# Methapors & Modality

James W. GARSON

## 1. *Introduction*

The purpose of this paper is to describe the formal and philosophical merits of an approach to quantified modal logic (QML) that doesn't require that terms be rigid designators.

Though the controversies of the sixties tended to focus on whether QML made sense in the first place, the more recent consensus seems to be that QML is possible, though perhaps not yet actual. The battles with Quine and his followers, however, did have the positive effect of making clear how deeply issues such as essentialism, transworld identification, and the distinction between *de re* and *de dicto* are intertwined with the development of a philosophical foundation for QML.

By the early seventies there were a host of approaches to QML, and their variety and complexity seemed to offer little hope that they would be of any use to the philosopher with but a passing interest in technical matters. In fact, without any philosophical underpinnings to motivate the application of these systems, it wasn't entirely clear that they were *modal* logics.

Kripke's 'Naming and Necessity' championed the first widely accepted philosophical account which motivates a QML. His remarks there sketch a framework which complements the formal structure of the systems he presented in 'Semantical Considerations in Modal Logic', systems which rely on the rigid designator treatment of terms and quantification. This formal choice colors Kripke's entire philosophical outlook. It is central to his 'resolution' of the problem of transworld identification, to his insistence that if anything, it is *de re*, rather than *de dicto* modality that is more perspicuous, to his sharp distinction between the world as it is known and the world as it is, and to his demythologization of the telescope metaphor.

The philosophical community has, no doubt, been relieved at the way Kripke has drawn order out of chaos. His philosophical views are beautifully tailored to his formal structures. The danger is that we will

assume that the neat fit of the philosophical clothing is a sign that Kripke's way of doing things is best.

In this paper we will show that Kripke's methods are parochial. They are not wrong exactly; it is just that alternatives are more general, and accommodate a wider range of philosophically interesting views. From this perspective, many of Kripke's philosophical positions (especially his sharp distinction between the metaphysical and the epistemological, and his modal realism) are burdens his theory of QML must bear, and not the ultimate consequences of the use of modal semantics which they are often taken to be.

We will show that a sensible philosophical account of a QML can be given without assuming that terms are rigid designators, and that it serves as a stronger and more flexible tool for dealing with problems concerning naming and necessity. Given that we want to avoid a logic which builds in, or strongly suggests, central and debatable philosophical assumptions, the more general approach should be preferred.

## 2. *Kripke's rigid designator theory*

In modal semantics, we assign a sentence not a truth value straight off, but an *intension*, which is a function from a set W of possible worlds to truth values that allows us to calculate the truth value of the sentence in each world. Similarly, an n-ary predicate letter is assigned an intension, which is a function from W into the set of extensions of predicate letters, (the extensions being sets of n-tuples of objects). This treatment of sentences and predicate letters is fairly standard in modal semantics. However, when it comes to the treatment of terms,([1]) we are confronted with an important choice. It is fairly clear that the extensions of predicate letters and sentences should vary from possible world to to possible world, but should the same be true for terms? We could simply maintain the parallel with our treatment of sentences and predicate letters by assign-

([1]) By 'term' we mean an expression of the formal language which can be substituted for free variables, to yield a well formed formula of the syntax. Strictly speaking, descriptions in Russell's theory of descriptions are not treated as terms since the notation '$\phi$Ix$\psi x$' (though it gives the appearance of the result of substituting Ix$\psi$x for y in $\phi$y) is in fact shorthand for something else.

ing terms functions from W into the set O of possible objects. These functions were called *individual concepts* by Carnap.

Kripke, however, chose not to maintain the parallel; he claimed instead that terms should be assigned objects directly. His intuition was that the referent of a term should remain the same regardless of the choice of possible world. He also realized that his theory would suffer from formal complications if he didn't insure that the range of reference of his terms, and the domain of quantification, include the same type of things. A theory, then, which chose individual concepts for terms would seem to require that the domain of quantification include individual concepts. Given that quantifers carry ontological commitment, it seemed foolish to assign individual *concepts* to terms. (2)

Following the tradition, we will call Kripke's approach the *rigid designator theory of terms* (or RD *theory* for short). But we should be warned that the expression 'rigid designator' is misleading. The interpretation of 'rigid' here that fits with the formal theory is that a rigid term doesn't 'change' its referent from one possible world to the next. This is a bit different from what we might take it to mean. A rigid designator could be a locution whose *intension* is a constant function, a function that picks out something for each world in such a way that all things picked out are identical. Now it makes no *formal* difference whether we assign terms objects directly as Kripke did, or assign them constant functions. Mathematically, the constant function with value 0, and 0 itself, are indistinguishable. However, there is a psychological difference, for the two readings lead to different pictures of what the formal theory is saying. Kripke's treatment of terms and their reference is simply classical. The relationship between terms and object is described exactly as in extensional logic, and so not especially problematic.

However, if we take the constant function picture, we encounter all sorts of worries, for here we have a function that picks out the *same* objects

(2) We should point out that given Kripke's treatment of terms, there are actually two choices to make concerning the domain of quantifiers. On the fixed-domain interpretation, the domain of quantification is the same for all possible worlds and consists of the set O of all possible objects. On the roving domain shifts from one possible world to another and consists in the possible objects that happen to exist at a given world. The choice between these two interpretations of the quantifiers, while interesting in other contexts, is not central to the points being made in this paper.

in each possible world, and this prompts the question of how to make sense of the idea of identity of objects across possible worlds.

One way to look at Kripke's discussion of transworld identity in part I of 'Naming and Necessity' is to think of him as insisting that we use the picture that goes with the direct-reference-to-object technique, while heading us off from tendencies we might have to adopt a constant function picture. (Unfortunately his use of the expression 'rigid designator' sways us in the opposite direction.)

