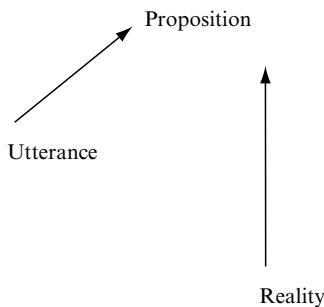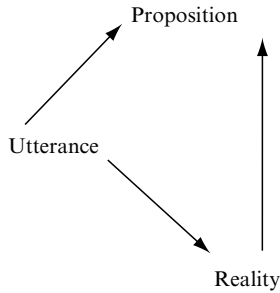
thereby determining whether the utterance is true. The alternative model, by contrast, might be depicted by Fig. 8.3. Here the utterance has a dual role: it not only determines a proposition but also the relevant reality; and whether the proposition or utterance is true is then determined by the reality.

If a tensed utterance is made at time t then there is a sense in which it is about that time just as there is a sense in which any utterance is about the actual world. But it is about the time or the world as a *target* rather than a *topic*. The utterance is not stating something about the time or the world; it is not stating—or, at least, not simply stating—that this is how things are at the given time t or the given world w, since that would give us a tenseless or a necessary proposition. Rather it is about the given time or world in the sense that it is facts that constitute how things are at the time or in the world that are relevant to determining whether the utterance is true. Normally there is no need to be explicit about the target of an utterance, since there is only one reality to which it can be directed. But once we adopt a non-standard form of realism, the target is no longer exogenously determined and must be regarded as a function of the utterance itself.

This view therefore gives a new meaning to the term 'indexical'. A sentence is usually taken to be indexical if its truth-value or if its content can vary with context (similar remarks apply, of course, to other categories of expression). But here it is the target—the tensed state of the world with respect to which the sentence is to be evaluated—that may vary with context. We may say, if you like, that an utterance of a tensed sentence has a *compendious* content, consisting of a tensed proposition and a specification of the time. But the two components of the content play completely different roles: the first is *factual* and serves to indicate the tensed conditions or facts under which the utterance is true, while the second is *focal* and serves to indicate where those facts are to be found.

The present realist's distinction between factual and compendious content corresponds to the anti-realist's distinction between character and propositional content and, indeed, the formal representation of the two may be the same. However, the realist's conception of compendious content is quite

different from the anti-realist's conception of propositional content. For what makes a propositional content true, for the anti-realist, is a tenseless fact, while what makes a compendious content true, for our realist, is a tensed fact. Thus the content is still tensed even though it incorporates the specification of a time.

The cardinal mistake of the anti-realist, according to the realist, is to assimilate focal content to factual content. He thinks that different utterances of the same tensed sentence will differ in their truth-value on account of a difference in their factual content. But what accounts for the difference in truth-value, according to the realist, is a difference in the focal content; for when the focal content targets a different state of the world, the utterance may change its truth-value even though the very same facts are relevant to its being true.

The realist can agree that a tensed utterance is in some sense about the time of utterance. He might even agree that the time of utterance is part of the factual content of the utterance—either directly, *qua* particular time, or indirectly, *qua* time of utterance—though this might be an odd view for him to have. But the anti-realist's mistake, he wants to say, lies in the refusal to recognize another, metaphysically more basic, way in which the utterance can be about the time of utterance. For it can be directed towards a time, or towards how things are at the time, in much the same way in which an ordinary utterance is directed towards how things are in the actual world. By attempting to incorporate the target of the utterance into its content, the anti-realist loses both what is semantically distinctive about its indexicality and what is metaphysically distinctive about its content.

We therefore see that the non-standard realist can effectively respond to the first version of the argument. But a little more needs to be said about the second version. This depended upon only the left-to-right direction of Link and, for the non-standard realist, this direction is unproblematic. For if an utterance is true, there will indeed be some fact (obtaining at the time of the utterance) that verifies what it states. How, then, is she to deal with the second version of the argument?

This employed three additional assumptions in place of the right-to-left direction of Link—Fact, Factuality, and Conditionality; and the third of these was derived from two further assumptions—Reality1 and Reality2. Fact and Factuality cannot sensibly be denied; and so that leaves the two Reality assumptions. We might naturally take a fact to belong to reality if it belongs to *a* reality. It is then the second Reality assumption, that any fact belonging to reality obtains, which should be given up; for the fact may relate to one reality and the obtaining to another. In stating that a fact belongs to reality, we adopt a general perspective but, in stating that a fact obtains, we adopt the current perspective; and it is because of this shift in perspective that we cannot generally assert that the facts belonging

to reality will obtain. Thus, once again, it is the absence of a single coherent reality that allows us to reject one of the assumptions upon which the argument depends.

10.   The Argument from Special Relativity

The final argument against tense-theoretic realism arises from the need to square it with Einstein's special theory of relativity (I ignore the complications arising from general relativity and quantum physics). Special relativity (SR) has weak readings in which it is compatible with there being an absolute notion of simultaneity and a strong reading in which it states, or implies, that there is no such notion. I am inclined to think that there are good scientific reasons for favouring the strong reading but, whether or not this is so, my concern here is with the compatibility of tense-theoretic realism with the strong reading.

This issue of 'compatibilism' has been much discussed in the literature.[25] But the focus of my own discussion will be somewhat different (though there will be some obvious points of contact). In the first place, my interest is in the compatibility of SR with tense-theoretic realism and not, as is more common, with presentism. Tense-theoretic realism is the weaker position and, consequently, incompatibility with realism is the stronger result. In the second place, I shall be careful to distinguish between the standard and non-standard forms of realism. The arguments for incompatibility are, I believe, effective against the standard forms of realism but not against the non-standard forms. Thus in one respect the negative conclusions of this section are more far-reaching in their intent than those to be found in the literature, since they are also meant to apply to non-presentist versions of tense-theoretic realism, and, in another respect, they are less far-reaching, since they are not meant to apply to the non-standard forms of realism.

Let us begin with the distinction between presentism and tense-theoretic realism. Presentism—or what, in the present context, might be called *ontic* presentism—is the view that only presently existing things are 'real' in some or another sense of the term.[26] There are familiar variants on this view. One could hold that only past or present things are real or that only past or present or determinately future things are real. These differences will not be important in what follows. There are also differences in what one might mean by 'real'. My own preferred reading is that for a thing to be real is for it to belong to reality, i.e. for it to figure in the facts that are constitutive of

---

[25] See e.g. Christensen [1974], Clifton and Hogarth [1995], Craig [2001], Crisp [2003], Godfrey-Smith [1979], Hinchliff [2000], Markosian [2002], Mellor [1974], Putnam [1967*b*], Rakic [1997], Savitt [2000], Sider [2001], Sklar [1981], Stein [1968, 70, 91], Tooley ([1997], ch. 11).

[26] Sider [1999] contains a recent discussion.

reality. But most philosophers take a thing to be real, in this context, if there is something that it is, in the broadest sense of the phrase. They are therefore of the view that what there is, in this broadest sense, is what there presently is. Thus whereas they would take there to be no such thing as Lincoln, we may allow that there is such a thing; it is just that it will not figure in the preferred account of how things are.

Tense-theoretic realism, by contrast, is the view that reality is tensed; reality comprises tensed facts (and perhaps tenseless facts as well). Tense-theoretic realism is, in its own way, a form of presentism; for, in so far as reality comprises tensed facts, it must be oriented towards the present. We might therefore call it *factive* as opposed to ontic presentism. It is concerned with the nature of *reality*, with how things are, rather than with the nature of the *real*, or with what things are.[27]

It should be evident that the two views are different. Ontic presentism is an ontological position; it is a view about what there is. Factive presentism, on the other hand, is a metaphysical rather than an ontological position; it is view about how things are, quite apart from what there is. In this respect, our formulation of the view in terms of facts may be misleading, since it suggests that the factive presentist will subscribe to a distinctive ontology of facts. But in endorsing tensed facts, he is merely endorsing a certain way things are. Moreover, ontic presentism is a negative view; it excludes certain things from what there is. Factive presentism, on the other hand, is a positive view; it includes certain ways of being in how things are. Consequently, an ontic presentist will have a more restrictive view than his opponent of what there is while the factive presentist may have a more inclusive view than his opponent of how things are.[28]

It is readily possible for a factive presentist not to be an ontic presentist. Indeed, he may endorse a full ontology of things past, present, and future: all such things may figure in his preferred account of reality; and he may think that there is a perfectly intelligible sense in which there are such things. He merely insists that some of the facts (if not all) should concern how things presently are. I am inclined to think that this version of factive presentism is much more plausible than the usual version, in which only present things are taken to exist; and it is a shame that a one-sided conception of the presentist issue has prevented philosophers from taking it more seriously.[29]

---

[27] The distinction should be a familiar one (it is made on the first page of Fine [1977*a*], for example). But it tends to be overlooked, especially in the context of the present discussion.

[28] This explains why presentism is usually taken to be an anti-realist position whereas I take it to be a realist position. The one is anti-realist in regard to what there is while the other is realist in regard to how things are. Presentists in my sense have sometimes been called 'A-theorists' or 'tensers', though there is no established terminology.

[29] Q. Smith ([1993], ch. 5) is an exception. Sider ([2001], 18) argues against the combination of the views on the grounds that it is 'unmotivated'; for why not give an analysis of tense if the

Ontic presentism, by contrast, does not really make sense except in the context of factive presentism. There is no strict implication from one to the other but, given that all the facts are tenseless, it makes no sense to restrict the ontology to presently existing things. Thus any argument against factive presentism is, *eo ipso*, an argument against ontic presentism.

Most philosophers who have worried about the compatibility of SR with presentism have worried about its implications for the ontology of presentism (or of some variant thereof). What, in the light of SR, should the presentist take to be real? If SR excludes *being present* as a criterion for being real, then what should be used in its place? But there is a more basic worry. For the presentist believes in tensed facts. But what, in the light of SR, should he take a tensed fact to be? Without an answer to this question, he is not even in a position to state an alternative criterion for being real, since any alternative criterion must presumably be tensed and hence must presuppose some alternative conception of tense.

The difficulty that SR poses for the conception of tense is this. Under the pre-relativistic conception of tense, a tensed proposition is one whose truth is simply relative to a time. Consider now any two events *e* and *f* and the tensed propositions that *e* obtains and that *f* obtains. If it makes sense to say that these propositions are true at any given time, then it makes sense to say that they are true at the same time. But for the propositions to be true at the same time is for the events to be simultaneous. Thus the classical pre-relativistic conception of tense presupposes an absolute notion of simultaneity.

What then, in the light of SR, should take its place? What should replace times as the standpoint from which the truth of tensed propositions is to be evaluated? There are two main options. Under the first, the truth of a tensed proposition is taken to be relative to a location in space-time. Thus the proposition that a given event is *here-now* may legitimately be regarded as tensed even though the proposition that the event is *now* or that it is *here* cannot be. Under the second option, a tensed proposition is taken to be true relative to an (inertial) frame of reference and a time. Each frame gives rise to its own framework of times;[30] and a proposition may then be taken to be true relative to the frame and one of its times. Thus the proposition that a given

---

materials for it are available? But the required materials may not be available if tense is taken as primitive in preference to the earlier–later relation; and even if the required materials are available, any proposed analysis might be rejected on the grounds that it fails to respect the distinctive metaphysical character of the tenses. I have given a similar argument against the combination of views in the modal case (Ch. 6, §1 above). The argument in this case is perhaps somewhat more plausible but might still be resisted.

[30] Which is subject to the condition that two events occur at the same time just in case they are simultaneous within the frame. We may suppose that simultaneity within a frame is characterized in the usual way.

event is *now* may legitimately be regarded as tensed, as may the proposition that a given thing is *now at rest*.

It seems to me that both these proposals are open to formidable objection. There is nothing wrong as such with the post-relativistic counterparts to the pre-relativistic notion of tense. The difficulty arises from taking tensed facts in this post-relativistic sense to be constitutive of reality. For if I take reality to be constituted in part by tensed facts, then I should be able to say what those tensed facts are. So here I am; and let me raise this very question.

Now which tensed facts I take to obtain will depend upon the standpoint from which I ask the question. Different tensed facts will obtain at different standpoints and so, when I ask this question, which is the standpoint from which I take the facts to obtain? Which of the various possible standpoints is the standpoint of reality?

There appears to be only one possible answer. The standpoint of reality is the standpoint that I occupy. Under the first proposal, this will be the space-time location from which I ask the question; and under the second, it will be the frame at which I am at rest when I ask the question and the time within that frame at which I ask the question. Indeed, if the standpoint from which the question is to be answered were not the standpoint that I occupy, then it is hard to see what else it might be. If a different space-time location or a different frame-time pair, then which?

But even if these are the only possible answers, can they be sustained? Consider the second proposal first; and imagine that you and I are in relative motion and that we coincide at the location at which I ask the question. Then what reason do I have to favour my own standpoint over yours? After all, the only possible relevant difference between us lies in our relative motion. But why should I think that reality is somehow attuned to my motion as opposed to yours?

A similar point holds in regard to the first proposal, though it is perhaps not quite as compelling. Suppose that you are standing next to me and that you also ask the question. Then why should I favour my standpoint over yours? Now if your question is asked in the absolute past or future of my question, then I do have a good presentist reason to favour my own standpoint. But what if the events of our asking the questions are space-like separated from one another (as they would be if I took you to be asking the question as the same time as myself)? What then? After all, space-like separation is as close as one can get to a purely spatial difference within the context of SR and so, if anything, it would appear to constitute a reason for admitting your standpoint rather than excluding it. But failing spatial separation, there is nothing about the difference between the two standpoints to which we can appeal in explaining why reality might be attuned to one as opposed to the other.

The force of the argument can be brought home by means of an analogy with the case of first-personal realism. Suppose I believe that reality is, in

part, constituted by first-personal facts; and I now ask what those facts are. Then surely I have no good reason to suppose that reality is somehow oriented towards my own standpoint as opposed to yours, that the only first-personal facts are those that concern me as opposed to you. This would appear to be metaphysical chauvinism of the worst sort. But similarly, it may be argued, for the presentist. Some sort of chauvinism about his own standpoint will be unavoidable once his conception of a standpoint is reconceived in the light of SR.

The general form of the argument in these cases is as follows. We wish to privilege our own standpoint as being the standpoint of reality. This requires that we *explain* why we should take our standpoint, and not also some other standpoint, to be *the* standpoint of reality. Thus given that s is a standpoint of reality and that t is not, we want to find a relation R that is such that t's standing in the relation R to s explains (or helps explain) why t is not also a standpoint of reality. It is then argued that in the cases at hand this cannot be done. In the first case, for example, the only possibly relevant relationship between the coincident observers is that they are in relative motion; in the second case, the only possibly relevant relationship between the neighbouring observers is that they are space-like separated; and in the third case (of first-personal realism), the only possibly relevant relationship between you and me is that we are different. But in none of these cases are the relationships adequate to the task. We are at loss to understand why a difference in relative motion or spatial separation or mere identity might prevent another standpoint from also being a standpoint of reality.

It might be countered that there *is* a relevant difference in all these cases. Our mistake has been to look to tenseless or non-indexical features to account for the difference in status. But once we take account of tensed or indexical features, the problem is readily solved. The standpoint of the coincident observer is not the standpoint of reality since he is not now at rest; the standpoint of my neighbour is not that of reality since he is not here-now; and you do not occupy the standpoint of reality since you are not *me*.

It must be conceded that features of this sort may be relevant to solving the problem in certain cases—and even essential. If the pre-relativistic presentist is asked why past and future observers do not occupy the standpoint of reality, then he can legitimately appeal to the fact that they are not present. And if an actualist is asked why a merely possible world does not constitute the standpoint of reality, then there is nothing he can do but appeal to the fact that the world is not actual. However, the responses in these cases are intuitively satisfying. Being present or being actual can sensibly be seen to bear upon the question of whether a given standpoint is the standpoint of reality. But not so in the cases in question. How can being at rest, or being here-now, or being me sensibly be seen to bear upon the question? From this point of view, then, the problem with the post-relativistic conception of tense

is that it no longer yields an intuitively satisfying *tensed* criterion for being the standpoint of reality.

This naked appeal to intuition may perhaps be reinforced by two other considerations. We may note first that, no matter how crazy an indexicalist view we might have, it is always possible to provide a corresponding indexical criterion for being the standpoint of reality. The spatiocentric realist, for example, might appeal to the fact that he was *here* to distinguish his standpoint from those who were elsewhere. This means that an indexical response has no probative value in itself and that it is therefore especially important to see if the response is intrinsically plausible or to see if some other form of response might be available. Second, a non-indexical form of response *is* available to the pre-relativistic presentist. If he is asked why the standpoint of a past or future observer is not the standpoint of reality, he can appeal to the fact that it is *earlier* or *later* than the present standpoint. But then it is hard to see why a response of this sort should not also be available to the post-relativistic presentist.

We have so far dealt only with the two most obvious ways of modifying our conception of tense in the light of SR. But might there not be others? There are two key respects in which the previous accounts might be extended. First, we might take cognizance of further information about the observer. Within the context of SR, perhaps all that might be considered relevant is his world-line and his location (when he asks his question). Thus a standpoint will at least be determined once we specify a world-line and a location on that world-line.[31] Second, we may wish to treat certain standpoints as giving rise to the same reality. This is especially true if we pack a great deal of information into the standpoint, since some of that information may then be irrelevant. If reality is taken to reside in a location, for example, then any two world-line/location pairs should be treated as the same when their locations are the same; and if reality is taken to reside in a frame-time pair, then any line/location pairs should be treated as the same when their world-lines correspond to the same frame and their locations to the same time. But even under a relatively meagre conception of what belongs to a standpoint, we may still wish to identify certain standpoints since this provides a possible solution to our previous difficulties; for if your standpoint and mine give rise to the same reality, then there is no longer any need to find a criterion by which they might be distinguished. However, it may be shown, once the problem is set up in this way, that there is no reasonable basis upon which a suitable equivalence between world-line/location pairs might be determined.[32]

---

[31] This is essentially the framework of Clifton and Hogarth [1995].

[32] I omit the details, which are somewhat messy. The result is an extension of proposition 3 of Malament [1977]. It is important to stress that our relation of equivalence is different from the

To make matters worse, there are some further, independent arguments against the proposals we have considered. Consider again the frame-theoretic proposal. The presentist takes there to be an absolute and objective sense in which a given frame-time is the standpoint of reality. He is therefore in a position to distinguish a particular frame as the frame of this standpoint; and this then enables him to characterize an absolute notion of simultaneity as simultaneity within this frame. Thus this proposal is not in keeping with the spirit of SR (under the strong reading in which it is denied that there is an absolute notion of simultaneity). The peculiarly and purely metaphysical way in which the absolute notion is derived might also be regarded as especially objectionable.

This leaves the locational proposal; and given that the argument from arbitrariness is less compelling in this case, one might well think that this relation constitutes the least undesirable of the various options. However, there is a telling objection to this proposal as well. One of the primary motivations for the presentist view is that it enables one to distinguish between space and time. Temporal indexicality is metaphysically significant, while spatial indexicality is not; there is an objective 'now', even though there is no objective 'here'. However, once we adopt the locational view, this asymmetry between space and time disappears. The two forms of indexicality collapse one into the other and reality can no more be said to be oriented towards a temporal standpoint than towards a spatial standpoint; the 'here' is as objective as the 'now'.

I suppose that it one were forced to treat space and time alike, then there might be something to be said for treating the indexicality of both in a realist manner, though it is far from clear why the presentist's post-relativistic views about the metaphysical status of the *here-now* should be taken to derive from his previous realist views about the *now* as opposed to his previous anti-realist views about the *here*. After all, anti-realism about tense is a straightforward view that has been held by many, while realism about spatial indexicality is a bizarre view that has been held by few, if any; and so it would seem more reasonable to move in the direction of a general form of anti-realism.

relation *R* that was introduced by Putnam [1967b] and subsequently discussed by other philosophers. Their relation is one that holds between *x* and *y* when *y* is real for *x* (thus *x* should be regarded as a standpoint and *y* a thing within the ontology of the standpoint). Our relation is one that holds between *x* and *y* when they are identifiable standpoints, i.e. ones that give rise to the same reality. It may reasonably be questioned whether the former relation is symmetric and perhaps even whether it is transitive. But this is not sensibly open to question in the case of our own relation. It should also be stressed that our negative result is meant to hold regardless of what the ontology of the presentist might be. He could even accept the full ontology of things past, present, and future; the difficulty of coming up with a reasonable conception of tense would still remain.

However, one is not forced to treat space and time alike. SR embodies enough of an asymmetry between space and time that the difference between the spatial and temporal forms of indexicality can be retained. Indeed, one can simply take spatial indexicality to be relativity to a frame-place pair and temporal indexicality to be relativity to a frame-time pair. One thereby obtains straightforward counterparts to the original forms of indexicality (with the frame of reference playing the role of a missing parameter). But given that the distinction between the two kinds of standpoint *can* be retained, it seems bizarre to adopt a view in which the distinction is abandoned and one of the principal motivations for the presentist position is lost.

In the light of these considerations, it is hard not to agree with those philosophers who have thought the situation is quite hopeless (though their focus has been on the ontology of presentism rather than the metaphysics).[33] There is an unwarranted arbitrariness in taking any given standpoint to be a standpoint of reality; and, depending upon which proposal is adopted, one is obliged either to accept an absolute notion of simultaneity or to relinquish the metaphysical distinction between space and time.

However, it seems to me that all the above arguments, with the exception of the last, depend upon taking for granted that standard realism is the only realist option. When it is asked 'how is reality?', it is presupposed that there is a single reality and hence a single standpoint (or class of equivalent standpoints) from which the question is to be answered. This then leads to the difficulty of saying what the standpoint is and of being committed, under certain responses, to an absolute notion of simultaneity.

But suppose that we give up this assumption. Each (representative) standpoint will give rise to its own reality and no one can be singled out as being *the* standpoint of reality. If it is asked 'why is your standpoint not also a standpoint of reality?', then the answer is that it *is*, regardless of your world-line or location and regardless of what might be true at your standpoint. Thus the problem of distinguishing my standpoint from yours does not arise and nor will it be possible to define an absolute notion of simultaneity in terms of a privileged standpoint.

In principle, one could adopt the non-standard realist view with respect to either the locational or the frame/time conception of tense. But the argument from the asymmetry of space and time still applies against the locational conception; and so this leaves the frame-theoretic conception as the only viable option. We may conclude that the frame-theoretic form of non-standard realism constitutes the only reasonable way of reconciling presentism with SR; and so this provides yet another reason for favouring a non-standard form of realism.

---

[33] See Savitt [2000], for example.

The resulting metaphysical view is quite remarkable. The usual view is that SR shows space-time to be Minkowskian rather than Newtonian; physical processes are to be seen as taking place within a physical space-time with the structure of Minkowskian rather than Newtonian space-time. But the present view is that what SR shows to be mistaken is not that space-time is Newtonian but that there is a single space-time. Thus we should picture physical processes as taking place within a plurality of physical space-times, each of them enjoying a common ontology of space-time locations and each of them Newtonian in structure, and yet differing in the spatial and temporal relationships that hold among the space-time locations.[34]

Of course, Minkowski space-time and the corresponding family of Newtonian space-times can be regarded as providing equivalent descriptions of the same underlying reality. But our interest is in the underlying reality itself; we want to know what *au fond* we should take the spatio-temporal facts to be. Are the different Newtonian space-times merely mathematical abstractions from Minkowski space-time? Or the other way round? Normally, a question of this sort would be taken to involve an awkward choice as to which entities and which properties or relations should be taken as basic. But in the present case, the issue turns on what *kinds* of fact should be taken to compose reality. Do we take the simultaneity of two events to be a fact of the right kind or only the simultaneity of two events relative to a frame of reference? In the former case, we must posit many different realities, or space-times, to accommodate the different spatio-temporal relationships that may hold between the same events whereas, in the latter case, we can make do with a single all-encompassing reality.

