MIDWEST STUDIES IN PHILOSOPHY, XI (1986)

Modal Paradox: Parts and Counterparts, Points and Counterpoints

NATHAN SALMON

I

There is a class of paradoxes that arise from the following (intuitively correct) modal principles concerning the possibility of variation in the original construction of an artifact:[1]

> If a wooden table x is the only table originally formed from a hunk (portion, quantity, bit) of matter y according to a certain plan (form, structure, design, configuration) P, then x is such that it might have been the only table formed according to the same plan P from a distinct but overlapping hunk of matter y' having exactly the same mass, volume, and chemical composition as y.
>
> If a wooden table x originally formed from a hunk of matter y is such that it might have been originally formed from a hunk of matter y' according to a certain plan P, then for any hunk of matter y'' having exactly the same matter in common with y that y' has, and having exactly the same mass, volume, and chemical composition as y', x is also such that it might have been originally formed from y'' according to the same plan P.
>
> (0) If a wooden table x is the only table originally formed from a hunk of matter y, then x is such that it could not have been the only table originally formed from entirely different matter, i.e., from a hunk of matter z having no matter in common with y (not even a single molecule, atom, or subatomic particle).

The last of these three modal principles, principle (0), is a nontrivial essentialist principle. It has been argued for by means of the following plausible, and perhaps more fundamental, essentialist principle concerning artifacts and their matter:

> (I) If a wooden table x is such that it might have been the only table originally formed from a hunk of matter z according to a certain plan P, then there could not be a table that is distinct from x and the only table formed from hunk z according to plan P.

The argument proceeds as follows: Let x be any arbitrary wooden table that is the only table formed from its original matter y, and let z be any nonoverlapping hunk of matter. Suppose for a *reductio ad absurdum* that table x is such that it might have been the only table originally formed from hunk z instead of from hunk y. Now necessarily, every table is formed according to some plan or other. Hence there is some plan P such that table x might have been the only table formed from hunk z according to plan P. It follows directly that hunk z is such that it might have been formed into a table (some table or other) according to the very plan P, and hence z might have been so formed only once. Since table x was actually the only table originally formed from hunk y, and since hunk z might have been formed into a table only once according to plan P, it might also have been that both obtained together. That is, it might have been that table x is the only table originally formed from hunk y, just as it actually was, while at the same time some table or other x' is the only table originally formed from hunk z according to plan P. (This is derived from a premise of the argument concerning the compossibility of certain possible states of affairs.) Of course, it is impossible for any one table to be originally formed entirely from one hunk of matter, and also originally formed entirely from some other, nonoverlapping hunk of matter. Thus, it is necessary that if table x is originally formed from hunk y, then any table formed from hunk z is not x. Hence, it is possible for there to be a table x' that is distinct from x and the only table originally formed from hunk z according to plan P. It follows by (I) and *modus tollens* that our original assumption that table x might have been the only table originally formed from hunk z is false.[2]

The first two principles cited above, taken together, imply that a certain *amount* of variation is possible in the original constitution of a table, whereas principle (0) implies that the amount of allowable variation is something short of total. A wooden table might have been originally formed from different wood, but not completely different wood; it might have been originally constructed with some different molecules, but not all. It follows that there is some threshold, some limit point—or if not a definite point, then at least some interval within which it is indeterminate—such that one more change in original constitution must by necessity result in a numerically distinct table.

It seems reasonable to suppose that the threshold consists in an interval of indeterminacy rather than a definite limit point. If a hunk of matter y' differs by only one molecule of wood from the original matter y of a table x, then clearly x is such that it might have been originally formed from y' instead of y. We have just seen an argument that if a hunk of matter z shares not even a single molecule of wood with the original matter of the table x, then x is such that it could not have been originally formed from z. Somewhere between these two extremes is the threshold—the minimum amount of required overlap, the maximum amount of allowable nonoverlap. The idea that this threshold amount should consist in an exact and specific number of shared molecules, or some other sort of sharp cutoff point, seems unrealistic. As with most of our concepts, our concepts of metaphysical possibility and impossibility do not seem to be quite that sharp. It seems more realistic to suppose that the threshold consists in some interval, perhaps some range of numbers of shared molecules. For any hunk of matter y' that shares a greater number of molecules with the actual matter y of the table

x than any number in this range, and that is otherwise just like y, it is determinately true of x that it might have originated from y' instead of from y. For any hunk of matter z sharing fewer molecules with y than any number in the range, it is determinately true of x that it could not have originated from z. For any hunk of matter y'' whose number of shared molecules with y lies within the range, it is indeterminate—vague, neither true nor false, there is no objective fact of the matter—whether x could have originated from y'' instead of from y.