Kripke points out that when we tell the story about the truth conditions of (1), we should *not* imagine that we are

(1) Muskie could have been president

looking into each world, trying to recognize Muskie (as if we had epistemological access to possible worlds in the first place!), for this prompts us to search for some set of observable and fully individuating properties with which to pick him out, and so determine his identity with the Muskie in the real world. We shouldn't do this because it raises pointless pseudo-epistemological questions. Instead we should imagine that 'Muskie' refers directly to an object in the ordinary way, and that the truth of (1) depends on whether the very object so designated has the property of being president in some situation. Since our ability to refer does not, typically, require that we be able to come up with any perfectly individuating set of observable properties of the object of reference, it is perfectly legal for us to refer to Muskie without knowing of any properties which would uniquely identify him in other situations.

Now I believe that this is the right tactic for handling the 'problem' of transworld identification. But though it may not be apparent from this description, we can employ roughly the same strategy for a semantics where terms are not rigid. Once we explain how to fashion a metaphor which is an accurate, rather than misleading, reflection of the formal machinery, the 'problem' of transworld identification cannot even be expressed.

## 3. *A Short History of Quantified Modal Logics*

One of the reasons Kripke chose the rigid designator theory of terms (RD theory) may have been that he suspected the existence of the formal

problems that turned up in individual concept theories (IC theories). We now know, however, that those formal problems can be overcome, and so it is worthwhile to explain how the recent history of QML has vindicated the IC approach.

Kripke's systems had one major drawback. (Kripke would call it a feature.) Since terms are rigid, it follows that

$$t = t' \supset \Box\, t = t'$$

is valid, which means we are commited to the view that any true term identity is a necessary identity. This automatically blocks any attempt to treat descriptions as terms in the logic, for descriptions are non-rigid designators, hence identities involving them are not necessary if true. For example, though Benjamin Franklin was the inventor of bifocals, it is not the case that he was necessarily the inventor of bifocals, since we can imagine situations where 'the inventor of bifocals' picked out somebody other than Franklin.

It follows that any QML with rigid terms will have to employ Russell's theory of descriptions, and so deny the 'termhood' of expressions of ordinary language which have initially all the semblence of terms. The interest in developing logics with a uniform treatment of the terms of ordinary language prompted some, notably Thomason and Hintikka to investigate semantics for formal language where terms are assigned individual concepts. ([3]) Thomason [13] describes three semantics Q1, Q2 and Q3, which represent what seem to the main options. Q1 is pretty much Kripke's system. It has rigid terms and quantification over a domain of objects. Q2 represents the most uniform alternative: terms are assigned individual concepts, and the domain of quantification is the set of all individual concepts constructable from the set W of possible worlds and the set O of objects. Although this semantics is quite pleasing and straightforward, there is a major formal obstacle. As long as the strength of the underlying modal logic is S4.3 or less, no set of axioms and/or rules can be given which picks out exactly the valid formulas of Q2. Because the domain of quantification is the set of *all* individual concepts, it turns out that the language has the expressive power of second order

([3]) See THOMASON [13], and HINTIKKA [7].

arithmetic, and this means that Godel's incompleteness result applies to it.([4])

So Thomason turned his back on Q2, remarking that the trouble with it is that it lacks a concept of substance. Since the domain of quantification includes all the individual concepts, no distinctions are made between the ICs for Benjamin Franklin (presumably a substance, or unified thing across possible worlds) and the inventor of bifocals (presumably a motley assortment of people across possible worlds).

He then investigated a system (similar to Hintikka's) where he 'cuts down' the domain of quantification to just those functions which represent 'uniform things', rather than 'collages'. His method for doing this is severe, for the domain of quantification is restricted to constant functions only. To put it in an equivalent way, the members of the domain of quantification are possible objects. Q3 than, is a hybrid system which assigns ICs to terms, but which quantifies over objects. The strain between the differing treatments of the terms and the quantifiers results in a complicated set of rules and axioms. Though Thomason showed that a formal system for Q3 can be constructed given the simplifying assumption that the underlying modal logic is as strong as S4, the complexity of the rules and those of similar systems due to Hintikka has left them unpopular.

So it seemed that on purely formal grounds that the rigid designator theory of terms was the only viable candidate for a simple QML. However, the work of Garson [5] shows that there is another way out([5]) For one thing, if we formulate Q2 with the substitution interpretation rather than the objectual interpretation of the quantifiers, the problem of the incompleteness of Q2 disappears. Not only that, the formal system that axiomatizes *this* semantics is simple. In fact, it is virtually Kripke's system with the difference that we lack

$$t = t' \supset \Box\, t = t',$$

([4]) For a detailed account of these results for Q2 see KAMP [8]. Similar results for related systems appear in GARSON [3].

([5]) I is worth pointing out that Bressan [1] develops a system of QML with quantification over individual concepts for a semantics similar to Q2, but unlike it on one respect. The *extension* of a one-place predicate letter (for example) at a world is a set of individual concepts, not objects.

In this system, the distinction between *de re* and *de dicto* applications of modality requires special technical devices. For a criticism of this approach see GARSON [4].

the bothersome assertion that term identities are necessary if true.

The use of the substitution interpretation is not necessary for rescuing Q2 from incompleteness. It is easy to see that the same system is obtained by adopting the objectual interpretation of the quantifiers with a domain of quantification which is some *subset* of the set of all individual concepts. This subset needn't be the set of all constant functions as Thomason supposed. The semantics merely says that a set theoretical model must make *some* distinction between the domain of quantification and the set of all individual concepts, while making no demands on how the distinction is made. This parallels the way models are defined in classical logic. The domain of quantification for a classical model can be any non-empty set we choose, and is not determined in any further way.