We can be either relativist or fragmentalist about the resulting realities. Under the relativist view, the different realities will be indexed to different frame-times. This appears to require, if the indexing is to have any real significance, that we have an independent conception of the frame-times, one that gives them an identity that is separate from the space-times to which they give rise. Under the fragmentalist view, by contrast, there will be a single though incoherent über-reality. Given a suitable notion of coherence, the content of the different frame-times can then be recovered as maximally coherent sets of facts.[35]

The present position may be understood by analogy with the usual pre-relativistic forms of tense-theoretic realism. Before, the fact that a given event was present was taken to be tensed, since its obtaining was relative to a time. Now, under the extended conception of tense, the fact that two

---

[34] It is perhaps this view—or, at least, a consequence of it—that Stein excoriates in several of his papers ([1968, 1970, and 1991]) but, non-standard as it may be, I do not see how it can so easily be dismissed as 'illegitimate'.

[35] It goes without saying that the fragmentalist viewpoint may have application to other areas of physics!

events are simultaneous is also regarded as tensed, since its obtaining is relative to a frame (the time in the frame-time pair here drops out of view). Before, the fact that a given event was present was taken to be absolute and capable of belonging to reality, notwithstanding its relativity to a time. Similarly, the fact that two events are simultaneous is now taken to be absolute and capable of belonging to reality, notwithstanding its relativity to a frame. Finally, in order to avoid privileging one time over another, we took the facts that a given event was present, past, and future to be equally capable of belonging to reality (whether to a fragmented reality or to one that is indexed to a time). Now, in order to avoid privileging one frame over another, we take the facts that two events are simultaneous or that either one is earlier than the other to be equally capable of belonging to reality (which, again, is either fragmented or indexed to a frame-time).

The big difference in the views is not only in the conception of tense but in what is taken to be tensed. Before the simultaneity or precedence of two events was taken to be a tenseless fact, but now it is taken to be tensed. Thus far from eliminating the metaphysical significance of tense under SR, the present view results in a broadening of its scope.

Of course, it is also possible for someone who is anti-realist about tense, as it is normally conceived, to adopt the present metaphysical gloss on SR. He may accept the many space-times and yet reject their orientation towards the present; and there may even be something to be said for such a view. But what is interesting about the realist is that he appears to have no choice in the matter; in accepting the absolute reality of the present, he is thereby forced to accept the absolute reality of ordinary temporal relations and the multitude of space-times to which they give rise.

11. Fragmentalism

This concludes our discussion of the arguments in favour of a non-standard form of realism. My concern, in the fourth and final part of the chapter, is to discuss the bearing of this position on a number of different topics: the question of relativism versus fragmentalism; the metaphysics of first-personal realism; and the interplay between our concept of reality and our conception of what is real.

I begin with the question of whether the non-standard realist should be a relativist or a fragmentalist. The previous arguments in favour of non-standard realism were largely neutral on this question. However, it seems to me that there are reasons for preferring the fragmentalist view, even though it is far more radical than the relativist view and might even be regarded as metaphysically repugnant. In the present section I would like to outline these reasons, though I am under no illusion that they can be regarded as anything more than suggestive.

One reason is specific to Special Relativity. I have suggested that the non-standard realist might think of reality as dividing up into different Newtonian space-times, each with its own version of what is present. In pre-relativistic physics, these centred space-times will be subject to the condition that the relative temporal status of events (as past, present, or future) must be preserved. Thus if two events are both present in one centred space-time they must both be past or both be future in any other centred space-time. This way of thinking is forced upon us if we think in a relativist way of presence as presence at a time. But suppose one just thinks of presence as the general manner in which an event becomes temporally manifest. There would then appear to be nothing to prevent us from envisaging a different rule of coherence, one that allows events that were both present in one manifestation of temporal reality to be past and present, say, or past and present in another manifestion of temporal reality. One can divide the tensed facts diagonally, as it were, and not just across the vertical (this would be analogous to allowing different minds to share the same experience). Thus fragmentalism has the advantage of more readily lending itself to the kinds of changes in our conception of space-time that SR seems to require.

There is, however, a more significant, and more general, reason for favouring the fragmentalist position. This concerns the ontological status of times (or whatever other kind of standpoint might be in question). Many philosophers have been tempted by the view that times are not among the fundamental constituents of reality. But this view is especially attractive for the tense-theoretic realist and, since it will be an important premiss in the argument to follow and in our subsequent consideration of first-personalism, it will be worth investigating further. (Similar reasons will also apply in the cases of modal and first-personal realism.)

If times are to be among the constituents of reality, then of which basic facts will they be constituents? They are two possibilities: (a) they appear in tensed facts; and (b) they appear in tenseless facts. If they appear in tensed facts, then which? Suppose, simply for the sake of illustration, that *raining* is a primitive feature of reality. Then could a particular time appear as a constituent in a tensed fact to the effect that it was raining? Clearly, it will not do to say that the fact is to the effect that it is raining at the particular time t, since this fact is tenseless, not tensed. One might, at this point, be tempted to suppose that in addition to times, as normally conceived, there is what one might call a tensed time, *the present*, and that this time can appear as a constituent in the tensed fact to the effect that it is raining at the present. But not only is it hard to know what the present is, if it is not just the time, as normally conceived, that is present, it is hard to see how the presence of a mere particular in a fact could make it tensed. A particular, one wants to say, is not the kind of thing that can be responsible for the tensed character of a

fact; what is responsible for that character, if anything, is the way the particular is picked out, not the particular itself.

But if times do not appear in ordinary tensed facts, then in which other basic facts might they appear? Only one plausible answer suggests itself. There is a primitive property of *presentness* and the way in which a time t appears in a tensed fact is through its possession of this property (and similarly, one might think, for other tensed determinations).

Now I take it that the only plausible view of this sort is one in which times also occur as constituents in ordinary facts such as the fact that it is raining at t. One could in principle have a view that admits the tensed fact that t is present and also the tensed fact that it is raining. But the ontology of times would then be strangely divorced from what was happening in time.

Once we have the fact that the given time t is present, there is no need for any other facts concerning what is going on at the present time. For suppose, by way of illustration, that it is presently raining. Then it will be a fact, under the proposed view, that it is raining at t; and from this fact and the fact that it is present, it will then follow that it is presently raining. In this way, the present disposition of reality may be determined from the tenseless facts and the fact that the given time t is present.

However, this is highly counter-intuitive. For consider tensed truths concerning what is presently happening, say 'it is raining' and 'it is cold'. These will be made true, in part, by tensed facts. Now we have a strong intuition that it is separate tensed facts that will help verify the two statements. But this is not an intuition that can be respected. For once we have the tensed facts that help verify the first statement, there can be no further tensed facts that might help verify the second. Thus the composite character of present reality must be denied.

If therefore appears that, if times are to be constituents of any facts whatever, they must be constituents of tenseless facts. But again, which will they be? One option is that they are constituents of ordinary facts. Thus we might take it to be a fact that it is raining at t. But we have seen that if there are to be any tensed facts whatever, we must also take it to be a fact that it is raining. We will therefore have the situation of two facts covering the very same metaphysical ground. And this seems bizarre. For why have two facts when we can get by with one? The other option is for times to be constituents of distinctively time-theoretic facts, such as the fact that one time is earlier than another, but not of ordinary facts. But the ontology of times will then again be strangely divorced from what is happening in time; for we will have some tenseless facts that specify the abstract structure of time and some tensed facts that indicate what is happening over time, but without any apparent connection between them. Thus we see that, in either case, there seems to be no room for an ontology of times within a realist tense-theoretic metaphysics.

We may now return to the main line of our argument. Given that times are not among the fundamental constituents of reality, it is plausible, once reality claims are expressed in their most basic terms, that they will involve no reference to time; for surely any reasons for thinking that times are not basic should apply across the board, not only to their role in the specification of the facts that are real but also in the formulation of the reality claims themselves. Thus if we wish to say that reality is relative to a time, then we must be able to find some more basic way of expressing the kind of relativity that might be in question. But what might this be? We cannot sensibly take such claims to be relative to an event or momentary object, even if our ontology allowed for such things.[36] We might take the facts to be relative to a centred world. But what is this centred world if not the facts that compose it? And so to say that the facts are relative to a world is simply to say that they divide up into different worlds. Thus it is not clear how there might be a significant and sensible form of relativism once all reference to times is dropped; and so we seem to be forced into favouring the fragmentalist position once we opt for a non-standard form of realism.

We might mention, in conclusion, another reason for favouring fragmentalism that may, at least, be persuasive to some. For the coherence of the relativism depends upon distinguishing between the internal and external forms of relativity—between reality's being a-certain-way-at-a-time and its being a-certain-way, at the time (or whatever the standpoint might be). But it is hard not to feel the temptation to collapse the one form of relativity into the other. Fragmentalism, by contrast, avoids this difficulty. Reality is simply a certain way and, even if one allows that certain facts might enjoy an internal relativity to a time, there exists no external form of relativity from which it needs to be distinguished. In this respect, then, the conceptual foundations of fragmentalism are more secure, even if the position itself is more bizarre.

12. First-Personalism

Our second topic, first-personal realism, has received very little attention in the literature, even by those sympathetic to other forms of aspectual realism. But the position is of great interest in itself, since it provides us with one of the most plausible ways of drawing the distinction between the subjective and objective aspects of reality. And it is also of general metaphysical

---

[36] Though I was interested to discover that Sprigge [1992] appears to hold such a view. He writes (p. 12), 'every event must be *present from its own point of view, and as it really is*' and, within our framework, we might take this to mean that each event is present relative to its own reality. But the notion of each event having its own 'perspective' on reality is decidedly odd; and I assume that there must be some more plausible way in which one might accommodate the universal presentness of events within our overall conception of what is real.

interest. For of all the different kinds of aspectual realism, it is the one in which the non-standard position might most plausibly be adopted. Thus consideration of this case provides us with an especially helpful context in which to understand how a non-standard realist position might be developed. Debate on the reality of tense has been very much shaped by the comparison with modality; and this has tended to push the discussion in the direction of the standard position. But a far better comparison, if only philosophers had been willing to make it, would have been with the first-person case.

I am not here concerned to defend first-personal realism (I do not know what to think on the question) but to see how it might plausibly be developed. I shall argue that the most plausible version of the view is one in which the self disappears, either as a subject or as the locus of subjectivity. An interesting aspect of the discussion is that it appears to make sense of positions that might otherwise seem puzzling or obscure. One can very well understand on this view why 'I' might fail to be referential or how one might distinguish between an empirical self, which lies within a world, and a metaphysical self, which lies beyond the world. Indeed, even if one rejects the underlying metaphysics, it is hard not to see the first-personal realist as providing us with an especially helpful way in which to understand the phenomenon of subjectivity.

The first-personal realist believes that there are distinctively first-personal facts. Reality is not exhausted by the 'objective' or impersonal facts but also includes facts that reflect a first-person point of view. But what are the basic first-personal facts? If we let ordinary language be our guide, then they will naturally be taken to be facts that are most directly specified by means of the first-person pronoun. Thus whereas impersonal reality might have been taken to contain the fact that KF is in pain, first-personal reality will be taken to contain the fact that *I* am in pain.

But it is a familiar point that the most basic forms of experience are ones in which there is no representation of the self. The world present itself, to me, as being a certain way; it does not present itself as being that way to me. This suggests that a better account of the first-personal facts would be one in which they would be specified in an egocentric language of the sort considered by Prior [1968*b*, 1969*a*].[37] We should not say 'I am in pain' but 'it is paining', where such a statement is taken to hold 'egocentrically' or relative to a subject in much the same way in which a tensed statement holds tense-logically or relative to a time. The self will be an implicit rather than an explicit subject of the first-personal facts.

---

[37] Though, given how we develop the position, we shall have no need for his strange egocentric counterparts to the modal operators.

This phenomenological consideration is supported by the metaphysical considerations from the previous section (which might even be regarded as providing an underlying explanation of the phenomenology). We there argued that the tense-theoretic realist should not accept the reality of times and similar considerations suggest that the first-personal realist should not accept the reality of selves. For we may ask of selves, just as we asked of times, how they might be capable of figuring in the fundamental facts; and there are difficulties either in the supposition that they occur in first-personal facts or in the supposition that they occur in impersonal facts. In the one case, there will be the embarrassment of having to accept a primitive property of me-ness;[38] and, in the other case, there will be the embarrassment of having to admit both the fact that it is paining and the fact that the subject is in pain. Indeed, in two respects the difficulties are even worse in the first-personal case. For, first, it is somewhat mysterious what this primitive property of me-ness might be (the corresponding difficulty concerning the property of being present is somehow not as acute). And, secondly, once the self is excluded from first-personal facts, it would appear to be possible for it to occur only in straightforward impersonal facts (such as the fact that KF is in pain). There is no counterpart to the 'structural' fact of one time being earlier than another.

However, there is one key respect in which the first-personal realist may wish to qualify his denial of the self. For he may wish to distinguish, in this connection, between two kinds of subject or self. On the one hand, there is the *metaphysical* self. This is the implicit subject of the egocentric facts (such as *it is paining*); and it might be regarded as the locus of subjectivity, since it is relative to such a self that the egocentric facts will obtain. On the other hand, there is the *empirical* self. This is the explicit subject of non-egocentric facts; and it might be regarded as the locus of subjecthood, since it functions as the subject of experience.

I might previously have appeared to argue that if one accepted the metaphysical self then one must reject the empirical self. For the self relative to which the egocentric facts obtain will not explicitly appear as a subject in the basic experiential facts. And so what room might there then be for a self as a subject of experience? But this is to assume that, if the empirical self were to exist, then it would be the same as the metaphysical self; and this is an assumption that may be doubted.

For one may have a conception of the empirical self in which it is a real object in the world, standing in a real relationship to its experiences. This might be true, for example, if one took it to be a Cartesian ego or a living organism or some kind of psychological unity. Thus the empirical self is in

---

[38] This is a primitive *subjective* property of me-ness—its application is relative to a subject. Most philosophers, in considering such a property, have taken it to be objective.

the nature of a substance; and the 'life' of an empirical self is given by the relationship between it and the various experiences it has. The metaphysical self, by contrast, is in the nature of an outlook; and the 'life' of the metaphysical self is simply given by the egocentric facts of which it is the locus.

We might say that the metaphysical self is 'embodied' in a particular empirical self (without meaning to imply that the empirical self is or has a body). There would appear to be nothing intrinsic to the metaphysical self (i.e. to the egocentric facts of which it the locus) that would require it to be embodied in one particular empirical self or even in one particular kind of empirical self. But once we have the empirical link between the two, we can slide between talking about the one in the same manner in which we talk about the other. Thus we might say that the metaphysical self is the subject of certain experiences simply because it embodies an empirical self that has those experiences; and we might say that the empirical self is an outlook on the egocentric facts simply because it embodies a metaphysical self that is an outlook on those facts. This makes it seem hard to distinguish between the two but, in each of these cases, the properties had by one are mediated through its link with the other; it is only because the one is an outlook or a subject that we may correctly say the same of the other.

We therefore see, given the distinction between the metaphysical and the empirical self, that the first-personal realist might acknowledge the reality of the empirical self while denying the reality of the metaphysical self. The empirical self is, as it were, the real world manifestation of the metaphysical self. In this respect, there appears to be a striking difference between the first-personal case of realism and the tense-theoretic and modal cases. If one believed in tensed or in worldly facts, then one could take there to be a time or a world that was the locus of such facts. But once one had done this, it would be decidedly odd to suppose that there might be any counterpart to the empirical self—a time or a world that was a genuine part of reality and yet was somehow distinct from the time or world at which the tensed or worldly facts were taken to obtain. A major part of what gives the issue of first-personal realism its distinctive character is the possibility of drawing a distinction between the metaphysical and the empirical self (just as a major part of what gives the issue of tense-theoretic realism its distinctive character is the passage of time).

Our discussion so far has been neutral between the standard and non-standard forms of realism. But, as I have already remarked, non-standard realism seems especially compelling in the first-person case. It seems quite bizarre to suppose that, from among all the individuals that there are, the subjective world-order is somehow oriented towards me as opposed to anyone else. However, this still leaves open the question of whether we should opt for a relativist or fragmentalist version of the non-standard position.

At this point, our previous considerations in favour of fragmentalism become especially relevant. If we opt for the relativist position, then we must take each subjective reality to be given relative to a metaphysical subject or self. But reality itself contains no metaphysical self. We therefore arrive at the conception of the *pure* metaphysical self—one that stands outside the world and yet is that by which the world (or the subjective world) is given. One can see why philosophers might have been attracted to such a position, given that they wished to give proper recognition to the multiplicity of different subjective viewpoints that, in themselves, were without a point of view.

But the position is barely intelligible; and the mystery of the pure metaphysical self no longer arises once we opt for the fragmentalist position. Über-reality will comprehend all the different subjective facts, both yours and mine, and there will be no more to the metaphysical self than the fragment of subjective reality to which it corresponds. The metaphysical self will dissolve, as it were, into the sea of facts of which it was previously regarded as the source. The sense in which the metaphysical self is an outlook is now especially clear since it will amount to no more than the facts by which the outlook is constituted.

Although this is a kind of 'bundle' theory, it differs from the usual Humean view in three important ways. First, it is a view about the metaphysical self, not the empirical self. Whatever we say about the empirical self—whether we adopt a Humean position, say, or a diametrically opposed Cartesian position—we are still left with the problem of explaining how the metaphysical self relates to the world. Second, there is no special problem of personal identity for the metaphysical self. For once that problem has been solved at the level of the empirical self, its solution may be used to explain how the different fragments of subjective reality cohere. Finally, the usual Humean view does not give any special recognition to the subjective or first-personal character of experience; experiences, for it, are just objects within the objective world-order. But for us, subjectivity lies not merely in the existence of experiences but in the fact that they are *experienced*, where this is an egocentric feature of experience, one whose possession is relative to the self; and the self that is the locus of all subjectivity is to be constructed from these experiential facts rather than from the experiences themselves. This is not so much a radical view about the self as a radical view about the nature of subjectivity.

Our discussion has so far focused on the metaphysics of the first-person. But it has some interesting connections with the semantics of the first-person, which we should now briefly consider. If there are first-personal facts, then presumably they can be stated, i.e. we can state true propositions that can be verified only with their help. And if first-personal facts can be stated, then presumably they can be stated by means of the first-personal

pronoun. For how else might they be stated? The first-personal realist therefore owes us an account of the semantic mechanism by which this possible. What is the semantics of 'I' that enables it to be used in the statement of first-personal facts?

The usual referential accounts are of no help, since what is stated on such an account is an impersonal fact. The sentence 'I am in pain', for example, will be used by the subject P to state that P is in pain (or something of that sort); and this is a proposition that, if true, is verified by the impersonal facts.

A radical response to this difficulty is to deny that 'I' plays a referential role in such sentences as 'I am in pain'. Or, rather, we might allow 'I' to have a referential role in such sentences but deny that it is relevant when the sentence is used, in a distinctively first-personal way, to state a first-personal fact. On this view, to say that 'I am in pain' is essentially to say that it is paining; and 'I' is used to indicate the egocentric character of the resulting proposition rather than to secure a reference to the self.[39] (One might think of 'now' functioning similarly in the sentence 'now it is raining'.)

But such a view is completely at odds with our syntactic and semantic intuitions. Surely, 'I' is used as a subject-term in utterances of 'I am in pain' just as 'KF' is used as a subject-term in utterances of 'KF is in pain'; and surely the sentence 'I am in pain' is used to say something about the self just as the sentence 'KF is in pain' is used to say something about KF. In both cases, the proposition expressed is subject-predicate in form, with one component corresponding to the use of the subject-term 'I' or 'KF' and the other corresponding to the use of the predicate-term 'is in pain'. It would therefore be preferable, if at all possible, to come up with a view that was in accord with these basic intuitions.[40]

If the referential role of 'I' is to be relevant in determining the first-personal content of a first-person sentence and if it is not the referent itself that is relevant to determining that content, then it must be the way the referent is referred to. There must presumably be some 'description' by which the referent is given and which is then partly responsible for the content of the sentence. The description itself must be given in first-personal terms if the content of the sentence is to be first-personal. Thus the usual token-reflexive accounts of 'I' (as with 'the speaker of this utterance') will not do, since they provide only for an impersonal content. Moreover, the description had better be expressible without the use of the first-person pronoun if we are to avoid any question of circularity; and it had better be expressible without appeal to a primitive property of me-ness if we are to avoid any hint of mystery.

---

[39] Anscombe [1975] has advocated a view of this sort for somewhat related reasons.

[40] We might note that Lewis's account of *de se* belief is subject to a similar difficulty. Lewis [1979] takes the content of a *de se* belief to be the property that it attributes to the self. But the self as an explicit object of the belief is then lost.

The problem of finding a descriptive equivalent for 'I' has usually been regarded as insoluble, since any way of identifying its referent would seem to require appeal to either irrelevant or ineffable content. But the problem is readily solved once one adopts the egocentric approach. For we may define 'I' as 'the subject at which the egocentric facts obtain'. It is intrinsic to the egocentric approach that we have an absolute understanding of what it is for an egocentric fact to obtain and a relative understanding of what it is for such a fact to obtain at a subject. The self may then be taken to be the link between the two; it is the subject for which the egocentric facts that obtain at it are the same as the egocentric facts that obtain *simpliciter*. (It is in an analogous way that the tense-theoretic realist might take 'the present' to be 'the time at which the tensed facts obtain' and the actualist might take 'the actual world' to be 'the world at which the worldly facts obtain'.)

It is evident on this view why 'I am in pain' has first-personal content (and hence must be verified with the help of the first-personal facts). For we understand 'P is in pain' to mean that it is paining at P; and so to say that I am in pain is to say that it is paining at the subject at which the egocentric facts obtain. But this then implies that it is paining; and so the fact that it is paining is required to make the sentence true.

It should be noted that the present view involves a difference in the structure of representation at the level of content and at the level of reality. For at the level of reality, there is—or plausibly may be taken to be—no metaphysical self, while at the level of representation there will be. Thus even though propositional contents may involve reference to the self, no appeal to the self need be invoked in explaining how such contents can be true. The self is, in a certain sense, a metaphysical illusion; and this may go some way to explaining why 'I' has been thought to lack any reference.

First-person identities (such as 'I am KF') give rise to special difficulties.[41] For it is clear that such identities have some non-trivial first-personal content. Indeed, in a sense they have maximal first-personal content since, in combination with all the impersonal facts concerning the subject, they will yield all the first-personal facts that hold relative to the subject. The question is to account for this distinctive first-personal content.

There is a way in which this question can be readily answered on our view. For plugging in the analysis of 'I', we see that the content of 'I am KF' is that the subject at which the egocentric facts obtain is KF. We might think of this sentence as attributing the property me-ness to the KF. But this property is no longer some mysterious primitive. It is a defined property, the property of

---

[41] These are also discussed in Nagel ([1983] and [1986], ch. 4) and in Stalnaker ([2003], ch. 14). Although there are several points of contact between Nagel's views and my own, it is not altogether clear to me what exactly he takes the first-personal content of the identity statements to be or how he wishes to deal with the problem raised by Neutrality, a problem that he himself raises.

being the subject at which the egocentric facts obtain, and can be understood in terms of the basic conceptual resources of the egocentric approach.