Moreover, even if there is a sharp cutoff point,[3] it seems quite unrealistic to suppose that one could ever establish—say by a philosophical proof—precisely where the cutoff point lies. Thus even if the threshold is some exact and very precise amount of overlap, from an epistemic point of view we can never be in a position to specify with adequate justification just what the threshold is—except by means of some vague locution like 'sufficiently substantial overlap'. We may assert the following:

> (II) If a wooden table x is the only table originally formed from a hunk of matter y according to a certain plan P, and y' is any (possibly scattered) hunk of matter that sufficiently substantially overlaps y and has exactly the same mass, volume, and chemical composition as y, then x is such that it might have been the only table originally formed according to the same plan P from y' instead of from y.
>
> (III) If a wooden table x is the only table originally formed from a hunk of matter y, and z is any hunk of matter that does not sufficiently substantially overlap y, then x is such that it could not have been the only table originally formed from z instead of from y.

It is to be understood that being exactly the same matter except for only one or two molecules counts as sufficiently substantial overlap, whereas complete nonoverlap (no shared molecules whatsoever) does not.

Paradox arises when it is noted that none of these modal principles is the sort of proposition that merely happens to be true as a matter of contingent fact. In particular, principle (II) is such that if it is true at all, it is necessarily so. Furthermore, (II) is such that if it is true at all, then it is necessary that it is necessarily true, and it is necessary that it is necessary that it is necessarily true, and so on. In fact, on the conventionally accepted system *S5* of modal propositional logic, *any* proposition is such that if it is necessarily true, then it is necessary that it is necessarily true, and it is necessary that it is necessary that it is necessarily true, and so on.

One paradox that arises from these observations I call the 'Four Worlds Paradox'. Elsewhere I have developed the paradox using the language and framework of possible-world discourse, i.e., language involving explicit reference to, and quantification over, possible worlds and possible individuals (instead of the ordinary modal locutions 'might have', 'must', or subjunctive mood). The paradox is constructed by considering four distinct but related possible worlds. The Four Worlds Paradox can also be developed within modal-operator discourse, i.e., the language of the modal operators 'necessarily' or 'must', 'possibly' or 'might', and subjunctive mood. The paradox goes as follows: We consider a particular wooden table, a, with its four original legs, L_1, L_2, L_3, and L_4. Let us call the (hunk of) matter from which the table a was originally formed 'h'. The original matter of the four legs is a proper part of hunk h. Suppose for the sake of simplicity (though this is by no means essential to the argument)[4] that the threshold for table a is such that (for example) any table having the same overall plan (form, structure, design, configuration) as a is such that it might have been originally constructed using one leg different from its four actual original legs, as long as whatever other parts there are to the table (the other three original legs, the original table top, original wood screws, original glue, and so on) and the overall plan are the same. Suppose further that no table of this overall plan could have been originally constructed using two or more different legs from the actual original four. Now instead of constructing table a as he did, the artisan who constructed a might have constructed a table according to the same plan using two different table legs L_5 and L_6 in place of L_3 and L_4, keeping everything else the same—where L_5 and L_6 are qualitatively and structurally exactly like L_3 and L_4 actually are, respectively. Let us call this (scattered) hunk of matter 'h''. Hunk h' consists of hunk h with the replacement of the matter in legs L_3 and L_4 (at the time of table a's construction) with the qualitatively identical matter in legs L_5 and L_6. By principle (III), any such (possible) table must be distinct from a itself, but there is no reason why the artisan could not have thus constructed a qualitative duplicate of a instead of a itself. In accordance with *S5* modal propositional logic, it follows by the necessitation of principle (II) that the artisan might just as well have constructed a table distinct from a according to the same plan using L_1, L_2, L_3, and L_6 as the four legs instead of L_1, L_2, L_5, and L_6 (keeping everything else the same), since this would involve a change of only one table leg. Let us call this hunk of matter 'h'''. Hunk h'' coincides exactly with hunk h' except for the replacement of the matter in leg L_5 with the matter in leg L_3. Now hunk h'' also coincides exactly with hunk h (table a's actual original matter) except for the replacement of the matter in leg L_4 with the matter in leg L_6. Since the original table a was actually formed according to the same plan from hunk h, it also follows by principle (II) that the artisan might have constructed a itself according to the same plan using the same parts—L_1, L_2, L_3, and L_6, keeping everything else the same. Thus, the artisan might have constructed a by shaping certain matter h'' according to a certain plan, and he also might have constructed a table distinct from a by shaping exactly the same matter h'' according to exactly the same plan. This contradicts (I).