So a simple IC semantics can be developed as long as we are willing to admit that it makes some sense to distinguish certain of the ICs for membership in the domain of quantification. Though we must introduce a concept of 'substance' to avoid formal problems, we do not need to draw the distinction in any particular way. If we want to obtain Thomason's system Q3 within this approach, we can, by laying down the condition that the set of 'substances' is the set of all constant functions. However, the lesson of the formal results is that simplicity lies in the more general treatment of the domain of quantification.

## 4. *Interlude: The Role of Metaphors in Modal Semantics*

Even though formal obstacles to IC theories have been overcome, we still need to show how to deal with two major philosophical problems: the nature of possible objects, and transworld identification of objects. If we are to understand how these issues are to count for or against the IC theory, we should have an account of the relationships between the purely technical machinery, and the metaphors we use to flesh it out.

The whole language of possible worlds, possible objects, and identity across worlds is no part of the purely formal account, but instead motivated its use. The situation here is something similar to what we find in the hard sciences. Theories of physics, for example, depend at the formal level on sets of mathematical equations, which, taken by themselves, have neither application nor motivation. Once we are able to set up appropriate correspondences between the equations and measurable

phenomena, we have a way of verifying and using a given theory. Given a metaphor, or picture, or model, of the equations, we feel we have a better grasp on what they are saying about the world. This helps us not only to develop new extensions of the theory, but to understand in a deeper way the relationships between the formal calculus and the world we experience.

In theories of modal logic, the set theoretical semantics plays the role of the mathematical framework, our intuitions about the norms of modal reasoning correspond to the evidence, and the metaphor of possible worlds serves to motivate our application of the theory to the problems of modal reasoning. The metaphor is perhaps more crucial to the use of set theoretical semantics than are models in science.

Philosophical criticism of modal theories has tended to focus on the metaphor, not on the formal calculus. It is as if the critics were attacking a scientific theory, not because it predicts the wrong things, but because the model presents a misleading or incoherent picture of the world. One simple way to avoid criticism of modal metaphors is to adopt a brand of modal 'instrumentalism', and claim that the metaphor was never intended to describe reality. The only questions worth asking, according to this line, is the success of the theory at predicting the norms of modal reasoning, and the heuristic value of the metaphor in developing useful extensions of the theory.([6]) However, for the purposes of the present discussion, we won't try to resolve the issue between modal realism and instrumentalism. Even if modal realism wins the day, there are still a number of ways in which an attack on the metaphors for modal semantics might be misguided. Two of these can be illustrate by considering analogous misguided attacks on models in physics.

One way we can go wrong is to criticize the model for properties which are not relevant to, or are incompatible with, the calculus which is being interpreted. For example, one might complain of the billiard ball model in statistical mechanics that struck billiard balls eventually slow down, while gasses do not spontaneously loose pressure (which is what would be expected if the molecules of a gas were to slow down). The proper reply is to point out that the friction caused by rolling billiard balls is not a relevant feature of the model used in statistical mechanics. To put it another way, the objector has gotten the model wrong, since on it, the

([6]) This position is spelled out in MORTON and MONDADORI [11].

'billiard balls' are supposed to undergo perfectly elastic collisions. We will call this sort of misguided complaint an attack on a fruitless feature of the metaphor.

A second way one can go wrong is to misdiagnose the connections between the primitive concepts of the formal theory and the metaphor. For example, someone might argue that the billiard ball model for statistical mechanics is incoherent on the grounds that mass-points, which are the fundamental 'objectlike' primitives of mechanics, cannot be said to have any 'real' size or mass. If one were to respond that having a finite mass and size are fruitless features of the metaphor, then our critic may direct his attack more on the ontological seriousness of the theory than on the model. The world, he complains, just couldn't be made up of things without definite mass or extension, because for one thing, we could have no epistemological access to 'objects' like that. (No matter how well you build your microscope, you will never be able to resolve a mass point.)

It isn't too hard to answer this kind of criticism. The molecules (or balls) of the model for statistical mechanics are not to be identified with the mass-points of mechanics, but rather with assemblies (of a superdenumerable number) of mass-points feeling forces that connect them rigidly. Mathematically one goes from properties of mass-points to properties of molecules by integrating over functions that assign density values to points of space. When this is done, we have mathematical representations, so to speak, of the molecules of our model, and these representations have features that correspond to the finite masses and sizes that we normally associate with objects.

So we see that quite unlike models in mathematics, models of science need not do the job of interpreting the *primitives* of the formal theory. More typically, they model the formal behavior of complexes. When it comes to determine how to apply the theory, what its ontological impact is (if it has any), and what our epistemological relationship is to the world it describes, we focus our attention on these complexes (which correspond to the central concepts of our model), not the primitives.

The scientist's typical attitude towards mass-points is that they are abstractions from the things (molecules) that are really there. They are central, nevertheless, to the formal techniques that make the theory work.

It is not essential to this paper that the account we have given of the use of models in physics is correct. The purpose of the discussion was to set the stage for diagnosing what is wrong with certain criticisms of

the possible worlds metaphor. I will show that some complaints against the IC theory attack fruitless features of the metaphor, and others misdiagnose the relationship between the primitives of modal semantics (the possible worlds and objects) and the world which we take to be real. This will serve as our main tactic for relieving ourselves from the philosophical pressures that allegedly compel us to adopt the rigid designator approach.