However, this is to answer the question at an intermediate level of representation, one at which the metaphysical self is still taken to exist. But we want an answer at the level of the reality; we want to know what *facts* might bear upon the sentence's being true. And since, at this level, there is no self, we want an account of the content in which all reference to the self has been made to disappear.

The corresponding problem in the modal case is readily solved. What we must do, in this case, is to specify a world-less content for 'w is actual', where 'w' is a term for a particular world. But each world is necessarily the way it is, and so the content of 'w is actual' can be taken to be that this is how things are. Thus to say the world w is actual is to say P, Q, R, . . . , where P, Q, R, . . . are how things are according to w.

But this answer is not appropriate in the first-personal case (or in the case of tense either). For each individual is not necessarily the way she is (and nor is each time necessarily the way it is). Thus in saying that I am KF, I am not in effect giving a complete account of how things are for KF. Indeed, it would appear that there is no special way things must go for me as opposed to you (or at one time as opposed to another). And so it would appear that there is no special egocentric content that might attach to the claim that I am KF (or to the claim that a particular time is present).

The usual formulation of the semantics for the first person in terms of centred worlds (with each world 'centred' on a given self) obscures this difficulty. For it leads one into thinking of the content of the identity sentence 'I am KF' as the set of worlds centred on KF. But we want to specify the content in terms of how things must go, at the most basic level, for the sentence to be true; and so we must provide some independent account of what it is for a world to be centred on a given self.

One might be tempted at this point to suppose that what gives the apparent identity 'I am KF' its distinctive first-personal content is the fact that the referents of 'I' and 'KF' are not the same: 'I' refers to the metaphysical self, 'KF' to the empirical self; and the sentence as a whole says that the one self is embodied in the other. Since the empirical self is in the real world, its embodiment of the metaphysical self can somehow be an egocentric fact about it.

I do not want to dispute the legitimacy of such a reading. However, the problem we face remains even when the reference of 'I' and 'KF' is taken to be the same. This is very clear from the tense-logical case. For we there have the analogous problem of accounting for the tense-logical content of 't is the present', even though we have no inclination to suppose that the reference of the two terms is not the same.

I can think of only one plausible way in which the present difficulty might be solved. I previously claimed that there was 'no special way things must go

for me as opposed to you (or for one time as opposed to another)'. Now this
might be accepted at the qualitative level (there is no special *kind* of way
things must go for me) though not at the level of particular goings-on. It
might be supposed, for example, that there are token experiences that only
I can have or token events that can occur only at a given time. Each of my
token experiences would then give rise to an egocentric fact, the fact of the
experience being *experienced*, that could obtain only for me and each
current token momentary event would give rise to a tensed fact, the event's
occurring, which could obtain only at the current time. The identity 'I am
KF' could then be seen as restricting egocentric reality to the experience of
experiences that only KF can have; it would require, in effect, that any token
experience that is experienced must be one that is either $e_1, e_2, \ldots$, where
$e_1, e_2, \ldots$ are all my actual and possible token experiences. And, similarly,
the identity of 't is present' could be seen as restricting tensed reality to the
obtaining of token events that can obtain only at t (or within arbitrarily close
neighbourhoods of t). In this way, we could provide each identity with a
distinctive aspectual content, one that had the effect of orienting reality
towards a particular individual or time though without any special commit-
ment to how things were at that individual or time.[42]

   This solution to the problem requires accepting an ontology of token
experiences or token events. If we are to do without the self or without
the present, then the particularity of the self must be reflected through the
particularity of experience and the particularity of the present through
the particularity of what goes on in time. Some tense-theoretic realists (for
example, Prior [2003], ch. 1) have been suspicious of a particularistic
ontology of token events or the like. But if I am right, they must accept
such an ontology if they are to do justice to what we convey in saying that a
particular person is me or that a particular time is present.

13.   The Form of Reality

I conclude with some general comments on the debate between the realists
and anti-realists on the nature of tense and other kinds of aspect. The debate
exhibits an interesting duality, which helps explain why it has been so
difficult to state or to settle.

   We may distinguish in a general way between the form and content of
reality (I do not make this distinction in the usual manner). A reality claim
is a claim to the effect that reality is a certain way. The *form* of such
claims—or, if you like, of reality itself—is constituted by their logical
form, broadly conceived, and the general logical principles by which they

---

[42] I might add that this kind of solution seems especially inappropriate in the modal case. For
we are strongly disinclined to accept a world-bound particularistic ontology.

are governed. Thus we might take a reality claim to consist of the sentential operator 'it is constitutive of reality that' followed by a sentence '...'; and we might take it to be part of the logic of such claims that if it is constitutive of reality that...then...is indeed the case. The *content* of such claims—or, of reality itself—comprises the particular things that are said to belong to reality. The naturalist, for example, will take the content of reality to be entirely constituted by natural facts.

The debate between the realists and anti-realists over tense (or other kinds of aspect) is ostensibly about content. For the realist will claim that reality is tensed, i.e. that it is partly composed of tensed facts, while the anti-realist will deny that this is so. The debate between the standard and non-standard realists is also ostensibly about content. For the standard realist will take only current tensed facts to belong to reality, whereas the non-standard realist will also include tensed facts from the past and the future. Seen from this perspective, then, the issue is about how comprehensive a view we should have of the tensed facts that belong to reality.

But this question of content turns on—or, at least, is closely connected to—a question of form. Take first the debate between the anti-realist and the realist. The realist wishes to claim that certain tensed facts belong to reality. But if he is a standard realist, he will be unable to see how a tensed fact, such as that I am sitting, could sensibly be said to be constitutive of reality unless the relation of constitution was itself taken to be tensed; the fact that I am sitting will be one that *currently* belongs to the constitution of reality. Consider now the debate between the standard and non-standard proponents of realism. The non-standard realist takes reality to be composed not just of tensed facts that currently obtain but also of tensed facts that obtain in the past or the future. Thus reality may contain the fact that I am sitting and also the fact that I am standing. But how can that be unless belonging to reality is a relative matter or unless reality is fragmented? We see in all these cases, then, that the only way to accommodate the new facts is to make successive adjustments in the form of reality; the concept of reality should be taken to be tensed, or relative, or receptive to fragmentation.

In each of these cases, the concept of reality is put under considerable strain. I have already pointed out that the formulation of the primary dispute between the realists and anti-realists requires that we make use of a substantive, metaphysical conception of reality; and the intelligibility of such a conception might be—and has been—doubted. But the secondary disputes between the different types of realist put the concept of reality under even greater strain. Not only do we require a distinction between how things are and how things are in reality, we must also allow that how things are in reality is a tensed matter, or relative to a time, or not even 'of a piece'.

Indeed, we can think of the debate as progressing through more and more problematic conceptions of reality as we move from one position to the next

(and it is curious that the more plausible the view of content, the less plausible is the conception of reality that it appears to require). The anti-realist's conception of reality has three desirable features—it is tenseless, absolute, and cohesive, i.e. resistant to fragmentation. Each of the alternative realist positions requires that we reject one of these features (and yet retain the others). Thus the standard realist abandons the tenseless conception of reality and yet still takes it to be absolute and cohesive; the external relativist abandons the absolute conception of reality and yet still takes it to be tenseless and cohesive; while the fragmentalist abandons the cohesive conception of reality and yet still takes it to be tenseless and absolute. We can think of each of these features as being less and less open to question as we move through the list; and this means that the corresponding conceptions of reality, in which the features are successively dislodged, become progressively more problematic.

We might call a dispute *doctrinal* if both sides to the dispute share a common understanding of the concepts in terms of which their respective positions are to be stated; and we might call a dispute *ideological* if there is no such common understanding. The statement of an opponent's position will be met with incomprehension rather than dissent. We have seen that the dispute between the realists and the anti-realist over tense is, in part, ideological. For the realist may well accept a conception of reality that his opponent will find unintelligible; and similarly for the non-standard realist *vis-à-vis* the standard realist or for the fragmentalist *vis-à-vis* the relativist. I believe that this aspect of the debate helps to explain why it has been so elusive. For an attempt has been made to formulate it as a straightforward doctrinal issue and, depending upon one's position, the formulation has then appeared to be either inadequate or unintelligible.

There are some philosophers for whom this conclusion will be disappointing. They recognize the need to clarify the debate over the reality of tense, but they seek clarification by way of an explanation in terms of concepts antecedently recognized to be unproblematic. But there are different ways in which clarification of a philosophical issue may be achieved. Most straightforwardly, it is achieved by way of a conceptually unproblematic explanation. But sometimes it cannot be achieved in this way; and then it is to be achieved through the recognition that this is so. It must be acknowledged that the issue presents us with a conceptual boundary, which is either to be transgressed or to be accepted as a genuine limit to our thought.[43]

9
Necessity and Non-Existence

Is it possible for Socrates to be a man and yet not exist? This is the kind of question that is likely to strike someone from outside philosophy as preposterous and that may not be taken seriously even by philosophers themselves. But I believe that the answer to this question has profound implications for our understanding of the concepts of existence, identity, and modality and for how these concepts connect to one another and to the world.

It is my central contention that, just as there is a distinction between tensed and tenseless sentences, so there is a distinction between worldly and unworldly sentences, between sentences that depend for their truth upon the worldly circumstances and those that do not. It is in terms of such a distinction that we should assess the possibility that Socrates might be a man and yet not exist, since his non-existence will be a matter of the circumstances while his being a man will not. But once the distinction is drawn, it will be seen to have consequences for a wide range of further questions. It will lead us to distinguish, within the realm of what are normally regarded as necessary truths, between the necessary truths proper, those that hold whatever the circumstances, and the transcendent truths, those that hold regardless of the circumstances. It will also lead us to make an analogous distinction, within the realm of what are normally regarded as necessary existents, between the necessary existents proper, those that exist whatever the circumstances, and the transcendent objects, those that exist regardless of the circumstances. Thus some objects will not properly be in the world just as it has been supposed that some objects are not properly in time. Finally, it will be suggested that the identity of an object—what it *is*—is not, at bottom, a worldly matter; essence will precede existence in the sense that the identity of an object may be fixed by its unworldly features even before any question of its existence or other worldly features is considered.

I begin by drawing the distinction between tensed and tenseless expressions (§1) and the corresponding distinction between worldly and unworldly expressions (§2). This latter distinction leads us to distinguish between different 'grades' of necessity and possibility in a way that has not hitherto being recognized (§3). These distinctions—between the worldly and the

unworldly and between the different grades of modality—are then put to use in the rest of the chapter. I first state a puzzle that appears to suggest that it is indeed possible that Socrates might be a man and yet not exist and consider two standard responses to the puzzle (§4). These responses are found to be wanting in a variety of ways (§5) and I then show how the distinction between the worldly and the unworldly enables us to provide a more adequate response (§6–7). Three other applications are considered; to the status of formal relations, such as *identity* or *membership* (§8); to the status of sorts, such as *man* or *set* (§9); and to the question of existence (§10). It is shown that the simple modal distinctions that one is inclined to draw in these areas should give way to a more subtle delineation of how it is that an object or a truth might engage with the world.

1.   Tenselessness

There is a familiar distinction between tensed and tenseless expressions. A sentence such as 'Socrates is a man' or 'Socrates is self-identical' is tenseless, it cannot properly be said to be true or false at a time, while a sentence such as 'Socrates does not exist' is tensed, it can properly be said to be true or false at a time. Similarly, a predicate such as 'is a man' or 'is self-identical' is tenseless, it cannot properly be said to be true or false of an object at a time, while a predicate such as 'exists' is tensed. There is a corresponding distinction between sempiternal and eternal truths, a sempiternal truth being a tensed sentence that is always true and an eternal truth being a tenseless sentence that is true *simpliciter*.

How one regards this distinction will depend upon whether one is an 'A-theorist' or a 'B-theorist'. The A-theorists favour the primacy of tensed over tenseless talk, while the B-theorists favour the primacy of tenseless over tensed talk. For many B-theorists, at least, there will be a straightforward grammatical basis for the distinction between tensed and tenseless sentences. For they will regard tensed sentences as incomplete expressions, implicitly containing an unfilled argument-place for the time at which they are to be evaluated. The tensed predicate of existence, for example, will be treated as involving two argument-places, one for the object that is taken to exist and the other for the time at which it is taken to exist.

Most A-theorists, however, have been inclined to reject the tensed/tenseless distinction. And, certainly, they are not in a position to make the distinction in the same way as the B-theorists since tensed sentences, for them, are already complete and stand in no need of missing argument-place for time. But there is nothing to prevent the A-theorist from also making the distinction, though in his own way. For he may take the tenses, properly so-called, to be meaningfully applicable only to certain kinds of sentence. Thus what will distinguish 'Socrates is self-identical' from 'Socrates exists' is that

one can properly say that Socrates once existed even though one cannot properly say that Socrates was once self-identical.[1]

Of course, there is room for scepticism over whether this distinction can legitimately be drawn, especially if one is an A-theorist. For it might be argued that even a supposedly tenseless sentence such as 'Socrates is self-identical' *can* properly be said to be true at a time; it is just that it does not matter what the time will be—the sentence will still be true. Thus the only distinction that can legitimately be drawn, on this view, is between those sentences that are always true and those that are sometimes false.

It has to be conceded that once we admit that there is restricted sense of truth-at-a-time, one for which only certain sentences can properly be said to be true or false at a time, then we should also grant that there is an extended sense of truth-at-a-time, one for which any sentence can properly be said to be true or false at a time. For sentences that were previously classified as tenseless can now be taken to be true at every time whatever if they are (tenselessly) true and false at every time whatever if they are (tenselessly) false. And once we operate with this extended notion of truth-at-a-time, the distinction between tenseless eternal truths and tensed sempiternal truths will disappear; the sentences 'Socrates is self-identical' and 'Socrates exists or does not exist', for example, will both be true at any given time.

However, it seems to me that even if we *start off* with the extended notion of truth-at-a-time, we may still draw an intuitive distinction between two different ways in which a sentence may be said to be true at a time. For what we previously took to be a tensed sentence will be true at a time because of how things are at that time while what we previously took to be a tenseless sentence will be true at a time regardless of how things are at the time. It is in this way that 'Socrates is drinking the hemlock' can *properly* be said to be true at a time, while 'Socrates is self-identical' can only *degenerately* be said to be true at a time. If we ask 'how must things be at a given time if the second sentence is to be true?', then nothing sensible can be said in reply; there are no transient states of Socrates' being self-identical, as there are of his drinking the hemlock.

We may likewise redraw the distinction between eternal and sempiternal truths. For an eternal truth will be true *regardless* of the time, i.e. regardless of how things are at the time, while a sempiternal truth will be truth *whatever* the time, i.e. however things are at the time. In the former case, there will be no genuine engagement with how things are at each time while, in the latter case, there will be. If the sun will always shine, then 'the sun will

---

[1] There has been a prejudice against thinking that the meaningful application of a sentential operator may be restricted to certain sentences, even when a similar restriction in the application of predicates has been allowed. But it is no more meaningful to say that I anticipate that the party was a success than it is to say that the number 3 is red.

shine' will be true at any given time because of how things are at that time while the apparently sempiternal truth of 'Socrates is identical' will not depend, in the same way, upon some ongoing feature of the universe.

2. Unworldliness

I now wish to argue that a similar distinction may be drawn in the modal sphere. Just as one may distinguish between tensed and tenseless sentences according to whether they can properly be said to be true or false at a time, so one can draw a distinction between *worldly* and *unworldly* sentences according to whether they can properly be said to be true or false in a world. And just as one may draw a distinction between eternal and sempiternal truths according as to whether they are true regardless of the time or whatever the time, so one can draw a distinction between *transcendental* and *necessary* truths according as to whether they are true regardless of the circumstances or whatever the circumstances.

'Donkeys bray', for example, will be a worldly sentence; its truth will depend upon the circumstances or how things turn out. 'Socrates is self-identical', on other hand, will be an unworldly sentence; its truth will *not* depend upon the circumstances or how things turn out. This same sentence is also a transcendental truth; it will be true regardless of the circumstances or how things turn out.[2] The sentence 'Socrates exists or does not exist', by contrast, is a necessary truth; its truth will indeed depend upon the circumstance (and, in particular, upon whether or not Socrates ends up existing) but in such a way that it is true whatever the circumstances, or however things turn out.

Of course, there is a sense in which the sentence 'Socrates exists or does not exist' *is* true regardless of the circumstances, for we can recognize it to be true on the basis of its logical form alone and without regard to the circumstances. But this is not the sense of 'regardless of the circumstances' that I have in mind. We should imagine ourselves evaluating the truth-value of the sentence on the basis of its logical form. This means that we should first evaluate the disjuncts 'Socrates exists' and 'Socrates does not exist' in turn and see what their truth or falsehood might depend upon; and since they depend for the truth upon the circumstances, then so does the disjunction.

The distinction in the modal case is harder to recognize than in the temporal case; and part of what makes it harder to recognize is that we are accustomed to operating with an inclusive conception of what is necessary and of what it is true in a possible world. We think of the actual world

---

[2] I hasten to add that the notion of *turning out* is a metaphysical, not an epistemological, notion for me. It is a matter of how the world itself turns out, not of how our knowledge of the world turns out.

as the *totality* of facts and so we think of any possible world as being like the actual world in settling the truth-value of every single proposition. Thus every proposition will be evaluable as true or false in a possible world; and the distinction between necessary and transcendental truths will therefore disappear.

We are also far less inclined than in the temporal case to suppose that there is a straightforward grammatical basis for the distinction between the two kinds of expression. If one were a B-theorist about modality and thought that ordinary worldly predicates contained an implicit argument-place for a world just as ordinary tensed predicates have been thought to contain an implicit argument-place for times, then one could, by this means, distinguish between a worldly predicate, which would contain an implicit argument-place for a world, and a transcendental predicate, such as identity, which would not. But this view of modality is not likely to appeal to most philosophers.[3]

All the same, it seems to me that we naturally operate with a more restrictive conception of what is necessary and of what is true in a possible world. A possible world, under this alternative conception, is constituted, not by the totality of facts, or of how things might *be*, but by the totality of circumstances, or of how things might *turn out*. A possible world, as so constituted, will only determine the truth-value of certain propositions (or sentences), those that *turn on* how things turn out. Thus in evaluating the truth-value of such a proposition, in accordance with its logical form, we will eventually be obliged to consider what the worldly circumstances are. The evaluation of an unworldly proposition, by contrast, will involve no such engagement with the world.

A necessary truth will then be a worldly proposition whose truth-value always turns favourably on how things turn out, while a transcendental truth will be a true proposition whose truth-value does not turn on how things turn out. Thus the proposition that Socrates exists or does not exist is a necessary truth, since its truth-value turns on whether or not Socrates exists, which is a matter of how things turn out, and its truth-value, as so determined, is always the Truth. On the other hand, the propositions that Socrates is self-identical or that $2 + 2 = 4$ are not ones whose truth-value turn on how things turn out; and they are therefore transcendental.[4]

We might think of the possible circumstances as being what is subject to variation as we go from one possible world to another; and we might think

---

[3] Curiously, it should not even appeal to a possible worlds realist such as Lewis since, given his counterpart theory, the simple predicates of his favoured language will contain no separate argument-place for a world.

[4] I recall Roger Albritton once expressing doubt over whether $2 + 2 = 4$ could properly be called a necessary truth. I did not understand his remark at the time but perhaps he had something like the above considerations in mind.

of the transcendental facts as constituting the invariable framework within which the variation takes place. Alternatively, we might think of the possible circumstances as being under God's control; it is what he decides upon in deciding to create one possible world rather than another. Thus he can decide whether Socrates exists or not and so he can do something that will guarantee that Socrates exists or does not exist. But there is nothing he can do that will guarantee that Socrates is self-identical or that $2 + 2$ is equal to 4; these are the facts that provide the framework in which he makes the decisions that he does, not the facts yet decided.

3. Grades of Necessity

Let it be granted that there is a restricted notion of truth at a world and a correspondingly restricted notion of necessity. These notions of truth and necessity may then be naturally extended; and we thereby obtain what I call the *extended* and the *superextended* notions of truth and necessity.

Under the first extension, we take every transcendental truth to be true at every possible world. This corresponds to taking a tenseless sentence to be true at every time and gives the *extended* sense of relative truth. Under the second extension, we 'recursively' extend the resulting notion of relative truth to all propositions whatever. Thus given that the proposition that Socrates does not exist is true at the world w in the unextended sense and that the proposition that Socrates is self-identical is true at any world in the extended sense, it will follow that the proposition that Socrates does not exist and is self-identical is true at the world w in this superextended sense. We might talk of truth *in* or truth *at* or truth *of* a world, depending upon whether it is the unextended or extended or superextended sense of relative truth that is in question.[5]

Each notion of relative truth applies to its own characteristic domain of propositions.[6] The first applies to purely worldly propositions, the second to purely worldly and purely unworldly propositions, and the third also applies to 'hybrid' propositions. These are propositions, such as that Socrates does not exist and is self-identical, that are composed of both worldly and unworldly components. Their truth-value in a given world turns partly on the worldly facts and also partly on the transcendental facts.

To each of these notions of relative truth will correspond different notions of necessity and possibility. Thus it will be:

---

[5] Other philosophers have distinguished, though in a rather different way, between various notions of world-relative truth.

[6] I have assumed that each notion does not *meaningfully* apply outside its domain. But another view is that each notion does meaningfully apply outside its domain, but not *correctly*. Some, though not all, of what I say can be adjusted to accommodate this other view; and I have sometimes not been too careful in distinguishing between the two views.

 (i) an unextended possibility that Socrates does not exist and an unextended necessity that Socrates does or does not exist;

 (ii) an extended possibility and also an extended necessity that Socrates is self-identical (though not an unextended possibility or necessity);

(iii) and a superextended possibility that Socrates does not exist and is self-identical and a superextended necessity that Socrates is self-identical if he does not exist (though not an extended possibility or necessity).

It is perhaps worth emphasizing that, even though it is true that Socrates is self-identical, this is not an unextended possibility. Truth does not imply possibility!

There are clear 'scientific' benefits in working with the superextended notions of possibility and necessity; and it is perhaps partly for this reason that they have become standard. One advantage is that they make it much easier to say many of the things that we wish to say. Consider, for example, the possibility that Socrates is the one and only philosopher. This is naturally expressed as: $\Diamond(Ps \ \& \ \forall x(Px \supset s = a))$. But this requires that the worldly predicate 'P' for being a philosopher and the unworldly predicate '=' for identity both occur within the scope of the modal operator and hence it calls for a superextended understanding of the operator. If we have an unextended, or even an extended, understanding of the operator, it is not clear how the possibility is to be expressed.[7]

Another advantage of taking the modalities to be superextended is that it is then much easier to state the general logical principles by which they are governed. It will be a general principle, for example, that whatever is the case is possibly the case ($A \supset \Diamond A$). But A must be suitably restricted if the modality is taken to be unextended or extended; and similarly for other principles.[8]

However, the theoretical benefits of the superextended notions should not blind us to the significance of the more restricted notions. They are more natural and also, I am inclined to think, more basic. The picture I have is one in which the unextended modal facts come first. The extended or superextended modal facts are then derived from these more basic modal facts, with the help of the facts of transcendence and the conventions governing the extended use of the modalities. Thus what ultimately accounts for its being an extended necessity that Socrates is self-identical is that it is a transcendental truth that Socrates is self-identical (just as what ultimately accounts for the fact that it is always the case that Socrates is self-identical is that it is a timeless truth that Socrates is self-identical). Similarly, what

---

[7] There are some interesting technical questions, which I shall not pursue, concerning the conditions under which it is possible to factor a 'mixed' modal claim into a purely modal and a purely transcendental part.