Formally, the Four Worlds Paradox proceeds from the following set of premises, where $\ulcorner M(\alpha, \beta)\urcorner$ means $\ulcorner\alpha$ is the only table originally formed from hunk of matter β according to such-and-such a plan$\urcorner$:

$M(a, h)$	[Given]
$\Diamond(\exists x)M(x, h')$	[Given]
$M(a, h) \supset \sim\Diamond M(a, h')$	[from (III)]
$\Box(x)[M(x, h') \supset \Diamond M(x, h'')]$	[from $\Box$(II)]
$M(a, h) \supset \Diamond M(a, h'')$	[from (II)]
$\Diamond M(a, h'') \supset \Box(x)[M(x, h'') \supset x = a]$	[from (I)].

From these (together with the trivial assumption that necessarily, if a table is formed from some matter, then it exists, and the quantified modal logical law of the necessity of identity) the following contradiction is immediately derivable in *S5*, and even the weaker *S4*, modal logic:

$$(C_1)\ \Diamond(\exists x)[x \neq a \,\&\, M(x, h'')] \,\&\, {\sim}\Diamond(\exists x)[x \neq a \,\&\, M(x, h'')].$$

I was once tempted by the view that this paradox is a *reductio ad absurdum* of the last premise cited above, and hence also a *reductio* of the cross-world identity principle (I). But to draw this conclusion is to miss the lesson of the paradox. Even if the last premise cited above is dropped from the list, an equally paradoxical argument can be constructed by invoking a slightly strengthened version of principle (II). To see this, let us first define the notion of a *materially complete* proposition. A proposition is materially complete if it is a complete enumeration of every particle of matter in the cosmos throughout all of a potential history of the world, as well as a complete specification of all the physical interactions and configurations of all the matter in the cosmos in exact chronological sequence throughout that potential history.

Needless to say, no materially complete proposition can be apprehended by the human mind, but of course, that is no reason to suppose that there are no such propositions. There are such propositions, and indeed one of them is true. Presumably, all true materially complete propositions are necessarily equivalent. On the modal logician's conception of propositions as sets of possible worlds (or as functions from possible worlds to truth values), exactly one materially complete "proposition" is true.

Let p be a (the) materially complete proposition that would have been true if the table a had been formed according to the same plan using leg L_6 instead of leg L_4. Notice that the materially complete proposition p surely strictly implies that some table or other is the only table originally formed from hunk h'' according to such-and-such a plan, in the sense that:

$$\Box[p \supset (\exists x)M(x, h'')].$$

Since p is a materially complete proposition that would have been true if table a had been formed from hunk h'' according to a certain plan, it is trivial that it might have been the case both that p is true and that a is the table formed from h'' according to that plan. By an argument that proceeds exactly as before, except invoking a stronger but still intuitively correct version of (II), it also might have been the case both that p is true and that the table formed from h'' is some table distinct from a. Hence in *S4* we may derive:

$$(C_2)\ \Diamond[p \,\&\, M(a, h'')] \,\&\, \Diamond[p \,\&\, {\sim}M(a, h'')].$$

This means that the question of which (possible) table is formed from hunk h'' (i.e., the question of the *haecceity* of the table formed from h'') is a question whose answer is not decided by a complete accounting of *all* the material facts in the cosmos—including the fact that hunk h'' exists as a physical unit and is table-shaped in such-and-such a particular way. This result is quite unpalatable. A table is in some obvious sense "nothing over and above" its matter and form. Perhaps some facts are underdetermined by the totality of material facts, but surely the question of whether a given actual table a is constituted by a certain hunk of matter h'' must be so determined. The fact that hunk h'' constitutes table a, if it does, is *supervenient* on a complete possible history of all the matter in the cosmos. If for some reason God had preferred to have table a originally formed from hunk h'' instead of from hunk h, once He has fixed all of the material facts—all of the facts concerning all of the matter in the cosmos—any further facts concerning which table is formed from which matter will take care of themselves. Hence, at a minimum, the following is true:

$$\Box[p \supset M(a, h'')] \vee \Box[p \supset {\sim}M(a, h'')].$$

This contradicts (C_2).