## 5. *The Telescope Metaphor and the Problem of Transworld Identity*

One of the supposed advantages of the rigid designator theory is that the problem of transworld identification is avoided using the direct-reference-to-object treatment of terms. Because reference is direct, we escape the temptation to think of the possible worlds as if they were different planets with their own special inhabitants. It is this spatialization of the possible worlds metaphor that prompts the worry about how the same object could inhabit different worlds, and how we can identify (different?) objects across worlds. Kripke has done us the service then of showing us that spatialization of the possible worlds is not a fruitful feature of the metaphor. It is not required by this semantics, and it merely raises problems where none need exist.

But what Kripke didn't say is that the same tactic not only makes room for the RD theory of terms, it paves the way for the IC theory as well. To help see this, let us review what a modal semantics using the IC theory of terms says about how we determine the truth conditions of (1).

(1) Muskie could have been president

We will assume that (1) is read *de dicto*, for it is only then that the problem of transworld identification arises. The theory sasys that (1) is true in world w just in case the following is true of all worlds v suitably related to w: that what is assigned to 'Muskie' at v is a member of the extension of 'is president' at v. Notice that nothing is said here about looking for the *thing* Muskie-in-v. The phrase 'in possible world v' does not attach to Muskie-the-object, but to the determination of what is assigned to 'Muskie'. In the RD-theory, we have the luxury of not needing the qualification 'in v', but even in the IC theory, the 'in v' does not attach to the object Muskie.

While it is more convenient to use the material mode in talking about the truth conditions of (1), and so say (2),

(2) Muskie in v is the president in v for some world v accessible from w

we must remember that (2) is shorthand for an expression that presupposes nothing about object's 'residence' in the different possible worlds. If we resist the temptation which (2) provides to assume that objects inhabit possible worlds, then we cannot even formulate the problem of transworld identification; we cannot do so because that problem is supposed to be about how we can identify objects (which are) in different worlds. If we attempt to pose the problem in the formal mode as we should, it becomes *not* the ontological question 'Which objects in a world are identical to which in another?", nor the epiostemological question 'How do I go about finding out which objects are identical in different worlds?", but a question about *reference*,([7]) a question about what is assigned to 'Muskie' for a world. If, as Kripke claims, we should take the notion of reference as primitive when we formulate a semantics, then no philosophical difficulties about the semantical account arise in evaluating the truth conditions of (1). The only reason people thought there was a problem of transworld identification was that they employed a metaphor which views possible worlds as spread out in space, inhabited by objects. This metaphor was prompted by a misinterpretation of what the formal theory says, and in particular by the failure to see that it is *reference*, not *objects* that is indexed by possible worlds in the semantics. Since this metaphor is not required to interpret the theory, and in fact distorts what it says, the charge that transworld identification is incoherent has all the force of the charge that statistical mechanics is incoherent because billiard balls slow down.([8])

([7]) Plantinga's main strategy [12] in defusing the TWI problem is by our account misguided. The main tactic is to dismiss the epistemological question 'How do I find the thing identical to this thing in that other world', and to postulate a metaphysical fact of the matter, so our inability to state the conditions for identity does not show that such conditions aren't there. If we are right, the problem of TWI does not even need this metaphysical rescue operation (though one might propose it for other reasons), for the problem in determining the truth conditions of (1) is not even about what is there, it is about reference.

([8]) So far we have proceeded without mentioning the 'roving domain' interpretation of the quantifiers. Here there is a domain of quantification $D_w$ for each world w. Surely here we have the need to talk of objects inhabiting (or at least existing in) different worlds. That

## 6. *Fixing Reference and Individual Concepts*

The reader may feel that I have merely swept the problem of transworld identity under the rug of reference. Granted that the semantical account of the truth conditions of (1) does not require that we fix identities across possible worlds, we still need to be able to fix the assignment of 'Muskie' for each world. But doing so would seem to raise exactly the same philosophical problems that are discussed under the heading of transworld identity. How are we to determine what the referent of 'Muskie' ought to be for a world? To answer this, we will have to know how to pick out the right object (not *in* but) *for* each world, and surely this is just a disguised version of the problem of determining which object in another world corresponds to Muskie in this one.

We can handle this response by reiteating what Kripke said in 'Naming and Necessity'. We reply that it is perfectly possible to refer successfully without having a complete, accurate, and individuating description of what it is that we are referring to. So the worry about how to find such a description to fix the referent of 'Muskie' misses the mark.

But the rigid designator theory (RD theory) seems to have a big advantage here. Kripke's terms refer to actual objects, so that the relationship between term and object is identical to what it is in extensional logic. Furthermore his causal theory of reference serves as an initially plausible (but not problem free) sketch about how reference fixing might go in the real world. In the individual concept theory, on the other hand, we are assigning terms to possible objects in different possible worlds, and so it is hard

is true, but it does not cause any problem of identifying objects across possible worlds. One might think there is a problem because if an object appears in two different worlds (by being a member (say) of $D_w$ and $D_v$), then we are required to give some sort of account of the relationship between its manefestations in these worlds that causes it to 'match up with itself'. But there is no problem at all. In the formal theory we simply have an object o that belongs to two sets: $D_w$ and $D_v$. That *doesn't* show that there should be two 'manefestations' of o which must therefore be pasted back together somehow. I fall into the set of philosophers, and into the set of people who like ice cream, but that doesn't show that there are two *mes*. It is only when we look at membership in the two sets $D_w$ and $D_v$ as incompatible (prompted no doubt by our talk of o-in-w and o-in-v) that we cannot ask for an explanation of how the *same* o can be in both of them. But there is no reason, apart from our spatialization of the possible worlds metaphor, that forces us to do that, and spatialization, I've been arguing, is an unfruitful feature of the metaphor of possible worlds.

to see how the causal (or any other) theory of reference could get off the ground.