[8] Again, there is some technical interest in developing the logics for the more restricted modal notions.

ultimately accounts for the superextended possibility that Socrates does not exist and is self-identical is that it is an unextended possibility that Socrates does not exist and a transcendental truth that Socrates is self-identical (just as what ultimately accounts for the futurity of Bush's not existing and being self-identical is the 'genuine' futurity of his non-existence and the timelessness of his self-identity).

4.   The Puzzle of Possible Non-Existence

I believe that the distinction between the worldly and the unworldly has some significant consequences for semantics and metaphysics and that it can be used in solving certain puzzles that would otherwise be quite baffling. Its utility in this regard is an important point in its favour. But it is worth pointing out that the distinction has a plausibility that is quite independent of these applications and that it is perfectly possible for someone to reject any or all of the applications and yet still accept the distinction.

I begin with a familiar puzzle concerning possible non-existence, whose difficulty and significance has not, we shall see, been fully appreciated. The puzzle may be stated by means of the following argument:

> (1) It is necessary that Socrates is a man;
> (2) It is possible that Socrates does not exist;
> (3) Therefore it is possible that Socrates is a man and does not exist.[9]

The first premiss of the argument seems clearly to be true. Not only is it intuitively acceptable as stated but there is a simple argument for its truth: for it is of the nature of Socrates to be a man; and from this it appears to follow that necessarily he is a man.

One might, of course, have substantive doubts about whether it is of the nature of Socrates to be a man, but these are not to the point, since the use of the term 'man' was for illustrative purposes only. The argument would work equally well with 'person' in place of 'man' or with 'Felix' and 'cat' in place of 'Socrates' and 'man'. And if one wanted to avoid appeal to any substantive metaphysical commitments, one could use 'self-identical' in place of 'a man'. One would then obtain the following purely logical version of the puzzle:

> (1)′ It is necessary that Socrates is self-identical;
> (2)′ It is possible that Socrates does not exist;
> (3)′ Therefore it is possible that Socrates is self-identical and does not exist.

---

[9] It is more idiomatic to say 'it is possible that Socrates never existed'. The reader who is unhappy with my stylized form of words may use this more idiomatic expression in its place. Strictly, instead of saying 'Socrates is a man and does not exist', I should say 'Socrates is a man and Socrates does not exist'. But I shall allow myself some obvious latitude in formulation.

For the most part, we shall confine our attention to the original formulation of the argument in terms of 'man' but it is worth keeping in mind the possibility of these other formulations.

The second premiss of the argument also seems clearly to be true. There may be items—numbers or propositions or the like—that necessarily exist. But Socrates (and Felix) are not among them. Indeed, it is surely possible that no men (or cats) exist; and from this it follows, given the appropriate version of the first premiss, that it is possible that Socrates (or Felix) does not exist.

The conclusion also appears to follow from the two premisses by impeccable modal reasoning. Indeed, the form of the argument is: $\Box p$, $\Diamond q$, $\therefore \Diamond(p \,\&\, q)$. And this is a form of inference that is sanctioned by the weakest systems of modal logic. However, the conclusion of the argument appears to be false. For how can Socrates be a man without existing? Unless *there is* an object to be a man then how can that object be a man?

Thus the argument appears to be sound, with true premisses and valid reasoning; and yet its conclusion is unacceptable. What are we to say? What mistake are we making? And what leads us to make the mistake?

There are two responses to the puzzle that are implicit in the literature and may be attractive to many readers. The first derives from the conceptual framework of Prior's system Q (Prior [1957], ch. 5). Within this framework, any proposition concerning an object is taken to be neither true nor false in any world in which the object does not exist. Since a proposition can lack a truth-value in a world, the modalities will be ambiguous: 'necessarily' can mean either 'always true' (the strong reading) or 'never false' (the weak reading); and, correlatively, 'possibly' can mean either 'sometimes true' (the strong reading) or 'sometimes not false' (the weak reading). Here, of course, the quantifiers 'always', 'never', 'sometimes' should be taken to range over possible worlds.

Our mistake, according to this proposal, is this. We accept the first and second premisses under the weak reading of 'necessity' and 'possibly' (since neither is correct under the strong reading). We accept, in other words, that it is never false (though *not* always true) that Socrates is a man and that it is sometimes not false (though *not* sometimes true) that Socrates does not exist. We are then in a position validly to infer the conclusion under the weak reading, i.e. to infer that it is sometimes not false that Socrates is a man and does not exist. But we mistakenly infer the conclusion under the *strong* reading, for which it is false, i.e. we mistakenly infer that it is sometimes true that Socrates is a man and does not exist; and what leads us to make the mistake is a confusion between the two readings.

The other response to the puzzle is in keeping with the standard semantics for modal logic. Every proposition is taken to be either true or false in a world, regardless of whether it concerns a non-existent, and, in particular,

the proposition that a given object does not exist will be true in a world if that object does not exist in the world and otherwise will be false. There is therefore no basis on this approach for a Q-ish distinction between the weak and strong reading of the modalities. What is recommended in its place is a related distinction between a qualified and unqualified reading. Under the 'unqualified' reading, a proposition concerning certain objects will be necessary if it is true in every world and, correlatively, the proposition will be possible if it true in some possible world. Under the 'qualified' reading, a proposition concerning certain objects will be necessary if it is true in any world in which those objects exist and, correlatively, the proposition will be possible if it is true in some world in which those objects exist.

Our mistake, according to this other proposal, is this. We accept the first premiss under the qualified reading of necessity (since it is not correct under the unqualified reading). In other words, we accept that it is necessary that Socrates is a man *if* he exists (though not without this proviso). We accept the second premiss under the unqualified reading (since it is not correct under the qualified reading). In other words, we accept that it is possible that Socrates does not exist (not that it is possible that he does not exist under the proviso that he exists). From this we then mistakenly infer the conclusion that possibly Socrates is a man and does not exist under the unqualified reading.

What leads us to make the mistake is an equivocation between the qualified and unqualified readings. Our acceptance of the first and second premisses requires us to adopt non-correlative readings of 'necessity' and 'possibility'—one qualified, the other unqualified. We then mistakenly infer the conclusion under the same unqualified reading of 'possibility' as in the second premiss, failing to recognize that the validity of the inference calls for a correlative unqualified reading of 'necessity' in the first premiss.

5. Problems with the Two Responses

These two proposals suffer from many problems, some common to both and some peculiar to one or the other. Since many philosophers, I suspect, will have been persuaded by one or the other of the proposals, it is worth spelling out in detail where the difficulties lie.

Here, for convenience, is a summary of how each proposal interprets the fallacious reasoning of the puzzle:

Priorian	*Standard*
$\boxed{w}$ p	$\boxed{q}$ p
$\diamondsuit_{w}$ q	$\diamondsuit_{u}$ q
$\therefore$ $\diamondsuit_{s}$ (p & q)	$\therefore$ $\diamondsuit_{u}$ (p & q)

There are five broad sets of difficulties in all:

(i) According to the Priorian, our mistake is to shift from a weak reading of the modalities in the premisses to a strong reading in the conclusion. But this is most implausible. For given that we must already adopt a weak reading of 'possibly' in order to accept the second premiss and given that the conclusion can be validly inferred under that reading, then why, in inferring the conclusion, should we switch to a reading under which it is not valid? Why gratuitously commit the fallacy of equivocation?

The standard response does not suffer from this difficulty, since the reading of 'possibility' in the second premiss and in the conclusion is the same.

(ii)(a) Both responses fail correctly to identify the sense of 'necessity' in which we are willing to accept the first premiss that it is necessary that Socrates is a man. For both take the relevant sense of 'necessity' (be it weak or qualified) to be one in which we are also willing to accept that necessarily Socrates exists. Now it is unclear that there is a sense of 'necessity' in which it is necessary that Socrates exists (under normal metaphysical views). But even if there is, it is hard to believe that someone who accepts the necessity that Socrates is a man is thereby committed to accepting the necessity that Socrates exists (assuming no shift in the sense of 'necessity'). The point can be made more dramatically still with 'existent man' in place of 'exists'. For someone who accepts that it is necessary that Socrates is a man is not thereby committed to its being necessary that he is an existent man.

Moreover, the standard response fails to provide an adequate reason for supposing that the first premiss is not to be accepted under the unqualified reading of 'necessity'. There is no corresponding difficulty for the Priorian, since he has a general horror of non-existence. The mere fact that an object fails to exist is taken to be a general barrier to truth; and so it will be impossible, given that Socrates fails to exist in some possible world, that he should enter into any strongly necessary truths. But the standard respondent has only a selective horror of non-existence; it is perfectly possible, given that Socrates fails to exist in some possible world, that he might enter into some unqualifiedly necessary truths. It will be unqualified necessity, for example, that Socrates exists or does not exist or that possibly Socrates is a man. But this means that if we are to reject the unqualified necessity that Socrates is a man, then it must be because of some special feature of the term 'man'.

If 'man' were a basic predicate (or *man* a basic property), then one can see how this might be taken to constitute a general reason for thinking that something could not be a man without existing. But on the face of it, 'man' is a defined term. There should be an account—at least, in principle—of what it is for something to be a man. But then how can we be sure that the

definition will result in a predicate that is existence-entailing? Indeed, I shall later argue (§9) that the proper form of definition will result in a definition that is *not* existence-entailing; and, at the very least, this makes it far from obvious that, in accepting the first premiss, we must somehow be committing ourselves to the qualified sense of 'necessity'.[10]

(ii)(b) This conclusion is further confirmed by considerations of nature or essence. It is of the nature of Socrates to be a man; this is what Socrates *is*. From this it appears to follow, for some suitable notion of necessity, that it is necessary that Socrates is a man. This inference is an instance of a general inferential connection between claims of essence and claims of necessity. For from:

It is of the nature of a to F (or a essentially Fs);

we may infer:

Necessarily, a Fs.

The question we now face is whether we are justified in inferring the modal conclusion in the stronger unqualified sense of necessity or in only the weaker qualified sense. In the latter case, all that we would be entitled to infer from it being of the nature of a to F is that it is an (unqualified) necessity that a Fs *if* a exists. But suppose that someone believes in God and takes it be of the nature of God to exist (the reasons he might have for believing in God's existence are not here in question). Then surely he is entitled to infer that it is an (unqualified) necessity that God exists. But all that we are entitled to infer under the weak reading of the connection is that it is (unqualified) necessity that God exists if he exists! This suggests that any plausible account of the connection between essence and necessity should make it strong.[11] Another difficulty with the weak reading is that it would also entitle us to infer the necessity that Socrates is an existent man from the fact that Socrates is essentially a man (this is related to our previous concern over the relevant sense of 'necessity'). Yet surely we are not entitled to make any such inference.

(iii) The Priorian response does not correctly identify the sense of 'possibility' in which we are willing to accept the second premiss that it is possible that Socrates does not exist. For it takes this to be a sense in which we are also willing to accept that it is possible that Socrates exists and does not exist. Now I very much doubt that there is a sense of 'possibility' in which it is possible that Socrates exists and does not exist. But even if there is, it is hard to believe that someone who accepts the possibility that Socrates does

---

[10] Similarly for the case of self-identity. It is very plausibly taken to be an unqualified necessity that Socrates is identical to Socrates. Indeed, this is taken for granted in the standard axiomatizations of modal logic.

[11] The connection between essence and modality is further discussed in Fine [1994].

not exist is thereby committed to accepting the possibility that Socrates exists and does not exist (assuming no shift in the sense of 'possibility').

We might call a possibility *non-existential* if it is the possibility that an object is a certain way and does not exist (something we might formalize as: $\diamondsuit(\varphi(a)\& \sim Ea)$, where a is a name for the object, $\varphi( )$ signifies the way in question, and E is the predicate for existence). The point is that we seem able to assert the second premiss under a notion of possibility that is capable of discriminating between different non-existential possibilities for a given object—that allows the possibility that it does not exist, say, but not the possibility that it exists and does not exist. But the conceptual framework of Q is incapable of recognizing any such notion of possibility.[12] The standard response does not suffer from this defect. For, given its conformity with the standard semantics for modal logic, it may allow the first possibility while disallowing the second.

(iv) Neither response correctly identifies the sense of 'possibility' in which we wish to reject the conclusion that it is possible that Socrates is a man and does not exist. This is evident in the case of Q, since we do not also want to reject the possibility, in this very sense, that Socrates does not exist. But it might appear to be an astonishing claim to make in the case of the standard response, since is it not just part of the underlying metaphysical view that Socrates cannot be a man without existing? But let us fix on the sense in which it is taken not to be possible that Socrates is a man and does not exist (this is presumably the unqualified sense though it makes no difference to the objection if we take it to be the qualified sense). If it is not possible in this sense that Socrates is a man and does not exist then it is also not possible in the same sense that Socrates is a non-existent man. But our attitude towards these two possibilities is quite different, despite their having logically equivalent contents. For we are somewhat disinclined to accept the one, the possibility that Socrates is a man and does not exist, and yet strongly inclined to accept the other, the possibility that Socrates is a non-existent man, for this seems to amount to no more than the possibility that Socrates does not exist and his being correctly classifiable as a man. (It is important to understand this latter possibility in the right way. Imagine someone listing the kinds of things that might not exist. Thus he might say that it is possible that Fido is a non-existent dog and possible that Socrates is a non-existent man). If this is right, then we have a sweeping objection to any normal reading of the modality, since any such reading will fail to distinguish between possibilities with logically equivalent contents.

(v) There is another, more general defect, with both responses. Let us say that we have a *dual* reading of 'necessity' and 'possibility' if 'necessarily' is so

---

[12] The conceptual framework of Q is further discussed in Fine [1977*a*, 1985] (Chs. 6 and 7 above).

understood as to be equivalent to 'not possibly not' (and, consequently, 'possibly' is so understood as to be equivalent to 'not necessarily not'). The dual reading of the modalities is commonly taken for granted in the formulation of modal logic; and it is natural to suppose that 'necessity' and 'possibility' will receive a dual reading in our understanding of the first and second premises.[13] In any case, we can, in effect, make the dual reading explicit by adopting a mono-modal presentation of the puzzle:

(1)* It is not possible that Socrates is not a man;
(2)* It is possible that Socrates does not exist;
(3)* Therefore it is possible that Socrates is a man and does not exist.

The first premiss, under this new formulation, is perhaps even more plausible than it was before.

However, it is not on the cards under either response to accept the two premises without equivocating on the meaning of 'possibility'. For the first premiss will be incorrect under the weak or the unqualified reading (since it is both weakly and unqualifiedly possible that Socrates is not a man), while the second premiss will be incorrect under the strong or the qualified reading (since it is neither strongly nor qualifiedly possible that Socrates does not exist). Yet surely we can accept both premises without any equivocation in the sense of 'possible'.

It might be thought that we could explain our acceptance of both premisses under a fixed meaning of 'possible' by appealing to the distinction between the 'internal' and 'external' use of negation. To say that Socrates is not a man is not simply to say that it is not the case that Socrates is a man; it is also to say (or to imply) that Socrates exists. It will then no longer be incorrect to accept the first premiss under an unqualified reading of 'possible' since it will not be possible both for Socrates not to be a man and to exist (this line of response, we might note, is not available to the Priorian, since the internal/external distinction is not one he can sensibly grant).

I am inclined to think that the first premiss is equally acceptable when the negation is taken to be external. For surely we do not wish to admit the possibility that it is not the case that Socrates is man. But we might also note that if the negation is taken to be internal in the first premiss, then it should also be taken to be internal in the second premiss, in which case the second premiss will no longer be acceptable. In order to make both premises acceptable we must posit an equivocation either in the use of 'possible' or in the use of 'negation'; and yet, in accepting both premises, it is hard to have any sense that we are guilty of any such equivocation.[14]

---

[13] 'Duals' are not the same as 'correlates' for the Priorian modalities: the dual of strong necessity is weak possibility, the dual of strong possibility weak necessity.

[14] Tony Anderson has suggested to me a response that is a kind of hybrid of the Priorian and standard responses: the existence-predicate is taken to be false of any object in a world in which

6.   The Worldly Response

I hope I have done something to convey just how puzzling the puzzle is. It is not just that the two most obvious ways of solving it both fail. We are also at a loss to account for all the various things we want to say in connection with the puzzle. We must somehow identify the peculiar senses of 'necessity' and 'possibility' in which we wish to endorse the premises of the argument and reject its conclusion and we must do so in such a way as to allow for: the unqualified endorsement of the necessity that Socrates is man; the possibility of a dual reading of the modalities; the sensitivity of our modal intuitions to the substitution of logical equivalents; a proper account of the connection between essence and modality; and an explanation of the special role played by such predicates as 'man' and 'self-identity'. The puzzle is a Pandora's box!

I would now like to show how we might employ the worldly/unwordly distinction in solving the puzzle and shedding light on these issues. It is worth noting that an analogous solution may be developed for the tense-logical version of the puzzle, a typical formulation of which might go as follows:

(1)#   It is always the case that Socrates is a man (or: it is never the case that Socrates is not a man);
(2)#   It is sometimes the case that Socrates does not exist;
(3)#   Therefore it is sometimes the case that Socrates is a man and does not exist.

Just as the worldly/unworldly distinction might be employed in solving the modal version of the puzzle, the tensed/tenseless distinction might be employed in solving the tense-logical version. Indeed, in some respects, the solution to the tense-logical version is more plausible and can be used to motivate a corresponding solution to the modal version. But rather than give a parallel treatment of both cases, I shall, for the most part, focus on the modal case and leave it to the reader to fill in the details of the tense-logical case for himself.

Two key hypotheses will be involved in our 'worldly' solution to the puzzle (though there are others). The first is that we are naturally inclined to use the modalities in an unextended or extended sense, though not in a superextended sense. One might attempt to justify the hypothesis along the the following lines, though this is not essential to its acceptance. First, we are

---

it does not exist while the predicate 'man' is taken to be neither true nor false of such an object. The modalities are then divided into weak and strong as under the Priorian approach. This response is able to avoid most of the difficulties of the other two responses. However, it will still fall afoul of a version of the second objection: for the sense of necessity in which it is necessary that Socrates is a man will be a sense in which it is necessary that Socrates exists; and all that will in general follow from its being of the nature of x to F is the weak necessity that x is F. It would also be desirable to be able to provide some deeper motivation for the differential treatment of the two predicates.

most naturally inclined to use the modalities in the unextended sense (just as we are most naturally inclined to use the tense-logical operators so as to apply only to tensed sentences). The extension of our unextended understanding of the modalities to the extended sense is relatively modest and one that we are naturally inclined to make (just as we are willing to go along with the idea that a tenseless truth, such as $2 + 2 = 4$, is always true). However, the extension to the superextended sense is artificial and one that we are disinclined to make (just as we are somewhat uncomfortable with the idea that it will be true that dawn breaks and $2 + 2 = 4$). In making the transition to the extended sense of the modalities, we must admit a new, degenerate, way in which a sentence may be true at a world; it may be true at a world simply because it is true regardless of how things are in the world. And in making the further extension to the superextended sense, we must admit a new, hybrid, way in which a sentence may be true at a world; it may be true at a world because of the (possibly different) ways in which its component sentences are true or false at the world. But for reasons hard to articulate, the first extension is much more of a 'stretch' than the second. The first simply involves the admission of a straightforward and independent way of being true at a world, while the second requires the recognition of an anomalous amalgam of the other two ways of being true at a world.[15]

The second hypothesis is that 'exists' is a worldly predicate while such predicates as 'man' or 'identity' are unworldly. I take it to be evident, given that the distinction between the worldly and the unworldly is viable, that 'exists' should be classified as a worldly predicate. The status of 'man' and 'identity' is less clear. But it is certainly plausible that they are both unworldly—whether an object is a man or self-identical is not something that appears to turn upon how things turn out for the object; and this is a view that I shall later try to defend.

Let us now see how these two hypotheses can be put to work in developing a solution to the puzzle.[16] We should say, in the first place, what accounts for the particular judgements that we are inclined to make. When it comes to evaluating the first premiss, that it is necessary that Socrates is a man, we implicitly treat 'man' as an unworldly predicate (in conformity with the first hypothesis). We then accommodate the sense of 'necessity' to the sense of 'man' by taking it in an extended sense (by the first hypothesis) and are thereby led to accept the first premiss. Given that 'man' *is* an unworldly predicate, our acceptance of the first premiss will indeed be correct. (It is for

---

[15] These remarks are merely suggestive. It would be desirable to have a more systematic framework within which to discuss the question of linguistic accommodation.

[16] Sometimes what matters is not the hypothesis itself but our implicit acceptance of its truth. But I assume that the truth of the various hypotheses makes our implicit acceptance of them all the more plausible!

a similar reason that we might accept that it is always true that Socrates is a man, in the tense-logical version of the puzzle).

When it comes to evaluating the second premiss, that it is possible that Socrates does not exist, we implicitly treat 'exists' as a worldly predicate (in conformity with the first hypothesis). We then take 'possibility' in the unextended sense (by the first hypothesis) and are thereby led to accept the second premiss. Given that 'exists' *is* a worldly predicate, the acceptance of the second premiss is indeed correct. (And similarly with 'it is sometimes the case that Socrates does not exist'.)

The evaluation of the conclusion, that it is possible that Socrates is a man and does not exist, is more complicated, since it involves a conflict that must somehow be resolved. On the one hand, we have some inclination to take 'man' to be unworldly and 'exists' to be worldly. But this would turn the proposition that Socrates is a man and does not exist into a hybrid and so would call for a superextended sense of 'possibility'. On the other hand, we are disinclined to use 'possibility' in a superextended sense (by the first hypothesis). Thus something must 'give' if we are to evaluate the conclusion. But what?

One option is to take 'man' to be worldly; another is to take 'exists' to be unworldly; and the third is to give 'possibility' a superextended sense. Of these three, it appears to be the first that does the least violence to our inclinations. The status of 'exists' as worldly seems non-negotiable in a way in which the status of 'man' as unworldly is not. And the slide to a worldly understanding of 'man' seems much less precipitous than the slide to a super-extended understanding of 'possibility'.

Indeed, there is a predicate, 'existent man', that is perhaps the closest worldly neighbour to the predicate 'man'; and in taking 'man' to be worldly, we naturally take it to mean *existent man*. Thus, under this option, we take the possibility that Socrates is a man and does not exist to consist in the existence of a possible worldly circumstance in which Socrates is a man and does not exist; and given that, in such a circumstance, Socrates can be a man only by existing, it is evident that the possibility should be rejected.