II

It is my view that both of the modal principles (II) and (III), and their multiple necessitations, are intuitively and literally true. Paradoxical conclusions are drawn from these principles by invoking defective rules of modal logic, by drawing fallacious modal inferences. Specifically, the conventionally accepted axiom of *S4* modal propositional logic,

$$\Box p \supset \Box\Box p,$$

or equivalently, the presumption that modal accessibility between worlds is transitive, is illegitimate and must be rejected in its unrestricted form. The modal logical system *S4* is fallacious. Its rejection invalidates a modal inference pattern critical to the Four Worlds Paradox:

$$\begin{array}{l} \Box(\phi \supset \Diamond\psi) \\ \Diamond\phi \\ \hline \therefore \Diamond\psi. \end{array}$$

Instead we have only the weaker inference:

$$\begin{array}{l} \Box(\phi \supset \Diamond\psi) \\ \Diamond\phi \\ \hline \therefore \Diamond\Diamond\psi. \end{array}$$

In particular, the hypotheses of the paradox yield the conlcusion that it might have been that it might have been that a table distinct from a was originally formed from hunk h'', but they do not yield the stronger conclusion that it might have been that a table distinct from a was originally formed from h''. There is no contradiction with (I).

The primary motivation for rejecting the *S4* axiom, as applied to the origins of artifacts (as well as other sorts of objects), is best given by means of an alternative modal paradox using a sorites-type construction, the main idea of which has been exploited by Roderick Chisholm. We begin with the same actual table a. The original matter h of table a consists of a certain number of molecules. Call this number 'n'. Now there is a finite sequence of hunks of matter, $h, h_1, h_2, \ldots, h_n$, where each element of the sequence h_i differs from its immediate predecessor h_{i-1} only in the replacement of one molecule by a qualitatively identical but numerically distinct molecule, in such a way that the last element in the sequence, h_n, has no overlap whatsoever with h, the original matter of table a. Now by the necessitation of principle (II), each of the following necessitated conditionals is true, where $\ulcorner M(\alpha, \beta)\urcorner$ again means $\ulcorner \alpha$ is the only table originally formed from hunk of matter β according to such-and-such a plan$\urcorner$:

$$\Box[M(a, h) \supset \Diamond M(a, h_1)]$$
$$\Box[M(a, h_1) \supset \Diamond M(a, h_2)]$$
$$\vdots$$
$$\Box[M(a, h_{n-1}) \supset \Diamond M(a, h_n)].$$

If we head this list with the true sentence '$M(a, h)$', we obtain a finite set of true premises that in *S4* logically entail the conclusion '$\Diamond M(a, h_n)$'. Let us call this argument (premise set plus conclusion) '(CP)', for 'Chisholm's Paradox'. The argument (CP) is *S4*-valid, and each of its premises is true. Yet by principle (III), '$\Box{\sim}M(a, h_n)$' is also true. Adding this to the list of premises of (CP), we obtain a set of true premises from which a contradiction is derivable in *S4*.

One can see what is amiss with *S4* by considering its import within the framework of possible worlds, to wit, the idea that the relation of modal accessibility between worlds is transitive. Since table a originates from hunk h in the actual world, it follows by (II) that there is a world w_1 possible relative to the actual world, i.e., accessible to the actual world, in which a originates from h_1. Hence by the necessitation of (II), there is a world w_2 possible relative to w_1 in which a originates from h_2. Hence by the *double* necessitation of (II), there is a world w_3 possible relative to w_2 in which a originates from h_3, and so on. Finally, by the $(n-1)$-fold necessitation of (II), there is a world w_n possible relative to w_{n-1} in which a originates from h_n. Thus, there is a world (w_n) bearing the ancestral of the accessibility relation to the actual world and in which a originates from h_n. But by principle (III), there is no *genuinely* possible world, i.e., no world possible relative to the actual world, in which a originates from h_n. Somewhere in the sequence $h_1, h_2, \ldots, h_n$, a hunk of matter h_m $(1 < m < n)$ is the first hunk to exceed the amount of allowable variation from h. Hunk h_m passes the threshold, and so, then, do all of its successors in the sequence. Hence, world w_m is not accessible to the actual world. World w_m is an *impossible world*. That is, w_m is impossible from the standpoint of the actual world, although it is possible relative to its immediate predecessor w_{m-1}, which is itself possible relative to the actual world. World w_m is a possibly possible impossible world.