The IC theorist, however, has a simple and effective reply. We need only claim that the 'normal' objects of everyday life correspond to the *individual concepts*, not the possible objects of the semantical theory. This is only natural and proper in the IC view because the domain of quantification (which is the prime candidate for determining ontological commitment) is a set of individual concepts.

The kind of reference which goes between terms and things described in the causal theory corresponds to the relationship between a term and its *intension*, not its extension. In one respect, the IC and RD theorists both agree in their account of the relationship between terms and the things of the world. On both theories, this kind of reference is determined without the help of possible worlds. The difference is that the formal reflection of a real thing on the IC view is complex, where it is simple on the RD view. The IC view, then, has extra resources for depicting the 'internal structure' of things.

So it is *complexes*, (individual concepts) that map over into the objects of our familiar world. The primitives (possible worlds and possible objects) are to be understood as abstractions from what is really there, in parallel with the treatment of mass points in physics.

The reader may still feel a worry. If reference goes from terms to individual concepts, how are we to give a coherent account of the sort of reference that goes between a term, a world, and a possible object? To help answer this question, and to prepare for the topic of the next section, let us consider a semantics for temporal logic where our intuitions are liable to be a bit clearer. If we adopt roughly the physicist's four-dimensional view of the world, temporal semantics might look something like this. Our set W now contains times. The members of O are the time-slices of objects (objects 'frozen' as they are at a given time). The individual concepts are now functions from times to time-slices, i.e. world lines. So the objects of daily life, the enduring and changing things, correspond to individual concepts, not to the members of O (which are time-slices). It is natural in this sort of semantics to adopt the IC account of terms, and to quantify over a set of individual concepts. Since reference in our world goes between terms and things, reference corresponds to the relationship between terms and their 'intensions', and not their 'extensions'.

Is there any special difficulty in this picture in making sense of the rela-

tionship between a term and its *extension* at a time? Not a bit. Given we have already got a relationship of reference between terms and things (intensions), and given that things include (in some sense) their time slices, then we can define a kind of time-slice reference *via* thing reference. We simply take the thing referred to by the term, and 'slice it' at the time in question, (i.e. take the object just as it is at that time).

Of course the process of fixing time slice reference is *abstract*. The process of 'taking a slice' has no physical realization, and the result of the process (the slice) isn't anything 'normal', since it has no temporal extension. Nor it is anything with which we could be properly said to have epistemological access. But such 'oddities' are no argument against our use of these abstractions. They would only be telling if we had been claiming that the process and resulting objects are real in the same sense that say aging and people are. This we are not doing. The time-slices are only a useful theoretical device for laying out the (real) structure of objects, as are mass points in physics.

A parallel should be clear between this discussion and an answer to the 'problem' of reference to possible objects. The reference of everyday life is mirrored in the formal machinery by the assignment of a term to an individual concept. Assignment of terms to possible objects in worlds is explicated *via* reference to individual concepts, and involves an abstraction from the modal nature of things. The possible objects are idealizations: things 'shorn of modality', i.e. stripped of their 'modal history', of their possibilities for being other than they are. Once we look at possible objects this way, worries about their ontological status seem no more well founded than similar worries about mass points. At the end of our search for the solution of the problem of transworld identification, we stumble across a solution to the 'problem' of possible objects as well. ([9])

([9]) This attitude towards modal semantics is completely *consistent* with many brands of modal realism, though not entirely with those brands that count possible objects real. There is no problem with adding to this view that essences of objects, haecities, etc. have ontological status, if we like. We have formal counterparts of these notions in the semantics, and it would seem an independent question as to exactly which one of these we would like to 'reify'.

However, the view we have presented also suggests an interesting brand of pragmatic conceptualism. Again the parallel with models in physics will be useful. In physics, the concept of a mass point is not only an abstraction, it is an idealization. In practice, one determines the limits of accuracy demanded by ones purposes, and provided by ones equipment, and than treats smaller chunks than that (roughly) as mass points. The detail that one uses in 'chopping up' the world then depends on ones goals and tools. Similarly one can abstract

## 7. *Are Rigid Terms Too Rigid?*

So far, we have done little more than defend the IC approach from potential criticism. Are there any considerations in its favor over the RD theory? I think we can find some, though they don't come in the form of a knock-down-drag-out argument. Our strategy will be to point to rigidities in Kripke's view which can be overcome with a more flexible treatment of terms.

Let us open that discussion by examing a likely reaction of an RD theorist to the time slice semantics of the previous section. 'Why', he says, 'do you introduce all that talk of slices in the first place?'. Terms just refer to objects, and objects, being atoms, do not need, or even have, a further analysis. Change is not to be mirrored in the semantics by setting up a worm composed of *different slices*, but rather by the way *objects* fall into the sets which are the extensions at different times of the predicate letters. We do not need to account for Howard Cosell's going bald by saying that there is a pair of non-identical Cosell-slices, an earlier one that isn't, and a later one that is bald. Why say such silly things? All we need to say is that the real honest to goodness *Howard Cosell* falls into the class of hairy things at an earlier time and into the class of hairless things at a later time. There is absolutely no need for your baloney metaphor and the formal bagage which accompanies it. Rigid designation will do.

This reply makes an attractive appeal to common sense. The problem with it is that without clear sense of what *counts* as common sense, the same tactic can be used to show that virtually any noun phrase is a rigid designator.