It might be wondered why the substitution of 'existent man' for 'man' serves to resolve the conflict. For have we not substituted one hybrid expression for another within the scope of the possibility operator? The answer turns upon a subtle distinction in the semantics for worldly and unworldly expressions. Given an unworldly predicate, such as 'man', and a worldly predicate, such as 'existent', we may use the worldly predicate to qualify the unworldly predicate, as in 'existent man', and thereby form a *worldly* predicate, that is true in each world of those objects that are in the world-free extension of the first predicate and in the world-bound extension of the second predicate. It is in the same way, if we are working within a tense-logical framework, that we may use a tensed predicate, such as 'existent', to

qualify a tenseless predicate, such as 'man', to form a tensed predicate, 'existent man', that is true of those objects that belong to the time-free extension of the one predicate and the time-bound extension of the other. 'Qualification' is not exactly conjunction, since the unworldly (or tenseless) predicate gains an argument, as it were, upon being combined with the worldly (or tensed) predicate. Thus the qualified predicate 'λx(x is an existent man)' is worldly while the conjunctive predicate 'λx(x is existent and x is a man)' is hybrid; and it is because of this difference that the modal operator need no longer be taken to govern a hybrid expression.

I should emphasize that, in giving the above explanations, we are merely attempting to account for a *tendency* in our judgements. This is compatible with there being tendencies in other directions. Thus we have some tendency not to accept the first premiss and to refrain from evaluating the conclusion; and this, I take it, is because we are inclined to use 'necessity' and 'possibility' in an unextended sense. We also have some tendency to accept the conclusion; and this, I take it, is because we have some tendency to use 'possibility' in a superextended sense. However, these tendencies are overridden—at least when it comes to our ordinary untutored judgements—by other, more compelling, tendencies.

On the present analysis, there is no mistake in the particular judgements that we make, given that we understand them in the way that we do. But we are mistaken in thinking that the conclusion, as we understand it, follows from the premisses; and we are also guilty of the more general error of using the term 'man' in an inconsistent manner in the context of the argument. For in evaluating the first premiss, we correctly treat it as unworldly while, in evaluating the conclusion, we incorrectly treat it as worldly.

There are two ways in which consistency in our use of terms might be restored. The first is to use the modalities in an extended sense throughout the argument (and taking 'man' to be unworldly). In this case, the proper response is to refuse to evaluate the conclusion. It should be regarded as anomalous in much the same way as 'it is possible both that Socrates is a man and he did not drink the hemlock' is anomalous. The second is to use the modalities in a superextended sense throughout the argument (again taking 'man' to be unworldly). This is the sense of the modalities that I previously recommended for 'scientific' purposes. In this case, the conclusion does indeed follow from the premisses and should therefore be accepted.

But any puzzlement we might have had in accepting the conclusion should disappear. For if I am right, the felt incompatibility between the non-existence of Socrates and his being a man arises from our implicitly assuming that his being a man is a worldly matter. There is then a genuine difficulty in seeing how he could both be a man and not exist. But, on a correct view, his being a man is an unworldly matter. It is something that holds 'off-stage',

regardless of how things turn out; and so, in particular, it is something that holds regardless of whether or not he exists. Thus it is not that he is possibly a man *despite* his not existing. His existence or non-existence is simply irrelevant to his possible status as a man; and all that the possibility of his being a man and not existing comes down to is the genuine possibility of his not existing and the unworldly, or circumstance-indifferent, fact that he is a man.

7. Advantages of Being Worldly

It will be helpful to compare our response to the puzzle with the two others and to show how it is able to satisfy the various desiderata upon which they foundered.

The underlying structure of the two classes of response is rather different. The two previous responses involved a double ambiguity in the modalities whereas ours involves a triple ambiguity. Our response also involves a possible ambiguity in or misconstrual of the predicate 'man', whereas theirs treats 'man' as unambiguous. More significantly, they take the predicate 'man' to be existence-entailing, it is only true in a possible world of objects that exist in that world. Our view is more subtle. Depending upon which notion of relative truth is in question, we either say that the predicate has no world-relative application or that it is true of men in each possible world independently of whether they exist.[17]

The underlying diagnosis of the error is also very different. The two previous responses take there to be a confusion over the sense of the modalities. But we also take there to be a confusion over the sense of 'man'; for, in evaluating the conclusion, we mistakenly take the predicate to be existence-entailing. Thus the mistake that we take to be at the heart of the puzzle is one that the other two responses reproduce!

Let us now consider the various desiderata that were listed at the beginning of the previous section.

(i) We wish to allow for the unqualified endorsement of the necessity that Socrates is man (so that it is not also a necessity, in the same sense, that Socrates exists). But the extended sense of necessity is just such a sense.

(ii) We wish to allow for a dual reading of the modalities under which both premises will be true. Again, the extended modalities will do the trick.

(iii) We wish to account for the sensitivity of our modal intuitions to the substitution of logical equivalents. Why, in particular, do we baulk at the possibility that Socrates is a man and does not exist but not at the possibility that Socrates is a non-existent man?

---

[17] I owe to Karen Bennett the thought that the ambiguity may lie not in the term 'man' but in the copula 'is'.

The answer turns on the difference in the semantics of the clauses 'x is a non-existent man' and 'x is non-existent and x is a man'. For the former is worldly, the latter hybrid; and so the former will give rise to an unextended possibility for Socrates, while the latter will not.

(iv) We should provide a proper account of the connection between essence and modality and we should explain, in particular, how it might follow from the nature of God that he necessarily exists. We can now accept the connection between essence and necessity in an unqualified form: given that a essentially Fs it follows that it is necessary that a Fs (where 'necessity' takes either an unextended, extended, or superextended sense, depending upon the predicate F in question). Thus when F is 'existence' or existence-entailing, it will follow that x necessarily exists.

(v) We should explain the special role played by such predicates as 'man' and 'self-identity' in the statement of the puzzle. This immediately falls out of our account since what is distinctive about these predicates, for the purposes of the puzzle, is their unworldly status.

Let us note, in concluding this section, that there is another familiar puzzle that can be resolved using the worldly/unworldly distinction.[18] Consider the following three assumptions:

(i) necessarily, if Socrates does not exist then the proposition that Socrates does not exist is true;

(ii) necessarily, if the proposition that Socrates does not exist is true then the proposition that Socrates does not exist exists;

(iii) necessarily, if the proposition that Socrates does not exist exists then Socrates exists.

From these three assumptions, it follows that necessarily Socrates exists (since it follows that necessarily if Socrates does not exist then Socrates exists). But surely it is possible that Socrates does not exist!

We may remove the air of puzzlement by noting that the claim that the proposition that S is true has a hybrid status. It is a matter of the proposition expressing that S, which is an unworldly matter, and it is also a matter of S, which is a worldly matter (as long as S is a worldly matter). We therefore see how it might be possible for the proposition that Socrates does not exist to be true even though the proposition does not exist. For this possibility simply turns on the proposition that Socrates does not exist expressing that Socrates does not exist, which is an unworldly matter holding regardless of the circumstances, and on the possibility that Socrates does not exist.

---

[18] This puzzle is discussed in §7 of Fine [1977] (Ch. 4 above). A more recent discussion, along somewhat different lines, is to be found in Williamson [2000].

8. Formal Transcendence

Once we recognize the distinction between tensed and tenseless sentences, we may make a corresponding ontological distinction between tensed and tenseless facts. A tensed fact will be one that can be stated by means of a tensed sentence while a tenseless fact will be one that can only be stated by means of a tenseless sentence. These facts, in their turn, will belong to two 'realms' of reality, one tensed or temporal and the other tenseless or timeless. Thus the fact that Socrates drank the hemlock will presumably belong to the temporal realm while the fact that $2 + 2 = 4$ will presumably belong to the timeless realm.

It should be observed that which facts are taken to belong to these two realms will depend upon whether one is A- or B-theorist about tense. If one is an A-theorist then the sentence 'Socrates drank the hemlock' will state a fact that in itself is tensed; but if one is a B-theorist, then it will state a fact to the effect that Socrates drank the hemlock at a particular time t, one that might also be stated by a tenseless sentence. Thus the tensed facts will be intrinsically tensed in the one case and merely a species of temporally 'located' fact in the other.[19]

Just as the linguistic distinction between tensed and tenseless sentences leads to an ontological distinction between tensed and tenseless facts, so the linguistic distinction between worldly and unworldly sentences leads to an ontological distinction between worldly and unworldly facts. The worldly facts are those that can be stated by worldly sentences while the unworldly facts are those that can be stated only by unworldly sentences. And these facts, in their turn, will belong to two realms of reality, one worldly or circumstantial and the other unworldly or transcendental. Thus the fact that Socrates is a philosopher is a worldly fact while the fact that he is self-identical is an unworldly fact. We adopt the A-theorist's perspective on modality; and so the distinction between the worldly and the unworldly can be taken to be intrinsic to the facts themselves.

There is a sense, which is hard to articulate, in which these various kinds of facts belong to different levels or orders of reality. At the bottom are the tensed or temporal facts. These are subject to the vicissitudes of time and hence of the world. Then come the timeless though worldly facts. These are subject to the vicissitudes of the world, though not of time. And finally, at the top, are the transcendental facts, subject to the vicissitudes neither of the world nor of time.

---

[19] There may perhaps be some facts which are stated by tensed sentences but which the B-theorist would not take to be temporal in the intended sense. Thus he might take the true sentence 't is present' to state the fact that the time t is identical to t. But this would presumably be classified as a timeless rather than as a temporal fact.

Once given the distinction between these different realms, one naturally wishes to know how the outlying realms might be constituted. What kinds of fact belong to the transcendental, as opposed to the empirical, realm (or to the timeless, as opposed to the temporal, realm)? What is the transcendental (or the timeless) constitution of the world?

I shall argue, in the remainder of the chapter, that the transcendental constitution is much more varied and interesting than one might have thought and that we should recognize transcendence both at the level of ideology, of how things are to be characterized, and at the level of ontology, of what there is. The opposed view is that of the worldly philosopher. He grants that there is an intelligible distinction between the worldly and the unworldly. Thus he is not sceptical of the distinction itself. But he denies that there are any unworldly facts; for him the world is entirely mundane.

Now there is a more and a less interesting ground upon which the worldly view might be challenged. Both sides to the debate can agree that there are modal facts—the fact, for example, that it is possible that Socrates does not exist or the fact that it is necessary that Socrates exists if he exists. And so the worldly view might be challenged on the ground that all such modal facts are transcendental.

The question of whether modal facts are transcendental is not without interest;[20] and I am inclined to think that they are and that the worldly view, for this reason alone, should be given up. But there is another, more interesting, ground upon which the view might be disputed. For it might be claimed that there are certain *basic* transcendental properties (or relations), ones whose application always gives rise to transcendental facts. Thus transcendence will then exist at the ground level, as it were, and will not simply arise from the possibility or necessity that certain circumstantial facts obtain.

There is a view that many philosophers (including my former self) have found attractive and that effectively rules out there being any basic transcendental properties. It is the ontological correlate of the 'Falsehood Principle' of Fine [1981b] and something like it also goes under the name 'Possible Actualism'. According to this view, any basic property will be existence-entailing, i.e. it will fail to be exemplified by an object that does not exist in a given world.[21] But if a property is transcendental, it will be

---

[20] The perceived transcendence of modal facts may help explain why there was once so much resistance to treating the iterated modalities as intelligible. For if 'necessity', say, is taken in an unextended sense, then it will no more have meaningful application to 'necessarily P' than will 'always' have meaningful application to 'always P'.

[21] I am aware that some philosophers will not be altogether happy with the notion of a basic property. A more cautious formulation of the view is that all properties either themselves are existence-entailing or can be defined in terms of properties that are existence-entailing. This is compatible with there being no basic properties or with there being some latitude as to what

exemplified by an object regardless of whether the object exists; and so no basic property could be transcendental (unless it happened to be exemplified only by necessary existents).

The thinking that lies behind this view is that a possible world is simply constituted by what one might call its 'inner' core, i.e. by the objects that exist within it and the properties and relations that they exemplify. Non-existents, that lie on the other side of the worldly divide, can only be understood negatively as not being among the objects by which the world is given and as not having the properties or entering into the relations by which the world is characterized. They belong to a predicational void, as it were, in which nothing of a positive character can be said about how they are.

However, once we are mindful of the distinction between worldly and unworldly facts, we see that the motivation for the view has been misconceived. A possible world, in the sense of *how things turn out*, will indeed be constituted by what exists and by the properties and relations that hold of the objects that exist. But transcendental properties will be exemplified by objects regardless of how things turn out and so should be taken to be exemplified by objects regardless of whether or not they exist. A world *qua* totality of circumstances will conform to the view, but a world *qua* totality of facts will not. Thus the view, properly stated, is that no basic *worldly* property or relation will be exemplified by non-existents; and this still leaves open the question of whether there are any basic *transcendental* properties or relations.

Now if one is going to argue that there are such properties or relations, then the case of identity would appear to be as strong as any. For surely the relation is basic; and surely, one is inclined to think, the identity of an object with itself is something that holds regardless of how things might turn out and regardless, in particular, of whether the object exists.

However, even if it is granted that there is a transcendent relation of identity and also a basic relation of identity, it might still be denied that the very same relation is both transcendent and basic. Let us distinguish between an *existential* relation of identity, which holds between an object and itself in a possible world only when the object exists in that world, and a *neutral* relation of identity, which holds between an object and itself in a possible world regardless of whether the object exists in that world. In the one case, self-identity will be compatible with non-existence and, in the other case, not. My opponent, I assume, takes the existential relation to be basic. But he can still define a neutral relation of identity; for he can take x to be identical to y in this sense if it is possible that x is identical to y in the existential sense (or, alternatively, if it is necessary that x is identical to y if x

they may legitimately be taken to be. It is also possible to give a formulation of the view in terms of 'supervenience' rather than 'definition'.

and y exist). As long as each possible object in the range of the variables 'x' and 'y' possibly exists, this will then yield the neutral relation of identity. Conversely, my ally, who takes the neutral relation of identity to be basic, can define the existential notion; for he can take x to be identical to y in this sense if x is identical to y in the neutral sense and if x and y both exist. Thus the question is not whether there is this or that relation of identity, but which of them should be taken to be basic. Should we understand neutral identity as the possibility of existential identity or existential identity as an existential strengthening of neutral identity?[22]

When we consider this question, then it seems to me that there are strong intuitive considerations against taking the existential relation to be basic. For surely the existential identity of an object with itself cries out for analysis into two quite distinct components, one the neutral identity of the object with itself, and the other its existence. The existence of the object appears to be merely a gratuitous addition to its identity, one that can be removed without any essential alteration to the underlying relation. Compare this with a genuine worldly relation, such as *hitting*. It makes no sense to suppose that such a relation is separable into two distinct components, one of which is neutral as to existence and the other of which is existence itself. Our very understanding of what it is to hit presupposes that the hitter and the object hit should exist.

There also appear to be strong intuitive considerations against taking the neutral relation to be defined. For if it is to be defined from the existential relation, then it must be taken to be modal; to be identical in this sense is *possibly* to be identical in the more basic sense (and thereby exist). But it seems bizarre to suppose that the identity of an object with itself, in a circumstance in which it does not exist, rests upon the *possible* identity of the object with itself in a circumstance in which it does exist. Or to put the point in terms of distinctness rather than identity, it seems bizarre to suppose that the distinctness of two objects in a circumstance in which both do not exist might rest upon the possibility of their being distinct in a circumstance in which both exist. We seem perfectly capable of understanding a non-modal notion of identity or distinctness in these cases. Indeed, suppose there were *impossible* objects, ones that do not possibly exist. Then we would have no more difficulty in understanding what it is for such objects to be identical or distinct than in the case of possible objects. But on the proposed view, it would be impossible to understand how the relations of identity or distinctness might sensibly apply to such objects.

Although I have focused on the case of identity, it should be clear that very similar considerations apply to other cases of what one might loosely call

---

[22] I here ignore the possibility, which I do not take seriously, that there might be a third, more basic notion of identity or that identity is to be defined in other terms altogether.

'formal' relations. Consider, for example, the relationship that holds between a set and its members or between an aggregate and its parts or between a proposition and its constituents. In each of these cases we have what is prima facie a transcendental relation, one that applies regardless of the circumstances. It also looks as if we face a choice between taking either a neutral or an existential version of the relation as basic; and it may be argued, on similar grounds, that it is the neutral version that should be chosen. Thus if basic transcendence is granted in the case of identity, it should be granted in these other cases as well.

9.   Sortal Transcendence

I have argued in favour of certain formal relations being transcendent. I wish now to consider the status of sortal such as *man* or *cat*.

It has to be admitted that we are much more prone in the case of *man*, say, than in the case of identity to treat the property or relation as one that entails existence. Even if something can be self-identical without existing, how, it might be wondered, can something be a man without existing? And it seems to me that this doubt can persist even when we are quite clear that the modality in question is to be given the broad 'superextended' sense. Thus our solution to the puzzle is not adequate, in the case of 'man', to dispel all doubt we might have as to the truth of its conclusion.

One source of doubt might lie in the following syllogism. A man must be flesh and blood (let us suppose). But nothing can be flesh and blood without existing; and so nothing can be a man without existing.[23] A similar difficulty does not arise in the case of identity, since there is no comparable descriptive content that can be associated with the property of being self-identical.

This doubt, however, is readily allayed. For to say that a man must be flesh and blood may simply mean that, necessarily, he is composed of flesh and blood; and to say that he is composed of flesh and blood, in this context, may simply mean that whatever composes him will be flesh and blood (or will have flesh and blood as a part). But a man's not existing will then be compatible with his being flesh and blood (indeed, will *require* that he be flesh and blood!) in this sense, since there will then be nothing of which he is composed.

I suspect that some deeper errors may have stood in the way of recognizing the existential neutrality of 'man'. Thus it is sometimes supposed—for reasons having nothing to do with the puzzle—that 'man' is, in effect, a tensed predicate, having application to an object *at a time*. Consider, for example, the usual way of drawing the distinction between a phase-sortal

---

[23] The argument (with actualist quantifiers) might be formalized as follows: $\Box \forall x \Box (M(x) \supset FB(x))$, $\Box \forall x \Box (FB(x) \supset E(x))$, $\therefore \Box \forall x \Box (M(x) \supset E(x))$.

term such as 'child' and a substance-sortal term such as 'man' (as in Wiggins [1980], 24–5, for example). It is thought that a phase-sortal need have only temporary application, applying to the object at some of the times at which it exists and not at other times, but that a substance-sortal must have permanent application, applying to the object at any time at which it exists. The one picks out what may be only a phase of the object, while the other picks out the object in its temporal entirety.

But this is surely to mischaracterize the distinction between the two kinds of term. A substance-sortal does not, properly speaking, have application to an object *at a time* at all. Socrates *is* (tenselessly) a man and it makes no sense for us to say that he *was* a man or for his parents-to-be to say that he *will be* a man.[24] Indeed, it seems to me that, even if we indulge in the fiction that substance-sortals have application at a time, there is still no possibility of distinguishing them from phase-sortals in the manner proposed. For 'existent man' is as much a phase-sortal as 'young man'; and yet 'existent man' has the same application to a man (when he exists) as 'man'.

I suspect that this error is not a mere case of sloppiness. Part of what may mislead is the superficial similarity in grammar between 'man' and 'child'. This may lead one to believe that their semantics is the same; and, since 'child' has application to an object at a time, so, it may be thought, should 'man'. But there may also be an underlying metaphysical error. There is a way a man must be at a time if he is to be a man; and it is principally through recognizing that someone is this way at a time that we recognize that he is a man. This may then lead one to believe that this temporary feature, by which we recognize something to be a man, is what it is to be a man.[25]

It seems to me that philosophers may have been guilty of a similar error on the modal side. Indeed, if the predicate 'man' is taken to be tensed then, *eo ipso*, it should be taken to be worldly; and the reasons there are for regarding it as tensed are equally good reasons, *mutatis mutandis*, for regarding it as worldly. But just as it is a mistake to think of 'man' as tensed, so it is a mistake to think of it as 'worldly'; whether an object is a man no more depends upon its worldly state than upon its temporary state.

These points become all the stronger once one considers how the term 'man' might be defined. As I have already remarked, it is implausible to suppose that 'man' is a basic predicate, one standing for a basic property; there is something it takes to be man even if there is some difficulty in saying exactly what it is. Suppose now that the predicate 'man' has application at a

---

[24] I here ignore the possibility of metamorphosis—of the very same object being a man, say, and then being a fly; I also ignore the possibility of using 'man' to mean 'adult man'. In neither of these cases would *man* be a genuine substance-sortal and so another example would be needed to make the point.

[25] This error permeates Locke's discussion of identity in the *Essay* and seems to have had an unfortunate influence on contemporary thinking about the topic.

time. Then what kind of definition might result in its being a predicate of this sort? Presumably (though this is not inevitable) the definition will relate to how the man is at a given time. But in this case, the definition will be unable to distinguish between a man and those objects that temporarily coincide with him, such as his time-slice at the given time or the quantity of matter that then makes him up. Suppose, on the other hand, that the predicate 'man' has application in a world. Then presumably the kind of definition that would result in its being a predicate of this sort is one that relates to how the man is over time; it will, in effect, specify the admissible 'careers' or temporal profiles that a man can assume. But in this case, the definition will be unable to distinguish between a man and the objects that coincide with him in the world, such as his tissue (something that would normally survive his death but that might come in and out of existence along with the man).[26] Thus it appears that no plausible definition of a man, either as a tensed or as a worldly predicate, will be extensionally correct.

But if we are not to be guilty of these various errors, then how should 'man' be defined? A natural suggestion is to extend the previous types of definitions into the modal dimension, as it were. To be a man is to have a certain kind of temporal-modal profile. Not only is it to have a certain kind of temporal profile in any possible world in which it might exist, it is also to have a related temporal profile in any other possible world in which it might exist; the various temporal profiles must be suitably co-ordinated across different worlds.

Even if we assume that a definition along these lines might be correct, it will be of no help to a philosopher who believes in existential 'man'. For, under the proposed definition, 'man' will signify a modal property; and so, like any modal property, its attribution to any object should hold of necessity, regardless of whether the object exists. We could, of course, take a man to be something that is a man in this neutral sense and yet also exists. But clearly, in this case, existence is merely a gratuitous addition to what is naturally taken to be the genuine, underlying, notion of manhood (just as in the case of existential identity).

A definition of this sort, if correct, may be of some comfort to the worldly philosopher, since it means that he need not recognize sortals as an independent source of transcendence, apart from the use of modality. However, I am very doubtful whether it is correct. Just as there are counterexamples to the temporal and worldly forms of definition, so, I believe, are there counterexamples to the modal form of definition

---

[26] I have assumed that the temporary and worldly coincidents exist and that they are distinct from the man. Fine [2000, 2003] contain a partial defence of this position. The case of Lumpl and Goliath from Gibbard [1975] provides another, more natural, example of coincidents within a world.

(Fine [2000]). But there is another, more basic misgiving that one might have. For one naturally thinks that it is an object's being a man that *accounts for* its having the kind of temporal-modal profile that it does. But on the present view, no such account or explanation is possible, since to be a man is simply to have a certain kind of temporal-modal profile. More generally, one might well think that an object's modal features should flow from *what it is* (i.e. from its essential features). But this means that one should not appeal to an object's modal features in stating what the object is, since they could not then be seen to flow from what it is. Since being a man is part of what a man is, it follows that there should be a non-modal account of what it is to be man.

The only plausible non-modal definition of 'man' is one that classifies the object under a sort; to be a man is to be an F (where this is the sort) differentiated in such and such a way. If we use only ordinary non-sortal properties in stating what a man is, then it is hard to see how the definition could have the required modal import. The thing, perhaps, should be fleshy but why *must* it be fleshy (when it exists)? If, however, we say that it is a fleshy *animal*, then we may take this to imply that it must be fleshy.[27]

This suggests that every object should be taken to fall under some general and basic sort, one that cannot be subsumed under any other sort or be defined in other terms; and these sorts will therefore constitute a further, independent, source of transcendence. Far from occupying a predicational void, the non-existents at a world will possess a rich panoply of sortal and formal features.