Similarly, there is a world w_{2m} in which table a originates from hunk h_{2m}. World w_{2m} is possible relative to a world w_{2m-1} in which table a originates from hunk h_{2m-1}, and w_{2m-1} is possible relative to w_m, but w_{2m} is not possible relative to w_m. World w_{2m} is an impossible world that is not even a possibly possible world. It is only a possibly possibly possible world. That is, w_{2m} is a possibly possibly possible impossibly possible world.

This means that the relation of modal accessibility between worlds is not transitive. The premises of the argument (CP) are all true, but its conclusion is false. The argument (CP) is logically invalid.

If there is any defect in this illustration of the intransitivity of modal accessibility, and the consequent illegitimacy of *S4*, it is the assumption that there is some hunk of matter h_m that is the *first* hunk in the sequence to pass the threshold. This is tantamount to the assumption that the threshold consists in some definite number of shared molecules. This assumption, however, is quite inessential to the illustration. Suppose instead that there is a range of hunks, h_k, h_{k+1}, ..., h_{m-1}, such that for any hunk in this range, it is indeterminate—vague, neither true nor false, there is no objective fact of the matter—whether table *a* could have originated from it. This results in two limit points where before we had only one, and one alone is sufficient for a failure of transitivity. In the sequence of worlds w_1, w_2, ..., w_n, each world is determinately accessible to its immediate predecessor. Furthermore, each of the worlds w_1, w_2, ..., w_{k-1} is determinately accessible to the actual world (since it is determinately true that table *a* could have originated from hunk h_{k-1} or any of its predecessors), whereas each of the worlds w_m, w_{m+1}, ..., w_n is determinately inaccessible to the actual world (since it is determinately false that table *a* could have originated from hunk h_m or any of its successors). Each of the remaining worlds w_k, w_{k+1}, ..., w_{m-1} is neither determinately accessible nor determinately inaccessible to the actual world (since it is neither true nor false that *a* could have originated from h_k, or from h_{m-1}, or from any intervening hunk). This would mean that the accessibility relation is only *partially defined,* in the sense that its characteristic function is not total but partial. There would be a failure of transitivity *via* a region of indeterminacy, but there would still be a failure of transitivity.

Thus the modal paradoxes turn on a fallacy special to *S4* modal logic. In deriving the paradoxes in *S4,* one commits the *fallacy of possibility deletion,* inferring $\ulcorner\Diamond\phi\urcorner$ from $\ulcorner\Diamond\Diamond\phi\urcorner$, or equivalently, the *fallacy of necessity iteration,* inferring $\ulcorner\Box\Box\phi\urcorner$ from $\ulcorner\Box\phi\urcorner$. In particular, though it is necessary that table *a* does not originate from hunk h_m ($= h'$), it is fallacious to infer that it is necessary that it is necessary that *a* does not thus originate. In the Four Worlds Paradox, though it might have been that it might have been that some table distinct from *a* is formed from hunk h'' ($= h_{m-1}$), it is fallacious to infer that it might have been that some table distinct from *a* is formed from h''.[5]

III

The primary (though not the only) rival to this approach to the modal paradoxes is derived from the modal theory of David Lewis, so-called counterpart theory. Versions of the counterpart-theoretic solution to the paradoxes have been suggested or advocated by a number of philosophers, including Hugh Chandler, Roderick Chisholm, Graeme Forbes, Anil Gupta, Saul Kripke, and Robert Stalnaker.[6] Forbes in particular has recently worked out many of the details of a counterpart-theoretic solution, defending it against criticisms I have made and raising objections to the intransitive-accessibility solution sketched above.[7]

Strictly speaking, one should speak of counterpart theory *with respect to* a certain kind of entity, e.g., artifacts. Counterpart theory with respect to a kind k makes use of a binary cross-world resemblance relation, counterparthood, between possible entities of kind k. The counterpart relation is fixed by considerations of sufficient cross-world similarity in certain relevant respects. Since distinct possible entities of kind k may bear sufficient resemblance to one another across possible worlds, an individual x of kind k will have counterparts at other worlds other than itself. Typically, it is a basic tenet of the theory that each possible individual of kind k exists in one and only one possible world, so that a pair of counterparts existing in distinct worlds are always themselves distinct.