Consider, for example, the following defense of the view that 'the president' is a rigid designator in the face of an opponent who claims that the referent of 'the president' changes from time to time depending on the person who holds the office. 'Why', our defender begins, 'all this talk about people?' Terms refer to objects, and they refer to the same object regardless of what time it is. Change of presidents is not to be mirrored

from things to functions from situations to possible objects in a number of different ways, each of which reflects a different set of objectives and techniques for dealing with the world. Apart from these contextual features, no one way to locate what is possible for objects is 'right'. For example, the question 'Is it possible that Jim Garson pole vault 16 feet?' receives different answers depending on exactly how we take the possibility involved. It is mere logical possibility, physical possibility, physiological possibility (with or without training), etc.?

in the semantics by setting up a 'worm' composed of this person at one time and this other person at another time. Why say such silly things? All we need to say is that the real honest to goodness *president* falls into the class of crooks in 1972, the class of bumblers in 1975 and the class of earnest bumblers in 1978. There is absolutely no reason for chopping up my perfectly good president into its component people-chunks.

Of course this defense is just crazy, because we just *do* chop our world up into people chunks. The description 'the president' simply does not pick out a single object through time, at least not in our conceptual scheme. There is just too much difference between the bumbler of 1975 and the earnest bumbler of 1978 for us to count these as if there were one thing with different properties; there was a radical change in *identity* of the president between 1975 and 1978, and we have every right to reflect that in our sematics. Still we must be ready to rebut our defender's reply that the fact that we *do* chop up the world that way doesn't show that we can't and in fact might profit from, chopping it up differently. From the point of view of the political system for example, it's the office that counts as an object not the people who 'wander' in and out of it.

This shows that if the RD-theorist's reply to the time slice semantics is not to open the door to indiscriminate rigidification, he must develop a theory about which sortals are to be accepted (person) and which sortals are to be rejected (time slice, president-office) in our account of what it is that a term picks out when it picks out something. To use Geach's way of saying it, the RD theorist needs to give a philosophically adequate account of nominal essences. Our request for this kind of theory puts the RD theorist on the defensive a bit, but it certainly does not weigh heavily against him. A theory of nominal essences may well be something achievable, and well worth having. Kripke doesn't have to solve all the interesting problems in philosophy at once. So our remarks so far have been inconclusive. Nevertheless, they have served the purpose of sketching out some of the main ideas we will use for a more damaging attack.

The reason we were upset with the defender of the rigidity of 'the president' was that he refused to admit the conceptual importance of the distinction between people. It's not that the properties of the president change, the president actually turns out to be a different thing at different times. So this rigid designator theorist is open to the charge of refusing to make important distinctions which are worth making, and that suggests that it may turn that an RD theorist, however he fixes the nominal

essences, is likely to have trouble dealing with substantial change.

To illustrate the difficulties, using the ordinary sortals of daily life, consider (3).

(3) This wine will be vinegar.

If we assume that 'this wine' refers rigidly and picks out an object with respect to the sortal 'wine', then the standard semantical treatment rules that (3) is true just in case there is a time in the future when wine *is* vinegar, and this is a contradiction. If we try another tack and claim that (3) is elliptical for something like (4), then

(4) There is a future time such that just before it there was this wine, and after it, this vinegar

we are left with no formal feature with which to connect the wine with the vinegar. (It is enough for (4) to be true that there be this wine, and later, any sample of vinegar, whether it resulted from this wine's turning or not.)

We have merely posed Aristotle's problem of substantial change. Perhaps the RD theorist can profit from the 'Medieval answer. ([10]) We claim that 'this wine' does not pick out with respect to the sortal 'wine', but rather with respect to some more general sortal like 'stuff'. Then we can claim that it is this *stuff* (not wine) that changed into vinegar. This line is certainly consistent, but it sends us on the quest of some featureless 'thisity' to serve as the object of reference. The result is that we lose sight of what was appealing about the RD theory, its straightforward account of what it is that terms refer to. Clearly *primum materium* is an abstraction from the familiar objects of daily life, and reference to abstractions should make us queasy.

The IC theorists, though he is in the business of abstraction, does not require that reference is determined *via* such nonstandard sortals as 'stuff'. On his account 'this wine' refers via the sortal 'wine' to this wine, and this wine can be treated as a worm with winey slices at some times and vinegary slices at others. Since there is no need to count these slices identical on the IC theory, there is no need to postulate the existence of a *primum materium* to underlie change. The unity of the wine-through-change is not sought in any constant or internal feature that underlies

([10]) As I see it, the Medieval answer, *via* primium materium, is *not* Aristotle's answer.

change, but in the fact that these slices are aspects of (were abstracted from) one thing.([11])

Furthermore, the IC theorist does not need to claim that 'this wine' refers with respect to sortals which we clearly do not use when we identify things in the world. The fact that the slice for 'this wine' is vinegar at a later time is no embarassment, since there is no reason why the term can't pick out slices that are different at different times.

Some RD theorists might object that 'this wine' isn't a name anyway, and so the whole discussion doesn't count against the theory. Well then, name the wine 'Charley' and consider 'Charley will be vinegar'; or does the RD theorist want to claim that we can't name things that undergo substantial change?

We have been talking too long about time. The real issue in this paper concerns modality. The preceding discussion can, at best, only serve as a training ground to illustrate some of the moves we can make concerning possibility. Here the difficulties for RD theorists involve their inability to explain the *possibility* (rather than the actuality) of substantial change.

I think that it is clear that (5) is true.

(5) Nixon could have been a chimpanzee

I believe this because I can imagine a situation where the following things happen to *Nixon*. *He* grows up pretty much as he actually did, but between the ages of 20 and 40 *he* underwent very gradual changes. By age 40, doctors certify that *he* is physiologically identical to a chimpanzee. I suppose that this is not a biological possibility, but it is 'better' than a merely logical one.