These features are not only significant as sources of transcendence, they are also significant as means by which we may specify what things *are*. Consider the case of sets by way of illustration. In specifying what a set is, we must state two things. First, we must state what general kind of thing it is—in this case, a *set*. Second, we must state how it is to be differentiated from other objects of the same sort—in this case, by its *members*. Thus the general sort, *set*, and the associated formal relation of *membership* come together in providing an account of what a particular set is. Since the sort and the relation are both transcendental, we therefore have a transcendental specification of what the object is.

Now I do not believe that the essence of an object is wholly given by its transcendental features. But I do believe that the transcendental essence of an object constitutes a kind of skeletal 'core' from which the rest of the essence can be derived. We therefore arrive at the view that the identity of an

---

[27] It is important, if this strategy is to work, that a fleshy animal should not simply be taken to be object that is both fleshy and an animal. There must be some other, non-predicational way in which 'fleshy' qualifies 'animal', but I shall not discuss the difficult question of what this might be. All that matters, for present purposes, is that it not involve an incursion into modal terrain.

object is independent of how things turn out, not just in the relatively trivial sense that the self-identity of the object is independent of how things turn out and not just in the relatively trivial sense that the identity of the object is something that will hold of necessity. Rather it is the core essential features of the object that will be independent of how things turn out and they will be independent in the sense of holding *regardless* of the circumstances, not *whatever* the circumstances. The objects enter the world with their identity predetermined, as it were; and there is nothing in how things are that can have any bearing on what they are.[28]

10. Transcendental Existence

I have argued that one should be willing to countenance an unworldly perspective if one is to have a satisfactory view of what things are; and I now wish to argue that one should also countenance such a perspective if one is to have a satisfactory view of what there is.

Just as there is a distinction between sempiternal and eternal truths, so there is a distinction between sempiternal and eternal existence. Some objects (perhaps some particles) always exist while others (such as numbers or facts) do not exist in time at all. This suggests that there should be a corresponding distinction between necessary and transcendental existence. Some things (perhaps an example is the lapse of time) will exist whatever the circumstances, while others (such as numbers) will exist regardless of the circumstances. However, there are great difficulties in making out these further distinctions; and in order to get a better handle on how they work in the modal case, it will be helpful if we first consider how they work in the more familiar case of tense.

Suppose that one is a B-theorist about tense. There are then three notions of existence or being that may in principle be distinguished. First, there is a time-relative notion of existence-at-a-time. It is in this sense of 'existence' that Clinton exists at the current time but did not exist 100 years ago and will not exist 100 years hence. There is then a corresponding timeless or derelativized notion of existence. This may be used in either an exclusive or an inclusive sense. In the exclusive sense, the number 2 will enjoy timeless existence while Clinton will not. In the inclusive sense, Clinton (and every other object in time) will also enjoy a timeless form of existence. This is a sense in which Clinton exists, not *at* a time, but *simpliciter*. Finally, there is

---

[28] Almog [1991] has emphasized the divide between an object's constitutive properties and its modal properties. I might note in this regard that it is only in an extended and therefore somewhat artificial sense of 'necessity' that an object can be taken necessarily to possess the features belonging to its transcendental essence. Thus, in this case, the nature of the features themselves constitutes a barrier to their straightforward transfer from the identity of the object to how it must be.

an *ontic* notion of existence. For something to exist in this sense is simply for there to be something that it is. This is the sense of existence that is tied to our understanding of the quantifier; where '∃y' is the unrestricted quantifier, x will exist in this sense if $\exists y(x = y)$.

We might talk of what there *is* when the third notion is in question and of what *exists* when the first and second notions are in question. Thus being is absolute while existence can be either relative or absolute. But even though I have introduced two absolute notions, one of existence and the other of being, and introduced them in a somewhat different way, there seems to be no reasonable basis upon which the B-theorist might differentiate between them; to exist, for her, is simply to be.

Suppose now that one is an A-theorist about tense (it is this position that provides the relevant analogy to the modal case). What then will be the corresponding notions of existence? Corresponding to the time-relative notion of existence will be an absolute (quasi-relative) notion of tensed existence. It is in this sense of 'existence' that Clinton currently exists but once did not exist and will subsequently cease to exist. Corresponding to the timeless notion of existence will be an absolute tenseless notion of existence—or, if one likes, it can be taken to be degenerately tensed, applying to an object at a time, regardless of what the time might be. Finally, corresponding to the ontic notion is another absolute and genuinely tensed notion. To exist or be, in this sense, is to be identical to something that there is; but since what there is, for the A-theorist, is tensed, so is the associated notion of being.

Thus there will be three absolute notions of existence or being for the A-theorist. The distinction between the first two of them should be evident, since the number 2 exists in the second sense though not in the first. However, the distinction between the third notion and the other two is not so clear. The difficulty arises from the fact that the A-theorist can have different substantive views about what there is. According to the presentist position, at least in application to temporal objects, what there is is what presently exists; according to the liberal position, what there is is what exists generally—in the past, present, or future; and according to an intermediate position, what there is is what exists in the present or past. It has been common to assume the presentist position in the context of the A-theory; and this then makes it look as if there is no genuine distinction to be drawn between what there is and what exists in the first (quasi-relative) sense.[29] But presentism constitutes a substantive doctrine about what there is and so any particular way of filling out what there is should be distinguished from our general conception of what there is. Indeed, under the intermediate position,

---

[29] Reasons for not tying the A-theory to an ontological form of presentation are given in Ch. 10, §10, below.

the objects that fall under what there is will differ from what exists in either the first or the second sense.[30]

The first two notions of existence are *object*-driven. What exists in either of these two senses is determined by relatively straightforward considerations concerning the character of the object in question. Thus it is not in the nature of the number 2 to exist in time and it is for this reason that we take it to enjoy an exclusive form of timeless existence. The third notion of existence or being, on the other hand, is *domain*-driven. What there is in this sense depends upon relatively problematic considerations of ontology. The question is one of ontic availability, of what is 'at hand'. And this question, for the A-theorist, may plausibly be answered on the basis of various more substantive conceptions of what exists.

We see that whereas it is difficult for the B-theorist to maintain the distinction between existence and being, there is no such difficulty for the A-theorist. Indeed, we might put the difference in the views this way: that whereas there is a single (absolute) notion of existence or being for the B-theorist, there is a distinction between the two for the A-theorist. For there is a domain-driven conception of being that is to be understood in terms of the quantifier; and there is an object-driven conception of existence that is to be understood by contrast with a genuinely tensed, or time-relative, form of existence.

This tripartite scheme is not the usual one—it is usually supposed that there is a single univocal notion of existence for the A-theorist; and so it is important to appreciate what might have led to the collapse of the one scheme into the other. First, it is not usually supposed that there is any need to distinguish being from existence in the first quasi-relative sense. After all, both notions may be expressed by means of a tensed monadic predicate; and given that one can be taken to be a criterion for the other, it is very easy to ignore the difference between them. And secondly, it is often supposed that there is no room within the A-theorist's ontology for a realm of timeless existents. Just as there has been a tendency to think that every sentence is tensed, so there has been a tendency to think that every object must enjoy a tensed form of existence.

Indeed, there is a powerful argument for this view, one that may have been in the back of the mind of those who have considered the presentist's position. For suppose that we are presentists—we believe that what there is is what presently exists. Then we should grant that there *was* such a thing as Socrates ($P\exists x(x = s)$) even though there (presently) is no such thing

---

[30] As we shall see, the notions of what there is and what presently exists might even come apart for the presentist once atemporal objects are taken into account. For, under any plausible presentist position, there will presently be such a thing as singleton Clinton even though this set will not presently exist in the same way that Clinton presently exists.

($\sim\exists x(x = s)$).[31] Suppose we also grant that Socrates enjoys a timeless form of existence. Then we should grant that there was such a thing as Socrates and that he timelessly exists ($P(\exists x(x = s)\&E(x))$). But what timelessly exists exists regardless of the time. Thus we may conclude that there was such a thing as Socrates and that he *now* exists ($P(\exists x(x = s) \& NE(x))$). And yet surely this is incompatible with presentism (and, in particular, with the claim that $\sim\exists x(x = s)$).

It therefore appears that we should deny that Socrates enjoys a timeless form of existence. But this is not strictly necessary. For one can simply deny that there is any conflict between the given assertion of existence and the doctrine of presentism. We wish to say that Socrates exists but that there is no such thing as Socrates. And it is only if we suppose that the relevant sense of 'exists' is 'ontic', that it implies that there is something that exists, that the two will be in conflict.

Philosophers have been wedded to the idea that existence should be explained in terms of the quantifier, of what I have called 'being'; for an object to exist is for it to be, i.e. for there to be something that the object is. But think of existence as a certain very general kind of status; to exist is to be some sort of existent. It is not then so clear that existence should be tied to being. For, if I am right, the status or kind of an object—its being a man, for example, or a cat—is not tied to being. And so why should the most general form of status or kind be tied to being? Indeed, we might naturally think that status of some specific kind should entail the appropriate form of existence. And this will then require that existence should be like the kind in being detachable from what there is.

A similar, and perhaps even more compelling, form of the difficulty we have considered arises in connection with 'hybrids'. These are objects (somewhat analogous to hybrid sentences) that have constituents that exist in time even though they themselves do not. Consider singleton Socrates, for example, the set whose sole member is Socrates. Even though we might be tempted to deny that Socrates timelessly exists, it seems much harder to deny that singleton Socrates timelessly exists. For surely sets, in general, are not the kind of thing to exist in time. Given that the null set does not exist in time, then why should it be any different for singleton Socrates? However, it seems clear that singleton Socrates once existed. And so, given that its existence is timeless, it presently exists. But how can that be unless Socrates exists?

In responding to this further difficulty, we should not deny that singleton Socrates timelessly exists. And nor should we deny that the existence of singleton Socrates implies, in the very same sense, of 'exists', the existence

---

[31] I use the name 'Socrates' for convenience only. The difficulty can equally well be stated with a predicate 'Socratizes' in place of the name.

of Socrates. We should simply allow that Socrates exists in the required sense. We may grant, as a general ontological principle, that there is no such thing as singleton Socrates unless there is such a thing as Socrates (the set is not 'at hand' unless its members are 'at hand'); and, as presentists, we may grant that there is no such thing as Socrates. From this it follows that there is no such thing as singleton Socrates. But this is entirely compatible, on the proposed view, with singleton Socrates being a timeless existent. Such sets, we might say, *are* in time but do not *exist* in time.

I wish to understand the A-theorist's position in the modal case by analogy with his position in the tense-logical case. Thus he will distinguish between three notions of existence or being. First, there is worldly or circumstantial existence. This is a form of existence that turns on how things turn out. Thus we may say, in this sense, that Socrates exists. Then there is unworldly or transcendental existence. This is a form of existence that holds regardless of how things turn out. It is in this sense of 'existence', construed exclusively, that numbers and sets exist; and it is also in this sense of 'existence', construed inclusively, that worldly existents may be said to exist. Finally, there is the ontic notion; to exist or be in this sense is to be identical to something that there is. Since what there is, for the modal A-theorist, is a worldly matter, so is the associated notion of being.

There are two main positions that the modal A-theorist might adopt on what there is. According to actualism, at least in its application to worldly objects, what there is is what is actual; and, according to possibilism, what there is is what is actual or possible.[32] But these are substantive positions; and so, again, we need to distinguish between the ontic notion and the standard actualist criterion for its application.

There are modal analogues of the previous two puzzles concerning existence; and they may be solved in an analogous manner. Thus we may take Socrates to enjoy an inclusive form of unworldly existence and take singleton Socrates to enjoy an exclusive form of unworldly existence. It therefore follows that, in a superextended sense of 'necessity, each object will necessarily enjoy its unworldly form of existence. But this should not be taken to imply that necessarily there *are* these objects or that necessarily they exist in a worldly sense. Singleton Socrates does not exist in a worldly sense at all and Socrates will only exist in a worldly sense when he *is*.

It might be thought odd that we take there to be a sense of existence in which Socrates necessarily exists. But we should bear in mind that this is an unworldly or transcendental form of existence; there is no worldly way in which Socrates has to be if he is so to exist. We might think of existence in this sense as the invariable concomitant of any object being what it is

---

[32] We can also imagine an intermediate position; for we might take a possible to be at hand when its 'ingredients' are at hand even if they have not been 'assembled' into something actual.

rather than of there being something that it is; and, since the object is what it is regardless of whether there is anything that it is, the same will be true of the form of existence that derives from its being what it is. To exist in this sense is simply to be an *existent*, i.e. the kind of thing that exists.

It might also be thought odd that we take singleton Socrates to be a transcendent as opposed to a worldly existent. For does not its existence turn upon the existence of Socrates; and so is not its existence a worldly matter? The question can be put into sharper focus if we contrast the set whose sole members are Socrates and Plato, say, with the *aggregate* of Socrates and Plato. We take the aggregate to be a worldly existent since its existence turns upon the existence of Socrates and Plato. So why not with doubleton Socrates and Plato as well?

The answer, I believe, lies in the role that existence plays within the structure of predication. An object will enjoy whatever form of existence is required to give it 'life', i.e. to allow it to have its characteristic properties or to enter into its characteristic relationships. Now part of the 'life' of an aggregate includes location and, in order for an object to have location, it must exist in a temporal (and hence also in a worldly) sense. However, there is no more to the 'life' of a set than its membership (and its being a set). These are not worldly matters; and so a worldly form of existence is not required.

Just as we have recognized different levels of reality, so we should recognize different levels of existence. Each object will exist at the lowest level at which it can enjoy its characteristic form of life. At the bottom are the objects that exist in time; these are the objects whose existence is subject to the vicissitudes of time. At the next level are the worldly objects, whose existence is subject to the vicissitudes of the world though not of time. Finally, at the top, are the transcendental objects, whose existence lies outside the world or time. Each type of existent will enter into facts of the corresponding type (and perhaps also of a higher type)—temporal objects into temporal facts, worldly objects into worldly facts, and transcendental objects into transcendental facts. It is just such a picture—stratified with respect both to how things are and what there is—that consideration of the initial puzzle seems irresistibly to lead.[33]

IV

Reviews

This page intentionally left blank

Review of *Counterfactuals* by David Lewis

This is an excellent book. It combines shrewd philosophical sense with fine technical expertise; the statement of views is concise and forthright; and the level of argument is high.

The book is centred upon an analysis of counterfactuals in terms of possible worlds. The first chapter presents the analysis: the counterfactual 'if it were the case that ϕ then it would be the case that ψ'—written '$\phi \square \rightarrow \psi$'—is true either vacuously or non-vacuously; vacuously if no ϕ-world is 'entertainable'; non-vacuously if, within some degree of similarity to the actual world, some possible world is a ϕ-world but none is a $\phi \,\&\sim \psi$-world. Later chapters deal with related matters: reformulations of the analysis; a comparison with rival accounts; the foundations in terms of possible worlds and comparative similarity; analogies with other notions; and questions of axiomatization and decidability.

Thus although the book is centred upon counterfactuals, it is by no means confined to them. There are interesting passages on context, vagueness, law, definite descriptions, and deontic logic. For reasons of space, I shall not discuss these topics. However, I shall consider, in addition to the actual analysis, the philosophical question of its foundations and some more technical questions arising from its justification and reformulations.

The main alternative to Lewis's analysis is what he calls the metalinguistic theory. On one version of this theory, the counterfactual $\phi \square \rightarrow \psi$ is true if ψ is a logical consequence of ϕ and some other sentences which are either implicit from the context or co-tenable with ϕ. Co-tenability is in need of explanation; but for a wide range of counterfactuals the co-tenables can consist of laws of nature and certain auxiliary conditions.

There is, of course, nothing essentially metalinguistic about this alternative account. Let $|\phi|$ be the truth-set of the sentence ϕ, i.e. the set of worlds in which it is true. Then we can say that the counterfactual $\phi \square \rightarrow \psi$ is true at a world i if there are sets of worlds which are co-tenable with $|\phi|$ and such that $|\psi|$ contains their intersection with $|\phi|$. Here co-tenability is a primitive relation between sets of worlds and dependent upon the given world i.

David Lewis [1973], *Counterfactuals* (Oxford: Blackwell).

   This reformulation suggests that the difference between the two accounts may not be one of substance. For we may follow Lewis and define co-tenability so as to make the accounts coincide; or we may equally well define similarity in terms of co-tenability. However, I do not think that the possibility of these definitions should be taken too seriously. For no metalinguistic theorist would accept that Lewis's definition provides the correct extension of *his* notion of co-tenability; and similarly for the other definition. Thus on Lewis's account, a counterfactual with true components is true and so the co-tenables of a true sentence will include or, at least, imply all truths. But no metalinguistic theorist will accept this. Again, if there is a closest ϕ-world i then all sentences true at i will be co-tenables of ϕ on Lewis's account. And no metalinguistic theorist will accept this either.

   In fact, the accounts differ in principle. For the metalinguistic theorist will verify a counterfactual by taking certain propositions as fixed and seeing if the consequent follows from them; whereas Lewis will consider a change in the truth-value of any proposition (other than the antecedent) so long as the overall dissimilarity from the actual world is thereby reduced. Verification of the one counterfactual will depend upon discernment of laws of nature and auxiliary conditions, say. Verification of the other counterfactual will depend upon judgements of overall similarity; and there will be only a tenuous link between what is preserved in all sufficiently closed worlds and what the metalinguistic theorist will wish to keep fixed.

   Implicit in this contrast is a serious objection to Lewis's analysis. The counterfactual 'if Nixon had pressed the button there would have been a nuclear holocaust' is true or can be imagined to be so. Now suppose that there never will be a nuclear holocaust. Then that counterfactual is, on Lewis's analysis, very likely false. For given any world in which antecedent and consequent are both true it will be easy to imagine a closer world in which the antecedent is true and the consequent false. For we need only imagine a change that prevents the holocaust but that does not require such a great divergence from reality.

   We can imagine that the holocaust's not occurring is a miracle, an exception to an otherwise universal law. This is especially plausible if there already are miracles in the actual world, but it is still plausible if there are not. For, arguably, a world with a single miracle but no holocaust is closer to reality than one with a holocaust but no miracle. The agreement of particular fact, in the first case, is so great as to override the disagreement of law in the second case. But there is no need to imagine a miracle. An undetermined break in an electrical connection will do or even a determined break, so long as the consequences are insignificant in comparison with a holocaust.

   To take another example, 'if Oswald had not shot Kennedy, then someone else would have' is probably true for Lewis on the grounds that the consequences of supposing that someone else shot Kennedy would make less

difference to the world than those of supposing that Kennedy was not shot after all. In general, it will be difficult for Lewis to make any great change from reality counterfactually consequent upon a small change, for the great change may be avoided by positing certain other small changes. One change will be traded for another so long as the overall dissimilarity from the actual world is thereby reduced.

There are two minor objections to Lewis's account that are of some independent interest. They differ from the preceding objection in being directly relevant to the logic of counterfactuals. The first of these objections is that Lewis's account incorrectly sanctions the inference:

$$\frac{\phi, \psi}{\phi \square \to \psi}$$

For if ϕ and ψ are both true then the actual world is a ϕ-world and no $\phi \,\&\, {-}\psi$ world is as close, since no world is as close to a given world as that world itself.

Lewis's response to this objection is that the assertion of a counterfactual whose antecedent is known to be true is misleading but not false. However, it may be appropriate to assert a counterfactual while in ignorance of its antecedent's truth-value. I may speculate on a student's prospects in an exam, the results of which are already settled, and assert: if he had worked hard he would have passed. My assertion is false if the student worked hard but was only able to pass through cheating. If I say the assertion is true I cannot generalize to: if a student of similar ability had worked hard then he would have passed. Indeed, such counterfactuals hold in virtue of a connection between antecedent and consequent that is not guaranteed by the truth-values of the components alone.

A related case is that of a counterfactual with false antecedent but true and unconnected consequent, for example: if I were to raise my little finger there would be a rainfall this Winter. This is false, but true on Lewis's analysis since sufficiently close antecedent-worlds will preserve the truth of unconnected propositions. Of course, the conditional 'There would still be rainfall this Winter, even if I were to raise my little finger' is true. But as Goodman has observed, this conditional seems to imply the negation of the counterfactual 'if I were to raise my little finger there would not be a rainfall this Winter' rather than the original counterfactual itself.

The second minor objection concerns counterfactuals with disjunctive antecedents. On Lewis's analysis, the inference-pattern

$$\frac{\phi_1 \mathrm{V} \phi_2 \square \to \psi}{\phi_1 \square \to \psi}$$

is not sound, for indefinitely close ϕ_1-worlds may be $\sim \psi$-worlds while all of the much closer ϕ_2-worlds are ψ-worlds. However, the counterfactual 'if

Thorpe or Wilson were to win the next General Election, Britain would prosper' does seem to imply 'if Thorpe were to win the next General Election, Britain would prosper'.

All the same, there is a difficulty in accepting this inference. For $\phi_1 \square \to \psi$ implies $(\phi_1 \& \phi_2) \vee (\phi_1 \& \sim \phi_2) \square \to \psi$ by substitution of tautologous equivalents, and this implies $(\phi_1 \& \phi_2) \square \to \psi$ by the inference in question. But surely the inference:

$$\frac{\phi_1 \square \to \psi}{\phi_1 \& \phi_2 \square \to \psi}$$

is not sound.

There are various ways out of this difficulty. One is to accept the substitution rule but to reject the inference-scheme concerning disjunction on the grounds that logician's '$\vee$' does not correspond to the 'or' that figures in what are apparently ordinary language instances of the scheme. Perhaps there are two senses of 'or', one genuinely disjunctive and the other always distributing conjunctively. Only a special construction, like 'if the proposition that ϕ were the case', will guarantee that an 'or' in ϕ has the first sense.

Another solution is to reject the substitution rule and, in particular, the free substitution of $(\phi_1 \& \phi_2) \vee (\phi_1 \& \phi_2)$ for ϕ_1. One semantic justification for this might be that each sentence refers to or indicates certain determinate states-of-affairs. $\phi \vee \psi$ would then refer to whatever states of affairs were referred to by ϕ or ψ (and perhaps also by $\phi \& \psi$), and the counterfactual $\phi \square \to \psi$ would be true if and only if each counterfactual $\chi \square \to \psi$ were true whenever χ singly referred to any state-of-affairs referred to by ϕ. Thus if ϕ_1, $\phi_1 \& \phi_2$ and $\phi_1 \& \sim \phi_2$ each referred to distinct single states-of-affairs then $(\phi_1 \& \phi_2) \vee (\phi_1 \& \sim \phi_2) \square \to \psi$, unlike $\phi_1 \square \to \psi$, would be true only if each of $(\phi_1 \& \phi_2) \square \to \psi$ and $(\phi_1 \& \sim \phi_2) \square \to \psi$ were true.

A compromise solution is to accept the substitution rule and to accept the inference concerning disjunctions for those disjunctions that express propositions which describe determinate states-of-affairs. If we call each such proposition (or truth-set) determinate, then $\phi \square \to \psi$ is true if and only if each $\chi \square \to \psi$ is true for $|\chi|$ determinate and $|\chi| \subseteq |\phi|$. This solution has the advantage of being classical and yet of explaining the truth-conditions of any counterfactual in terms of those for counterfactuals whose antecedents express determinate propositions.

A similar difficulty arises for the notions of permission and preference. Let us write 'Pϕ' for 'it is permitted to see to it that ϕ' and 'Pr(ϕ, ψ)' for 'that ϕ is preferred to that ψ'. Then given the substitution of tautologous equivalents, we cannot accept the inference P($\phi \vee \psi$)/P(ϕ) and reject P(ϕ)/P($\phi \& \psi$) and nor can we accept the inference Pr($\phi_1 \vee \phi_2$, ψ)/Pr(ϕ_1, ψ) and reject Pr(ϕ_1, ψ)/Pr($\phi_1 \& \phi_2$, ψ). In both of these cases, we can adopt solutions analogous to those above.

Let us now turn to the question of foundations. The two important primitives of the analysis are the notions of possible world and comparative similarity; and Lewis defends these notions against attacks of various sorts. The issues involved call for a very thorough discussion, and I shall have to confine myself to one or two comments.

On taking the ordinary modal idioms (possibly, necessarily, etc.) as primitive, Lewis remarks that 'this is not an alternative theory at all, but an abstinence from theorising'. But I do not see why the modal idioms are more in need of explanation than the notion of possible worlds, unless there is an objection to taking any non-extensional notion as basic. Indeed, there is a strong reason for identifying possible worlds with maximal collections of propositions X such that it is possible that all of the propositions in X are true. For, ultimately, there is no way of identifying a merely possible world other than through the propositions that are true in it. In this way, possible worlds become logical constructions defined in terms of the more usual notion of possibility.

Lewis suggests that merely possible worlds are like the actual world, 'differing not in kind but only in what goes on at them'. Indeed, for him there is no absolute difference between the actual world and the others: the difference is relative to a particular possible world as point of reference. A similar view has been held about the present time, but it is hard to accept for possible worlds. On the logical construction view, the actual world is distinguished by the property that all of its propositions are true. Here 'true' is an absolute term. It is not defined as truth in the actual world but, on the contrary, truth-in-a-world is defined as set-theoretic membership.

Whether possible worlds are fundamental is not too important for Lewis's analysis, for any definition of them is hardly likely to lead us back to counterfactuals. However, the notion of comparative similarity gives rise to an immediate danger of circularity. For similarity is a matter of agreement in propositions; and among those propositions will be counterfactual ones. So to evaluate a counterfactual one needs to compare worlds for similarity to the actual world and this would seem to require the evaluation of further counterfactuals. Perhaps there is impredicativity rather than circularity here: counterfactuals are defined in terms of similarity, which is explained in terms of quantification over counterfactual propositions, among others. But the impredicativity is metaphysically and epistemologically troublesome. For it is no longer clear what the truth of a counterfactual consists in or how we can ever come to know that counterfactuals are true or false.

To avoid the circularity, one might make a prima facie distinction between indicative propositions, which describe occurrences and are understood independently of counterfactuals, and all others: similarity would then be explained in terms of indicative propositions alone. This solution rests upon a distinction that some may find untenable. But it also makes an important

metaphysical assumption, namely that if two possible worlds agree on indicative propositions, on what occurs, then they also agree on counterfactual propositions, on what would occur *if*. For since similarity is judged in terms of indicative propositions alone, two worlds which agree on such propositions must enjoy the same relation of comparative similarity and so make the same counterfactuals true.

Another danger of circularity arises from the vagueness of comparative similarity. Lewis is quite right to insist that vagueness is, as such, no objection to using a notion in definitions. Carnap's old term 'explication' covers the separable tasks of analysing and of making more precise. Thus to say 'a man is bald if he has few hairs on his head' is to analyse but not to precisify. However, in the present case, it is not clear that the vagueness of counterfactuals and similarity exactly match, either in possessing corresponding border-line cases or in admitting corresponding ways of being made more precise. Comparative similarity suffers from multi-criterial vagueness; many different criteria of comparison and ways of balancing them are compatible with the meaning of the term. Worlds may be similar in respect of particular fact, or of general but non-causal patterns, or of humanly but not scientifically important events, and so on. Some counterfactuals are vague, but not in the same way. It would be an odd man who denied a counterfactual on the grounds that similarity was to be judged mainly in respect of particular fact alone so that the antecedent could be causally insulated from all other propositions.

If not all standards of overall similarity are relevant to counterfactuals then there is the possibility that the delimitation of the relevant senses may require counterfactual or related notions. Indeed, what makes Lewis's theory hard to judge is that the notion of similarity tends to accommodate itself to the counterfactual at hand; that sense of similarity is chosen that will yield the correct truth-conditions. The question is whether this accommodation can be characterized without circularity.

The earlier objection, in terms of great and small changes, was intended to show that none of the more central standards of similarity would give the correct results. It may be that some fairly sophisticated standard will work. As a first attempt, one might say that j is as similar to i as k if the first time t at which either j or k diverge from i is such that (i) j at t is as similar to i as k at t, and (ii) laws in i apply at t onwards in j if they do in k. This formulation already involves the problem of law. Improvements could also involve the problem of auxiliary conditions in that these might be required to specify a dependence between the antecedent of a counterfactual and an appropriate relation of similarity. Thus it is not clear if similarity can be suitably specified without circularity or without appeal to some independent theory of counterfactuals.

Finally I shall discuss two technical aspects of Lewis's work. The first arises from his objection to treating $\phi\square \rightarrow \psi$ as a strict conditional, i.e. of

the form $\square(\phi \supset \psi)$ for some notion of necessity. Essentially, the objection is that the inference-pattern:

$$\frac{\phi_1 \square \to \psi}{\phi_1 \& \phi_2 \square \to \psi}$$

is then made sound, although examples show that it is not.

There are many other objections of this sort: for example, that treating $\square \to$ as a strict conditional makes the inference-pattern:

$$\frac{\phi \square \to \psi}{-\psi \square \to -\phi}$$

sound. In these arguments, the reasoning is purely formal: it depends only upon the logical properties of the definiendum and proposed definiens. Thus the above arguments appeal only to the negative fact that $\square \to$ does not satisfy the inferences and to the positive fact that $\square$ satisfies the inference $\square\phi/\square\psi$ for ψ a tautologous consequence of ϕ in the first case and ψ tautologously equivalent to ϕ in the second case.

So perhaps there are general formal grounds for supposing that no definition of $\square \to$ in terms of $\square$ can be correct. There would then be a result of the form: if $\square$ satisfies certain logical properties then it cannot be used to define a binary connective that satisfies certain other logical properties. Certainly, philosophers have had some such view in regard to the notions of commitment and entailment. The natural definitions in terms of obligation and necessity, respectively, have led to the wrong logic, and so it is conjectured that there is no satisfactory definition at all.

There is, in fact, a more general issue here, which is whether any non-extensional connectives are essentially many-place in the sense that they are not definable in terms of unary modal connectives alone. The view that there are no essentially many-place non-extensional connectives might be called modal monism, by analogy with classical monism, which states that all many-place predicates can be defined in terms of one-place predicates.[1] Not that the two forms of monism are entirely independent, for there may be many-place predicates that can be defined in terms of one-place predicates and one-place modal operators but not in terms of the former alone. Thus it might be argued that all connections reduce to nomic ones which, in their turn, reduce to the application of a one-place nomic operator.

Given a semantics for the connectives in question, it may be easier to settle whether some case of modal monism is correct. For example, Hans Kamp has shown that under very weak assumptions the two-place tensed

---

[1] This sense of 'modal monism' should be distinguished from the sense discussed in Ch. 9 above.

connectives Since and Until are not definable in terms of one-place ones. But this result depends upon not having propositional quantifiers in the language; for with them there is a definition, due to Geach and reported in Prior's *Past, Present and Future*. Again, Lewis shows (pp. 45–6), how $\Box \rightarrow$ can be defined in terms of propositional quantifiers and certain one-place connectives. Thus within these wider definitional resources, the question of a general reduction remains open.

The second technical aspect arises from the two main formulations of the analysis. The first formulation is in terms of spheres: each sphere is a set of worlds within a certain degree of similarity to the given world; the non-entertainable worlds are those that lie in no sphere; so that $\phi \Box \rightarrow \psi$ is true if either no ϕ-world is in a sphere or some sphere contains a ϕ-world but no $\phi \& \sim \psi$-world. A later formulation is in terms of the three place relation $\leq_i$ of comparative similarity among possible worlds: 'j $\leq_i$ k' is read 'j is as similar to i as k'; and $\phi \Box \rightarrow \psi$ is true at i if either no ϕ-world is within the field of $\leq_i$ or some ϕ-world k is such that $(\phi \supset \psi)$ is true at each j for which j $\leq_i$ k.

The rationale and status of these two formulations are not altogether clear. It is surely the formulation in terms of similarity that is the more basic. For spheres must be explained as sets of worlds any of whose members is more similar to the given world than a non-member, and the conditions on spheres must be justified on the basis of corresponding conditions on the comparative similarity relation.

The formulation in terms of comparative similarity is more fundamental in another way. It is elementary or first-order in the double respect of having an elementary truth-clause for $\Box \rightarrow$ and of imposing only elementary conditions upon the similarity relation. Thus the notions of 'true in a model' and 'model' can be defined without resort to quantifiers over sets of world. On the other hand, the formulation in terms of spheres is not elementary in either sense. It would be interesting to have more information on the scope of elementary semantics for modal languages and, in particular, to know if there is a modal logic complete for some non-elementary semantics but no elementary semantics.

The formulation in terms of spheres highlights the semantic difference between counterfactuals and strict conditionals, with one appealing to a fixed sphere of possible worlds and the other to a system of such spheres. However, the main significance of the formulation is that it makes the underlying topological ideas explicit. With each world is associated a topological space, i.e. a system of open sets or neighbourhoods, defined upon some subset X of possible worlds. The truth-clause for $\Box \rightarrow$ is as before, but with 'open set' in place of 'sphere'. $\phi \Box \rightarrow \psi$ then expresses a natural separation property on disjoint and non-empty sets Y = $|\phi \& \psi| \cap$ X and Z = $|\phi \& \sim \psi| \cap$ X, namely that the closure of Z does not contain Y.

Lewis does not use this more general account. Since his relation of comparative similarity is connected, he assumes that the spheres or open sets satisfy some special assumptions. However, for a non-connected relation of similarity or for other applications altogether, it may be useful to have the more general account.

Review of *The Nature of Necessity* by Alvin Plantinga

This book discusses several topics in the theory of modality: the *de re/de dicto* distinction, possible worlds, essences, names, possible objects, and existence. In the final two chapters, the preceding material is applied to the problem of evil and the ontological argument. In its philosophical (though not theological) parts, the book is close to Kripke's *Naming and Necessity*. There are similar accounts of the a priori/necessary distinction, proper names, transworld identity, and the identity theory.

The first chapter clarifies the notion of metaphysical necessity and distinguishes it from notions with which it is often confused; being logically necessary (in the narrow sense), unrevisable, a priori, or self-evident. It is an excellent antidote for someone who is still in the grip of the positivist conflations.

The next two chapters defend modality *de re*. First, it is shown that standard objections to *de re* modality fail once the objectual interpretation of the quantifiers is kept firmly in mind. But for the sceptic who is still not convinced, it is shown how modality *de re* can be reduced to modality *de dicto*. The problem is to make the transition from a necessity predicate or connective for sentences to quantified modal logic. Plantinga does not take the problem in quite this form. For he attempts to go from a necessity predicate of propositions to an essentialist relation between objects and properties. The first problem appears to be more basic than the second, but, in any case, both problems should be solved by a successful reduction of *de re* modality.

Plantinga's treatment of the reduction problem is complicated by his unusual use of the terms 'necessarily', 'essentially', and 'necessary'. He wants to say:

(1) 'necessarily' ($\Box$) and 'essentially' are synonyms (p. 14);
(2) x has P essentially iff x has P in each world in which x exists (p. 56);
(3) a proposition is necessary iff it is true in each possible world (p. 55).

These definitions are perfectly consistent, though it is never made clear whether the sentence Fa is false or lacks a truth-value in case of a's non-existence. It is simpler and more natural, however, to maintain (2) alongside:

(4) The sentence $\Box A$ is true iff A is true in all possible worlds;

(5) $\Box A$ is true iff the proposition expressed by A is necessary.

Plantinga must deny (5) whereas we can replace (2) with the equivalence of '*a* has F essentially' to '$\Box(a$ exists $\supset Fa)$'.

Given this simplification, Plantinga's solution to the original reduction problem would take the form: replace $\Box\varphi(x)$ with 'for some proper name *a* of *x*, the sentence $\ulcorner\varphi(a)\urcorner$ is necessary'. This is close to the account in Kaplan's 'Quantifying In'. Although Plantinga lays great stress on proper names, their use is not essential in an account of this form. One might hold, for example, that proper names are descriptions but that logically proper names would yield a correct reduction.

The main objection to this reduction is that it requires that proper names, or whatever is used in their place, be rigid designators even though the notion of rigid designation would not be intelligible to the *de re* sceptic. For example, it requires that the sentence $a = b$ is necessary when *a* and *b* are distinct proper names for the same object (see p. 35). Now this result is certainly correct on essentialist intuitions. But the *de re* sceptic need not accept it. He may say that necessary sentences are analytic, that they are the logical consequences of meaning postulates. The sentence $a = b$ will then not be necessary, for it is only plausible to suppose that the meaning postulates contain no essential occurrence of proper names. Even with these identity-sentences as meaning postulates, the reduction may still not be satisfactory. For it will yield only trivial *de re* truths, with $\exists x\Box\varphi(x)$ implying $\Box\forall x\varphi(x)$ for $\forall x\varphi(x)$ a name-free sentence. It is only when the meaning postulates distinguish between names that the *de re* superstructure becomes more interesting. Thus only essentialist intuitions will generate non-trivial *de re* truths.

Another objection is that the reduction should apply to sentences $\varphi(x)$ already containing $\Box$. The necessity predicate *for* the given language will then need to be a predicate *of L*. But like semantical closure, this can give rise to contradiction. For if *L* is suitably expressive, it will contain a formal analogue of the modal Liar, namely, 'This sentence is necessarily false'; and if necessity satisfies the logic $S5$, this sentence will lead to paradox. Thus a rigorous account of the reduction may well require a hierarchy of necessity-predicates or some other device for avoiding paradox.

The next three chapters, IV–VI, deal with possible worlds, essences, and transworld identity. In the first of these chapters, Plantinga takes a possible world to be a maximally possible state of affairs or proposition. He then uses this account in defining some related notions, such as existence, actuality, truth-in-a-world, and essential property. I agree with many of his points, but I do not accept that propositions or possible worlds necessarily exist. I think that, in general, the propositions that *x* has P exists only in a world in which

x does. Plantinga does not like propositions to the effect that an object has a certain property. But his arguments against such propositions betray a double standard; for they are exactly analogous to the arguments Quine uses against *de re* modality. If it is asked which sentence expresses such propositions, then Plantinga, at least, can say it is that sentence in which a proper name refers to the object.

In the next chapter, an essence of an individual is defined as a property that is essential to the individual and that is not had by any other individual in any possible world. Examples of essence are given and its connection with proper names is discussed at some length. The examples include world-indexed properties 'having-*P*-in-*W*', where exactly one individual *a* has *P* in *W*. Some philosophers have boggled at this use of worlds, and so it is worth noting that the essence can be expressed without them. The world *W* already contains the fact that *a* has *P* uniquely, and so a tighter formulation of the same essence is given by $\Box(\forall y(Py \equiv y = a) \supset Px)$.

In chapter VI, Plantinga argues that transworld identity is a pseudo-problem and that counterpart theory is inadequate on both semantic and metaphysical grounds. I think, however, that the whole dispute is, at bottom, metaphysical. If one's view of individuals is such that a possible world must first be described in general terms, then there is a problem of cross-world identification and counterpart theory is a plausible, though perhaps not unique, solution to that problem. If, on the other hand, a possible world can, without further analysis, be described with the help of names, then there is no problem of cross-world identification and counterpart theory is most implausible. It is right to insist upon the legitimacy of the second position; but that is not to refute the first position or to overcome the problems to which it gives rise.

The next two chapters, VII and VIII, deal with the question of merely possible objects. Plantinga is an actualist; he believes that only actual objects exist, and so he is concerned to defend this doctrine against attacks of various kinds. The first problem arises from the possible worlds semantics. For if one accepts that some possible object is not actual, the semantics requires that one quantify over merely possible objects. Plantinga's solution to this problem is extremely weak. He merely shows (p. 132) that one possibilist statement, that some object does not exist in the actual world, can be given an actualist reformulation. But what is required is an actualist reduction of *all* possibilist discourse. I think that this can be given, but Plantinga does not show how.

Another problem relates to what Plantinga calls the Ontological Principle. This states that an object has no properties in a world in which it does not exist. The problem is to reconcile this principle with the possible truth of negative existentials, such as 'Socrates does not exist'. Plantinga's answer is that this sentence *denies* a property of Socrates. But this will not do. For to

deny a property of Socrates is to assert its complement of him. To this, Plantinga replies that the complement of a property also requires the existence of its subject. But this will not do either. For if a sentence asserts a property of an object, let the *external* complement of the property be that property which the negation of the sentence asserts of the object. Then the external complement of a property will not, in general, require the existence of its subject.

I myself think that the Ontological Principle should be rejected, though one could follow Prior and deny the possible truth of any negative existential. Part of the attraction of the Principle may lie in the confusion of $\Box(\forall x \, (\varphi(x) \supset Ex))$, which is valid, with $\forall x \Box (\varphi(x) \supset Ex)$, which is invalid when $\Box$ is subject to (4). A counter-instance to the latter formula may require properly possibilist quantifiers in the possible worlds semantics; but this takes us back to the first problem and is not an additional difficulty.

The third problem is that fictional names appear to denote merely possible objects. Plantinga argues that this is not so, and that when we tell a story $\varphi(\text{John})$, we are exhibiting the general proposition expressed by $\exists x \varphi(x)$ (or $\exists x \, (x \text{ was named 'John' } \& \varphi(x))$). I think that he is right in his positive contention and wrong in his negative. First, he has got the speech-act wrong. In telling a story, we do not merely exhibit a proposition, we also pretend that it is true. Second, the proposition is wrong. We do not pretend that a general proposition is true, rather we pretend that the story expresses a true singular proposition. Thus names have the same function in fiction (but within a pretence) as they have in ordinary assertion.

In an appendix, Plantinga discusses Quine's objections to quantified modal logic and some of the responses to these objections. Many of Plantinga's remarks are interesting, but others are puzzling. For example, he is opposed to the distinction between trivial and non-trivial essentialism and the view that quantified modal logic is only committed to trivial essentialism. But the distinction can be made clear in various ways; for example, by saying that a logic or theory is trivially essentialist if $\forall x \varphi(x)$ is a theorem whenever the sentence $\exists x \Box \varphi(x)$ is (If $\varphi(x)$ contained the sole name a then $\varphi(a) \vee \forall x(x \neq a \supset \varphi(x))$ would be a theorem, and similarly for more names). Moreover, given this definition it would appear that many reasonable theories are trivially essential. Plantinga's world-based essences or essential properties are not really counter-instances, for it is clear that once the names implicit in the world-description are made explicit, the above condition is satisfied.

The final two chapters of the book deal with the problem of evil and the ontological argument. A detailed discussion of these topics is out of place in a review, but let me attempt to state some weak points in his arguments. In avoiding the problem of evil, he points out that God's omnipotence does not imply that he has the power to create every possible world. After all, he

cannot create a world in which he does not exist (assuming that there is such a world). This then leaves open the possibility that an omnipotent God can only create worlds that contain moral evil. Let us grant this, though it might be more plausible to say that the maximal possibilities God can bring about are not always possible worlds (complete possibilities). But there is still a problem of evil. For first, there is natural evil (which Plantinga attributes to demons or the like); and, second, would it not be better for the world to be so constituted that no man had the power to bring about the great evil that many men can and do bring about?

Plantinga's version of the ontological argument is stated in terms of possible worlds. A modal formulation of the argument runs as follows. Let $Me(x)$ stand for 'x is maximally excellent', $Mg(x)$ for 'x is maximally great', and $D(x)$ for 'x is omnicient, omnipotent, and omnigood'. Then the argument is:

(6) $\Diamond \exists x Mg(x)$
(7) $\Box \forall x (Mg(x) \supset \Box Me(x))$
(8) $\Box \forall x (Me(x) \supset D(x))$
∴(9) $\exists x D(x)$.

The argument is sound (at least, if $\Box \forall x \Box (Me(x) \supset Ex)$ is assumed), but are the premises correct? There are arguments against (6)–(8) being true. For they imply $Ex \Box D(x)$ and if it is necessary that any object satisfying $D(x)$ creates a good world then it is necessary that any possible world is good. There appear to be no particular arguments, however, *for* (6)–(8) being true.

This book is a substantial contribution to the study of modality. It isolates some key issues in this area and argues, with great vigour, for a position that combines an actualist ontology with a respect for the ordinary modal notions. Sometimes there is more vigour than rigour, but, in its central contentions, the book is always both clear and convincing.

References

ADAMS, R. M. [1974], 'Theories of Actuality', *Noûs* 8: 211–31; repr. in Loux [1979], 190–209.

—— [1981], 'Actualism and Thisness', *Synthese* 49: 3–42.

ALMOG, J. [1991], 'The What and the How', *Journal of Philosophy* 88/5: 225–44.

ANSCOMBE, E. [1975], 'The First Person', in S. Guttenplan (ed.), *Mind and Language* (Oxford: Clarendon), 45–65.

ARMSTRONG, D. M. [1983], *What is a Law of Nature?* (Cambridge: Cambridge University Press).

—— [1989], *A Combinatorial Theory of Possibility* (Cambridge: Cambridge University Press).

BACON, J., CAMPBELL, K., REINHARDT, L. [1993], *Ontology, Causality and Mind: Essays in Honor of D. M. Armstrong* (Cambridge: Cambridge University Press).

BAKER, J. R. [1978], 'Some Remarks on Quine's Arguments Against Modal Logic', *Notre Dame Journal of Formal Logic* 19: 663–73.

BALDWIN, T. [1999], 'Back to the Present', *Philosophy* 74: 177–97.

BARCAN MARCUS, R. [1990], 'A Backward Look at Quine's Animadversions on Modalities', in R. B. Barrett and R. F. Gibson (eds.), *Perspectives on Quine* (Cambridge, Mass.: Blackwell), 230–43; repr. in Barcan Marcus [1993], 216–32.

—— [1993], *Modalities: Philosophical Essays* (Oxford: Oxford University Press).

BIGELOW, J. [1991], 'Worlds Enough for Time', *Noûs* 25/1: 1–19.

BOOLOS, G. [1984] 'To Be is to be the Value of a Variable (or Some Values of Some Variables)', *Journal of Philosophy* 81: 430–50, repr. in *Logic, Logic, Logic* (Cambridge Mass.: Harvard University Press).

BOYD R. [1988], 'How to be a Moral Realist', in G. Sayre-McCord (ed.), *Essays on Moral Realism* (Ithaca: Cornell University Press); repr. in Darwall et al. [1997].

BRICKER, P. [1987], 'Reducing Possible Worlds to Language', *Philosophical Studies* 52: 331–55.

BRINGSJORD, S. [1985], 'Are There Set-Theoretic Possible Worlds?', *Analysis* 45/1.

BRINK, D. [1989], *Moral Realism and the Foundations of Ethics* (Cambridge: Cambridge University Press).