Suppose first that the RD theorist agrees that (5) is true. Then he will need to postulate some one 'thisity' which is in one situation a person and in another a person-chimpanzee. To handle a full range of sentences like (5) he will at the very least have to individuate the world with respect to some sort of animal stuff, and have his terms refer to that. But then

([11]) Of course whether we adopt the RD or IC theory, we will need some practical account of 'what makes things cohere through substantial change'. I imagine that spatio-temporal continuity is the right place to look despite all the arcane counterexamples. Whatever the specific answer, it won't weigh for or against the RD and IC theories, since neither makes specific claims about coherence.

reference goes not to familiar objects but to 'thisities' shorn of their essential properties.

Suppose the RD theorist tries to deny that (5) is true. I think my story shows it is, but the RD theorist may counter that something which was a chimpanzee at a point in its life couldn't be the referent of 'Nixon', because being a chimp is not compatible with what Nixon is: a person. He claims, then, that failing to be a chimp is part of Nixon's essence, and so claims Nixon couldn't have that happen to him. ('If, anything', he says, 'your story was about somebody very like Nixon in certain respects, but that thing doesn't deserve the name "Nixon"'.)

If he is to handle a full variety of cases like (5), he will have to assume that what is picked out by a term is individuated by sortals in such a way that it is impossible for that thing to undergo any substantial changes (change of essential properties). The result is that he will be hard pressed to account for the *possibility* of substantial change, and I take it as axiomatic that we need such an account.([12])

The RD theorist can surely bite either horn of the dilemma here. As I said, I have no crushing argument. However, the IC theorist doesn't have to bite anything at all. From his point of view (5) can be true, and assuming it is does not require the location of anything novel as the object of reference. 'Nixon' refers to the real Nixon; on reflection about what is possible for *him*, we discover that he can be other than *what he is* as of this world. So 'Nixon' picks out 'modal slices' that belong to different natural kinds, and are therefore not identical. The possibility of Nixon's substantial change is accounted for because we do not have to say that any one thing is a person in one situation, and a person-chimp in another.

To illustrate that the RD theory faces the same problem for natural kind terms as well as names consider (6).

(6) Whales could have been quadrupeds

We should count (6) true because we know that had the course of evolution gone differently, the evolutionary pressures which lead certain hippopotamus like creatures to become sea dwellers might have been different, freeing these animals from the necessity of evolving fins.

([12]) One attractive point made by Kripke in section I of [9] is that when we talk about what is possible for Nixon, we are not talking about what happens to somebody just like him in another world; we are, instead, talking about *him*. So why should it be any different when Nixon happens to have the possibility of substantial change?

Should we claim that (6) is false on the grounds that being always a quadruped is inconsistent with what it is to be a whale? Should we conclude that 'whale' picks out some sort of undifferentiated animal stuff which doesn't have fins as part of its essence? ([13]) The IC theory does not require that we do either. (6) is true because there are situations we can imagine where whales turn out to be other than they *are*, i.e. where the natural kind term 'whale' rightly picks out things ([14]) different from what it picks out in the real world. The IC theory of reference allows us the flexibility to see to it that the essential properties of objects do not cross classify the sortals of individuation which help us fix reference, i.e. it lets us match up essences with nominal essences if we like. This is exactly what we may want to do.

## 8. *Rigidity in Metaphysics and Epistemology*

We have discussed the ackwardness of the RD theory's treatment of the possibility substantial change. The theory has another, related problem: true term identities are necessary. Because of the claimed rigidity of 'Hesperus' and 'gold', for example, we end up saying that Hesperus could not have been other than Phosphorus, and that gold could not be

([13]) Some people might object that (6) is really not about whales, but rather about evolutionary branches, and so (6) expresses that the gene pool for whales had the potentiality of producing quadrupeds. Hence the *line* terminating in whales did not have that particular termination as part of its essence.

But then how do *whales*, which represent, I suppose, the end of one line, manage to have the essential properties that they *do* have, i.e. being water dwellers, finned etc., given that the line doesn't?

It is much more straightforward to say that 'whale' is used to pick out the whole line, roughly as our objector contends, but that *that* underwent substantial change, at least with respect to the ordinary taxinomic categories.

In any case, *my* treatment of (6) turns out to be much more in the spirit of Kripke's remarks on essences [9] note 57, p. 351 (point (1)):

"Ordinarily when we ask intuitively whether something might have happened to a given object, we ask whether the universe could have gone on as it actually did up to a certain time, but diverge in its history from that point forward so that the vicissitudes of that objects would have been different from that time forth. Perhaps this feature should be erected into a general principle about essence. Note that the time in which the divergence from actual history occurs may be sometime before the object itself is actually created".

([14]) 'Things' here picks out the modal slice, not the object of reference of 'whale'.

other than a certain element, (given that we are right about what they are). Kripke realizes that this is one part of this theory that doesn't conform to what we are likely to say, so he gives the following explanation. ([15])

If we feel that Hesperus just might not have been Phosphorus, it is because we can imagine being in a 'qualitatively identical situation' to our own where we *call* two non-identical bodies 'Hesperus' and 'Phosphorus'. The fact that we might use these labels in other ways does not show that Hesperus could have been other than Phosphorus, for that qualitatively identical epistemic situation would not have been one where the labels refered to Hesperus and Phosphorus, since Hesperus and Phosphorus are the same thing.

Similarly, we feel that gold might have been a certain compound because we imagine ourselves in 'qualitatively identical evidential situations' where what we pick out with the label 'gold' is a compound. But what we *do* pick out would not be gold, since gold just *is* an element.