BROAD, C. D. [1938], *Examination of McTaggart's Philosophy* (Cambridge: Cambridge University Press).

—— [1959], *Scientific Thought*, (Paterson, NJ: Littlefield, Adams).

BROCK, S. [1993], 'Modal Fictionalism: a Reply to Rosen', *Mind* 102: 147–50.

CARNAP, R. [1947], *Meaning and Necessity* (Chicago: University of Chicago Press).

CARROLL, J. W. [1994], *Laws of Nature* (Cambridge: Cambridge University Press).

CHALMERS, A. [1999], 'Making Sense of Physical Laws', in Sankey [1999], 3–15.

CHIHARA, C. S. [1998], *The Worlds of Possibility* (Oxford: Clarendon).

CHISHOLM, R. [1967], 'Identity through Possible Worlds: Some Questions', *Noûs* 1: 1–8.

CHRISTENSEN, F. [1974], 'McTaggart's Paradox and the Nature of Time', *Philosophical Quarterly* 24: 289–98.

CHURCH, A. [1943], Review of Quine's 'Notes on Existence and Necessity', *Journal of Symbolic Logic* 8: 45–7.

—— [1951], 'A Formulation of the Logic of Sense and Denotation', in P. Henle, H. M. Kallen, and S. K. Langer (eds.), *Structure, Method and Meaning: Essays in Honor of Henry M. Sheffer* (New York: Liberal Arts Press).

CLIFTON, and HOGARTH, [1995], 'The Definability of Objective Becoming in Minkowski Space-Time', *Synthese* 203: 355–87.

CRAIG, W. L. [2001], *Time and the Metaphysics of Relativity* (Dordrecht: Kluwer).

CRESSWELL, M. J. [1990], *Entities and Indices* (Dordrecht: Kluwer).

CRISP, T. M. [2003], 'Presentism', in Loux and Zimmerman [2003], ch. 8, 211–45.

DARWALL, S., et al. [1997], *Moral Discourse and Practice: Some Philosophical Approaches* (Oxford: Oxford University Press).

DAVIDSON, D., and HINTIKKA, J. (eds.) [1969], *Words and Objections: Essays on the Work of W. V. O. Quine* (Dordrecht: Reidel).

DIVERS, J. [1995], 'Modal Fictionalism Cannot Deliver Possible Worlds Semantics', *Analysis* 55: 81–9.

DRIER, JAMES [1992], 'The Supervenience Argument against Moral Realism', *Southern Journal of Philosophy* 30/3: 13–38.

DUMMETT, M. [1960], 'A Defence of McTaggart's Proof of the Unreality of Time', *Philosophical Review* 69: 497–504; repr. in Dummett [1978] and in Levenson & Westphal [1993].

—— [1973], *Frege: Philosophy of Language* (London: Duckworth).

—— [1978], *Truth and Enigmas* (Cambridge, Mass.: Harvard University Press).

DYKE, H. [2002], 'McTaggart and the Truth about Time', *Philosophy* 50 (suppl.): 137–52.

ELLIS, B. [1999], 'Causal Powers and Laws of Nature', in Sankey [1999].

EVANS, G. [1985], 'Does Tense Logic Rest Upon a Mistake?' in *Collected Papers,* ed. A. Phillips (Oxford: Clarendon).

FALES, E. [1993], 'Are Causal Laws Contingent?', in Bacon, Campbell, and Reinhardt [1993].

FINE, K. [1970], 'Propositional Quantifiers in Modal Logic', *Theoria* 36/3: 336–46.

—— [1976], 'Review of *The Nature of Necessity* (A. Plantinga)', *Philosophical Review* 86: 562–66. (Ch. 11 of the present volume.)

—— [1977a], 'Postscript to *Worlds, Times and Selves*', from *Worlds, Times and Selves* (with A. N. Prior) (Duckworth: England). (Ch. 4 of the present volume.)

—— [1977b], 'Properties, Propositions and Sets', *Journal of Philosophical Logic* 6: 135–91.

—— [1978], 'Model Theory for Modal Logic Part I: The De Re/De Dicto Distinction', *Journal of Philosophical Logic* 7: 125–56.

—— [1980], 'First-order Modal Theories, II—Propositions', *Studia Logica* 39/2–3: 159–202.

—— [1981*a*], 'First-Order Modal Theories I—Sets', *Nôus* 15: 177–205.

—— [1981*b*], 'Model Theory for Modal Logic—Part III: Existence and Predication', *Journal of Philosophical Logic* 10: 293–307.

—— [1982], 'First-Order Modal Theories III—Facts', *Synthese* 53: 43–122.

—— [1984], 'Reference, Essence and Identity'. Unpublished manuscript.

—— [1985], 'Plantinga on the Reduction of Possibilist Discourse', in J. E. Tomberlin and P. van Inwagen (eds.), *Alvin Plantinga*, Reidel Profiles Series 5 (Dordrecht: Reidel), 145–186. (Ch. 5 of the present volume.)

—— [1989], 'The Problem of De Re Modality', in J. Almog et al. (eds.), *Themes from Kaplan* (Oxford: Clarendon), 197–272. (Ch. 2 of the present volume.)

—— [1994], 'Essence and Modality', *Philosophical Perspectives* 8: 1–16.

—— [1995], 'Senses of Essence', in W. Sinnot Armstrong, *Modality, Morality, and Belief* (Cambridge: Cambridge University Press), 53–73.

—— [1998*a*], 'A Defence of Cantorian Abstraction', *Journal of Philosophy* 95/12: 599–634.

—— [1998*b*], 'The Limits of Abstraction', in M. Schirn (ed.), *The Philosophy of Mathematics Today* (Oxford: Clarendon), 503–630.

—— [2000], 'A Counter-example to Locke's Thesis', *Monist* 83/3: 357–61.

—— [2001], 'The Question of Realism', *Philosophers Imprint* 1/1.

—— [2002], *The Limits of Abstraction* (Oxford: Oxford University Press).

—— [2003], 'The Non-identity of a Thing and its Matter', *Mind* 112: 195–234.

FORBES, G. [1985], *The Metaphysics of Modality* (Oxford: Oxford University Press).

—— [1989], *Languages of Possibility* (Oxford: Blackwell).

—— [1997], 'Belief Reports and Speech Reports', in W. Künne, A. Newen, and M. Anduschus, *Direct Reference*, CSLI Lecture Notes (Cambridge: Cambridge University Press), 313–30.

FORREST, P., and ARMSTRONG, D. M. [1984], 'An Argument Against David Lewis' Theory of Possible Worlds', *Australasian Journal of Philosophy* 62: 164–8.

FRIEDMAN, M. [1994], 'Geometry, Convention and the Relativized A Priori: Reichenbach, Schlick, and Carnap', in W. Salmon and G. Wolters (eds.), *Logic, Language, and the Structure of Scientific Theories* (Pittsburgh: University of Pittsburgh Press); repr. in P. Grim et al. (eds.) [1994].

GIBBARD, A. [1975], 'Contingent Identity', *Journal of Philosophical Logic* 4: 187–221.

GODFREY-SMITH, W. [1979], 'Special Relativity and the Present', *Philosophical Studies* 36: 233–44.

GÖDEL, K. [1944], 'Russell's Mathematical Logic', in P. A. Schilpp (ed.), *The Philosophy of Bertrand Russell* (Evanston, Ill.: Northwestern University Press); repr. in *Kurt Gödel, Collected Works* (Oxford: Oxford University Press, 1990), ii.

—— [1949], 'A Remark About the Relationship between Relativity Theory and Idealistic Philosophy', in P. A. Schilpp (ed.), *Albert Einstein: Philosopher-Scientist* (La Salle, Ill.: Open Court), repr. in Kurt Gödel, *Collected Works* (Oxford: Oxford University Press, 1990).

Goodman, N. [1978], *Ways of Worldmaking* (Cambridge: Hackett).

Grim, P. [1986], 'On Sets and Worlds: A Reply to Menzel', *Analysis* 46: 186–91.

—— et al. [1994], *The Philosopher's Annual*. 18/94 (Atascadero, Calif.: Ridgeview).

Gupta, A. [1980], *The Logic of Common Nouns* (New Haven, Conn.: Yale University Press).

Hale, B. [1995], 'Modal Fictionalism—A Simple Dilemma', *Analysis* 55: 63–7.

Hinchliff, M. [2000], 'A Defense of Relativism in a Relativistic Setting', *Philosophy of Science* 67 (Proceedings), S575–86.

Hintikka, K. J. J. [1956], 'Identity, Variables and Impredicative Definitions', *Journal of Symbolic Logic* 21: 225–45.

Horwich, P. [1989], *Asymmetries in Time* (Cambridge, Mass.: MIT); repr. in Levenson and Westphal [1993].

Jackson, F. [1998], *From Metaphysics to Ethics: A Defence of Conceptual Analysis* (Oxford: Oxford University Press).

Kamp, H. [1971], 'Formal Properties of "Now" ', *Theoria* 37: 227–73.

Kaplan, D. [1969], 'Quantifying In', in Davidson and Hintikka [1969]; repr. in Linsky [1971].

—— [1972] 'What is Russell's Theory of Descriptions', in D. F. Pears (ed.), *Bertrand Russell* (New York: Anchor Books).

—— [1975], 'How to Russell a Frege-Church', *Journal of Philosophy* 72: 716–29; repr. in Loux [1979], 210–24.

—— [1979], 'Transworld Heir Lines', in Louix [1979], 88–109.

—— [1986] 'Opacity', in L. E. Hahn and P. A. Schilpp (eds.), *The Philosophy of W. V. Quine* (La Salle, Ill.: Open Court).

—— [1989], 'Demonstratives', in J. Almog et al. (eds.), *Themes from Kaplan* (Oxford: Clarendon), 481–564.

—— [1995], 'A Problem in Possible-World Semantics', in W. Sinnot-Armstrong (ed.), *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus* (Cambridge: Cambride University Press).

Keller, S. [2003], 'Presentism and Truthmaking', in Zimmerman [2003].

Klagge, J. [1984], 'An Alleged Difficulty Concerning Moral Properties', *Mind* 93: 370–80.

Kripke, S. [1963], 'Semantical Considerations on Modal Logic', *Acta Philosophica Fennica* 16: 83–94; repr. in Linsky [1971].

—— [1980], *Naming and Necessity* (Cambridge, Mass.: Harvard University Press).

Levenson, C., and Westphal, J. (eds.) [1993], *Time* (Indianapolis: Hackett).

Lewis, D. [1973], *Counterfactuals* (Oxford: Blackwell).

—— [1979], 'Attitudes De Se and De Dicto', *Philosophical Review* 88: 513–43.

—— [1986], *On the Plurality of Worlds* (Oxford: Blackwell).

—— [1989], 'Dispositional Theories of Value—II', *Proceedings of the Aristotelian Society*, Suppl., 113–37.

Lewy, C. [1976], *Meaning and Modality* (Cambridge: Cambridge University Press).

Linsky, L. (ed.) [1952], *Semantics and the Philosophy of Language* (Urbana: University of Illinois Press).

—— [1967], *Referring* (New York: Humanities Press).

—— (ed.) [1971], *Reference and Modality* (Oxford: Clarendon).

LOUX, M. (ed.) [1979], *The Possible and The Actual* (Ithaca, NY: Cornell University Press).

LOWE, E. J. [1992], 'McTaggart's Paradox Revisited', *Mind* 101: 323–6.

—— [2001], *Kinds, Essence, and Natural Necessity* (forthcoming).

MCFETRIDGE, I. [1985], 'Supervenience, Realism, Necessity', *Philosophical Quarterly*, 35: 245–58.

MCMICHAEL, A. [1983], 'A Problem for Actualism about Possible Worlds', *Philosophical Review* 92: 49–66.

MCTAGGART, J. M. E. [1908], 'The Unreality of Time', *Mind* 17: 457–74.

—— [1927], *The Nature of Existence* (Cambridge: Cambridge University Press).

MALAMENT, D. [1977], 'Causal Theories of Time and the Conventionality of Simultaneity', *Noûs* 11: 293–300.

MARKOSIAN, N. [2002], 'A Defense of Presentism' in Zimmerman [2003].

MELLOR, D. H. [1974], 'Special Relativity and the Present Truth', *Analysis* 34: 74–8.

—— [1986], 'Time's Tenseless Truth-Conditions', *Analysis* 46: 167–72.

—— [1998], *Real Time II* (London: Routledge).

MENZEL, C. [1986a], 'On Set-theoretic Possible Worlds', *Analysis* 46: 68–72.

—— [1986b], 'On the Iterative Explanation of the Paradoxes', *Philosophical Studies* 49: 37–61.

MONTAGUE, R. [1970], 'Universal Grammar', *Theoria* 36: 373–98; repr. in R. H. Thomason (ed.), *Formal Philosophy* (New Haven: Yale University Press, 1974).

MOORE, A. W. [1997], *Points of View* (Oxford: Clarendon).

MOORE, G. E. [1922], *Philosophical Studies* (London: Routledge & Kegan).

NAGEL, T. [1983], 'The Objective Self', in C. Ginet and S. Shoemaker (eds.), *Mind and Knowledge* (Oxford: Clarendon).

—— [1986], *The View from Nowhere* (Oxford: Clarendon).

NOLAN, H., and HAWTHORNE, J. O. [1996], 'Reflexive Fictionalism', *Analysis* 56: 23–32.

NOONAN, H. [1994], 'In Defence of the Letter of Fictionalism', *Analysis* 54: 133–9.

PARSONS, T. [1969], 'Essentialism and Quantified Modal Logic', *Philosophical Review* 78: 35–52; repr. in Linsky [1971].

PEACOCKE, C. [2002], 'Principles for Possibilia', *Noûs* 36/3: 486–508; also to appear in A. O'Hear (ed.), *Logic, Thought and Language*, Royal Institute of Philosophy, suppl. 51 (Cambridge: Cambridge University Press), 155–81.

PERCIVAL, P. [1989], 'Indices of Truth and Temporal Propositions', *Philosophical Quarterly* 39: 190–7.

—— [1991], 'Knowability, Actuality, and the Metaphysics of Context-Dependence', *Australian Journal of Philosophy* 69: 82–97.

—— [2002], 'A Presentist's Refutation of Mellor's McTaggart', *Philosophy* 50 (suppl.), 91–118.

PLANTINGA, A. [1974], *The Nature of Necessity* (Oxford: Clarendon).

—— [1976], 'Actualism and Possible Worlds', *Theoria* 42: 139–60; repr. in Loux [1979], 253–73.

—— [1979], 'De Essentia', in E. Sosa (ed.), *Essays on the Philosophy of R. M. Chisholm* (Amsterdam: Rodopi), 101–21.

PLANTINGA, A. [1983], 'On Existentialism', *Philosophical Studies* 44: 1–20.

PRIOR, A. N. [1957], *Time and Modality* (Oxford: Clarendon).

—— [1967], *Past, Present, and Future* (Oxford: Clarendon).

—— [1968a], *Papers on Time and Tense* (Oxford: Clarendon).

—— [1968b], 'Egocentric Logic', *Noûs* 2: 101–19; repr. as ch. 2 of Prior and Fine [1977a] and Prior [2003].

—— [1968c], 'Modal Logic and the Logic of Applicability', *Theoria* 34.

—— [1969a] 'Worlds, Times and Selves', *L'Age de la Science*, 3: 179–91; repr. in Prior and Fine [1977].

—— [1969b], 'The Possibly True and the Possible', *Mind* 78: 481–92.

—— [1969c], 'Tensed Propositions as Predicates', *American Philosophical Quarterly* 6: 290–7; repr. in Prior [2003], ch. 15.

—— [2003], *Papers on Time and Tense* (new edition), ed. P. Hasle et al. (Oxford: Clarendon).

—— and FINE, K. [1977], *Worlds, Times and Selves* (London: Duckworth).

PUTNAM, H. [1967a], 'The Thesis that Mathematics is Logic', in R. Schoenman (ed.), *Bertrand Russell: Philosopher of the Century* (London: Allen & Unwin), 273–303.

—— [1967b], 'Time and Physical Geometry', *Journal of Philosophy* 64: 240–7.

QUINE, W. V. O. [1943], 'Notes on Existence and Necessity', *Journal of Philosophy* 40: 113–27; repr. in Linsky [1971].

—— [1947], 'The Problem of Interpreting Modal Logic', *Journal of Symbolic Logic* 12: 43–8.

—— [1953], 'Reference and Modality', in Quine [1963].

—— [1956], 'Quantifiers and Propositional Attitudes', *Journal of Philosophy* 53; repr. in Quine [1966] and Linsky [1971].

—— [1960], *Word and Object* (Cambridge, Mass.: MIT).

—— [1963], *From a Logical Point of View* (New York: Harper & Row; originally printed by Harvard University Press [1953]).

—— [1964], 'Ontological Reduction and the World of Numbers', *Journal of Philosophy* 61: 209–16; repr. in Quine [1966].

—— [1966], *The Ways of Paradox* (New York: Random House).

—— [1969], *Ontological Relativity* (New York: Columbia University Press).

—— [1976], 'Worlds Away', *Journal of Philosophy* 73: 859–63.

—— [1977], *Intensions Revisited*, Midwest Studies in Philosophy 2 (Minneapolis: University of Minnesota Press).

RAILTON, P. [1986], 'Moral Realism', *Philosophical Review* 163–207; repr. in Darwall [1997].

RAKIC, N. [1997], 'Past, Present, Future and Special Relativity', *British Journal for the Philosophy of Science* 48: 257–80.

RICHARDS, B. [1974], 'A Point of Reference', *Synthese* 28: 361–454.

ROSEN, G. [1990], 'Modal Fictionalism', *Mind* 99: 327–54.

—— [1993], 'A Problem for Fictionalism about Possible Worlds', *Analysis* 53: 71–81.

—— [1995], 'Modal Fictionalism Fixed', *Analysis* 55: 67–73.

RUSSELL, B. [1956], 'The Philosophy of Logical Atomism' in R. C. Marsh (ed.), *Logic and Knowledge* (London: Allen & Unwin).

SANKEY, H. [1999], *Causation and Laws of Nature* (Dordrecht: D. Reidel).

SAVITT, S. [2000], 'There's No Time Like the Present (in Minkowski Spacetime)', *Philosophy of Science* 67 (Proceedings), S563–74.

SCHLESINGER, G. N. [1991], 'E Pur Si Muove', *Philosophical Quarterly* 41: 427–41.

—— [1993], 'A Short Defense of Temporal Transience', *Philosophical Quarterly* 43: 359–61.

—— [1995], *Timely Topics* (London: Macmillan).

SEGERBERG, K. [1970], 'Some Modal Logics Based on a Three-Valued Logic', *Theoria* 36: 301–21.

SHOEMAKER, S. [1980], 'Causality and Properties', in P. van Inwagen (ed.), *Time and Cause* (Dordrecht: D. Reidel); repr. in *Identity, Cause and Mind: Philosophical Essays* (Cambridge: Cambridge University Press, 2003).

—— [1985], Review of S. Blackburne's 'Spreading the Word', *Noûs* 18/3: 438–42.

—— [1998], 'Causal and Metaphysical Necessity', *Pacific Philosophical Quarterly* 79: 59–77.

SHORTER, J. M. [1984], 'The Reality of Time', *Philosophia* 14: 321–39.

SIDELLE, A. [2002], 'On the Metaphysical Contingency of Laws of Nature', in T. Szabo and J. Hawthorne (eds.), *Conceivability and Possibility* (Oxford: Clarendon), ch. 8.

SIDER, T. [1999], 'Presentism and Ontological Commitment', *Journal of Philosophy* 96: 325–47.

—— [2001], *Four-Dimensionalism: An Ontology of Persistence and Time* (Oxford: Clarendon), 42–52 for SR.

—— [2002], 'The Ersatz Pluriverse', *Journal of Philosophy* 99: 279–315.

SKLAR, L. [1981], Time, Reality and Relativity in R. Healy (ed.), *Reduction, Time and Reality* (New York: Cambridge University Press), 129–42.

SMITH, Q. [1993], *Language and Time* (New York: Oxford University Press).

SMITH, M. [1994], *The Moral Problem* (Oxford: Blackwell).

SPRIGGE, T. L. S. [1992], 'The Unreality of Time', *Proceedings of the Aristotelian Society* 92: 1–19.

STALNAKER, R. C. [1976], 'Possible Worlds', *Noûs* 10: 65–75; repr. in Loux [1979], 225–34 and Stalnaker [2003], ch. 1.

—— [2003], 'On Thomas Nagel's Objective Self', in *Ways a World Might Be* (Oxford: Clarendon).

STEIN, H. [1968], 'On Einstein-Minkowski Space-Time', *Journal of Philosophy* 65: 5–23.

—— [1970], 'A Note on Time and Relativity Theory', *Journal of Philosophy* 67: 289–94.

—— [1991], 'On Relativity Theory and the Openness of the Future', *Philosophy of Science* 58: 147–67.

STURGEON, N. L. [1984], 'Moral Explanations', in D. Copp and D. Zimmerman (eds.), *Morality, Reason and Truth: New Essays in the Foundations of Ethics* (Totowa, NJ: Rowman & Littlefield).

SWOYER, C. [1982], 'The Nature of Causal Laws', *Australasian Journal of Philosophy* 60: 203–23.

TAYLOR, B. [1997], 'Dummett's McTaggart' in R. Heck (ed.), *Language, Thought and Logic* (Oxford: Clarendon) 183–99.

THOMASON, R. H. [1979], 'Home is Where the Heart Is', in P. French, T. Uehling, and H. Wettstein (eds.), *Contemporary Perspectives in the Philosophy of Language* (Minneapolis: University of Minnesota Press).

THOMSON, J. J. [2001], 'McTaggart on Time', *Noûs* (suppl.) 15: 229–52.

TOMBERLIN, J. E. and VAN INWAGEN, P. (eds.) [1985], *Alvin Plantinga*, Reidel Profiles (Dordrecht: Reidel).

TOOLEY, M. [1997], *Time, Tense and Causation* (Oxford: Clarendon).

TWEEDALE, M. [1984], 'Determinable and Substantival Universals', in Radu J. Bogdan (ed.), *D. M. Armstrong*, Profiles 4 (Dordrecht: Reidel).

VAN FRAASSEN, B. [1977], 'The Only Necessity is Verbal Necessity', *Journal of Philosophy* 74/2: 71–85.

VAN INWAGEN, P. [1980], 'Indexicality and Actuality', *Philosophical Review* 89: 403–206.

—— [1981], 'Why I Don't Understand Substitutional Quantification', *Philosophical Studies* 39: 281–5.

VLACH, F. [1973], ' "Now" and "Then": A Formal Study in the Logic of Tense Anaphora', Ph.D. Dissertation, UCLA.

WIGGINS, D. [1980], *Sameness and Substance* (Cambridge, Mass.: Harvard University Press).

—— [1987], *A Sensible Subjectivism*?, Aristotelian Society Series 6 (Oxford: Blackwell).

WILLIAMSON, T. [1998], 'Bare Possibilia', *Erkenntnis* 48: 257–73.

—— [2000], 'Necessary Existents', in A. O'Hear (ed.), *Logic, Language, and Thought* (Cambridge: Cambridge University Press).

ZANGWILL, N. [1996] 'Moral Supervenience', *Midwest Studies in Philosophy*, 20: 240–62.

ZIMMERMAN, D. (ed.) [2003], *Oxford Studies in Metaphysics* (Oxford: Clarendon), i.

Index