Kripke's examples are compelling because we assume that astronomical and chemical sortals are the right one to use for individuating objects and natural kinds, i.e. we assume 'Hesperus' refers to the *planet* Venus, and that 'gold' refers to an *element*. But, of course, our present theories might be hopelessly wrong. Suppose, then, that the correct theory identifies samples of gold and platinum, despite their seeming differences. (Then we would be roughly in the same spot as the 'silly' Babylonians who didn't know (according to Kripke) that 'Hesperus' and 'Phosphorus' refer to the same thing.) If Kripke's theory is right, we will have to conclude that 'gold' and 'platinum' refer to the same thing even though we have no clue (and perhaps no way of ever finding out!) what the concepts are that could be used to individuate things the right way. Suppose on the other hand that the right theory classifies samples of gold down the middle so that roughly half the gold has a positive and the rest a negative sense. It doesn't seem fair to claim that the real referent all along of 'gold' was either plus or minus gold, or that 'gold' just never referred at all (though we thought it did). Kripke's account of reference, then, is best described by the doggerel: 'I shot an arrow in the air, and if it lands I know not where'. This, I take it, amounts to the admission that God only knows for sure what we are talking about when we talk. We encouter

([15]) See KRIPKE's [9] pp. 208ff and 333ff.

a problem similar to the one we sketched in the previous section: objects of reference in the RD theory are opaque, in the sense that we don't know exactly what it is we are refering to. But how is it possible that the success and direction of reference could be so different from what we expect and intend?

Once we admit that we may be no better off than the Babylonians, we are likely to take a more practical line about how reference works, a line that does not sit well with Kripke's story. Clearly the Babylonians did not use 'Hesperus' to refer via the sortal 'planet', any more than the Greeks refered to gold *via* the sortal 'chemical element'. They didn't because they couldn't, seeing as how they never heard of these sortals. In fixing the referent of 'Hesperus' they clearly used some sortal like 'heavenly light'. Given this, they were *right* to consider Hesperus and Phosphorus different things (heavenly lights), just as *we* are right (no matter what the ultimate theory) to say gold and platinum are not the same thing (elements).

It overworks the principle of charity (and ignores the fact of conceptual change) to claim that the Babylonians were talking, exactly as we do, about planets. It overworks it even more to claim that both Babylonians were, and we are, talking about things in exactly the right way, as if there were one ultimately right way to individuate the world. The fact is, we know what we are talking about,([16]) and this knowledge reflects our competence in our language, our mastery of our conceptual scheme, as well as our knowledge about how the world is.

If people like Fodor [2] are right, even the ultimate scientific theory will not yield any one right way to individuate the world, since each of the special sciences *should* do so in different ways. Given this, there wouldn't have to be any one right way to say what sortal to use for 'Nixon'. Depending our purposes and the level of prediction and explanation we are interested in, we can individuate him as a person, an organism, an aggregate of molecules or fundamental particles, or even (to go the other way) as a role. (If Hitler had been born in Whittier California, he could have been Nixon.)

Now I might be entirely wrong about this. I don't pretend to have proven a case for my own story about reference. Maybe when we all get to heaven we will find out what the right sortals are. But then again, Kripke might be wrong.

([16]) That doesn't mean that we know everything *about* what we are talking about.

In the meantime, it seems best to adopt Pascal's strategy and hedge our bets. If we adopt the IC theory, we won't have to say that true term identities are necessary, and so we won't need an explanation of that which opens a potential rift between what we think we know about reference and what we are actually doing. Furthermore, even if the RD theory is right, it is easily accomodated in the IC approach by letting our domain of quantification be a set of constant functions. It would seem best then to pick the more inclusive option. This way we will avoid the risk that we will need to seek refuge in ephemeral features of the 'world behind our senses' to rationalize the claim that term identities, if true, are necessary.

*University of Houston* James W. GARSON
Houston TX
TEXAS MM 004
USA

BIBLIOGRAPHY

[1] BRESSAN, A. *A General Interpreted Modal Calculus*, Yale University Press, New haven, 1972.

[2] FODOR, J. *The Language of Thought, Introduction*, Thos. Crowley Company, New York, 1975.

[3] GARSON, J. "Unaxiomitizibility of an Intensional Logic", *Journal of Philosophical Logic*, vol. 9 (1980) pp. 59-72.

[4] — "Prepositional Logic" *Logique et Analyse*, vol. 24 (1981) pp. 4-33.

[5] — "The Substitution Interpretation and the Expressive Power of Intensional Logics" to appear in the *Notre Dame Journal of Formal Logic*.

[6] GEACH, P. *Mental Acts*, section 16, Routledge Kegan Paul, London, 1957.

[7] HINTIKKA, J. "Existential Presuppositions and Uniqueness Presuppositions" in *Philosophical Problems in Logic*, K. Lambert (ed.), D. Reidel, Dordrecht, 1970, pp. 20-55.

[8] KAMP, H. "Two Related Theorems of Scott and Kripke" (xerox).

[9] KRIPKE, S. "Naming and Necessity" in *Semantics of Natural Languages*, S. Davidson and G. Harman (eds.), D. Reidel, Dordrecht, 1972.

[10] — "Semantical Considerations in Modal Logic", *Acta Filosophica Fennica*, vol. 16 (1963), pp. 83-96.

[11] MONDADORI, F., and A. MORTON, "Modal Realism: The Poisoned Pawn", *Philosophical Review*, vol. 85 (1976), pp. 3-21.

[12] PLANTINGA, A. "Transworld Identity and World Bound Individuals" in *Logic and Ontology*, M. Munitz (ed.), New York University Press, New York, 1973, pp. 193-212.

[13] THOMASON, R. "Modal Logic and Metaphysics" in *The Logical Way of Doing Things*, K. Lambert (ed.), Yale University Press, New Haven, 1969, pp. 119-146.