There are certain theoretical constraints on the counterpart relation. For example, any possible individual of kind k is its own counterpart at any (the) world in which it exists. Another minimal constraint is that if a possible individual x of kind k has a counterpart at world w that exists in w, then all of x' s counterparts at w exist in w. In the typical case, a counterpart of x at w is something that exists in w and (as it is in w) sufficiently resembles x as it is in its own world. Alternative versions of the theory provide for a possible individual to have a special counterpart at a world even though the counterpart does not itself exist in that world, as does Forbes's, but this happens only when the individual has no existing counterparts at the world in question. Yet another minimal constraint typically imposed is this: if a possible individual y is a counterpart of a possible individual x at a world w, and y itself has counterparts at w that exist in w, then all of y' s existing counterparts at w are also counterparts of x at w, i.e., all of a possible individual's existing counterparts at a given world are counterparts at that world of anything that the individual is itself a counterpart of at that world. This constraint can be trivially satisfied by means of the stronger constraint, typically but not always imposed, that any possible individual y that exists in w is its own sole counterpart at w. One condition typically not imposed, however, is transitivity. Since counterparthood is a cross-world similarity relation, and similarity is not transitive, there will be possible individuals x, y, and z, such that y exists in some world w and is a counterpart of x at w, and z exists in some world w' and is a counterpart of y at w', but z does not sufficiently resemble x to be a counterpart of x at w'.

Counterpart theory (with respect to kind k) provides for a possible-world semantic theory that differs in important respects from standard Kripkean possible-world semantics for modal-operator discourse. Let us first briefly review the main ideas that differentiate standard Kripkean possible-world semantics from classical Tarskian semantics. In standard Kripkean possible-world semantics, the extensional semantic attributes—such as singular term reference, predicate application, and sentence truth value—are relativized to possible worlds. In the case of reference and truth value, this relativization to worlds is in addition to the usual Tarskian relativization to assignments of values to individual variables. (Suppressing any reference to a model) if α is an individual variable, *the referent of* α *with respect to a world w under an assignment s*, or $Ref_{w,s}(\alpha)$, is simply the possible individual assigned to α by s, i.e., $s(\alpha)$. If α is a simple individual constant, it is assigned a referent (or to use Kripke's phrase, its "reference is fixed") independently of any possible world or assignment of values to variables. Thus, simple individual constants and individual variables are *obstinately rigid designators,*[8] expressions that refer to the same thing with respect to every possible world. If π is an n-place predicate, and $\alpha_1, \alpha_2, \ldots, \alpha_n$ are singular terms, then the atomic formula $\ulcorner\pi(\alpha_1, \alpha_2, \ldots, \alpha_n)\urcorner$ is *true with respect to a world w under an assignment s,* or *true* $_{w,s}$, if and only if π *applies with respect to w,* or *applies*$_w$, to the n-tuple consisting of the referents of each of the α_i with respect to w under s, $<Ref_{w,s}(\alpha_1), Ref_{w,s}(\alpha_2), \ldots, Ref_{w,s}(\alpha_n)>$. The connective and quantifier cases similarly follow standard Tarskian semantics. A formula $\ulcorner\Box\phi\urcorner$ is true $_{w,s}$ if and only if ϕ is true$_{w',s}$, for every world w' accessible to w. A formula $\ulcorner\Diamond\phi\urcorner$ is true$_{w,s}$ if and only if ϕ is true$_{w',s}$, for some world w' accessible to w. A sentence is *true (simpliciter)* if and only if it is true $_{\text{the actual world},\, s}$, for every assignment s.

Following the lead of Lewis, counterpart theorists typically formulate their theory in terms of translations of sentences (open or closed) involving modal operators into sentences of possible-world discourse, sentences involving explicit attribution of a counterpart relation between individuals in different worlds. This standard sort of formulation of counterpart theory may be regarded as providing a partial semantics for modal-operator discourse, in that it provides truth conditions in terms of possible worlds and counterparts for each sentence (open or closed) of modal-operator discourse. However, the semantics is only partial, since nothing is said explicitly concerning the semantics of subsentential expressions (such as singular terms and predicates) or how the truth conditions of sentences are computed from the semantics of their components. If one wishes to understand the compositional nature of the semantics of modal-operator-discourse sentences in terms of the semantics of their component expressions, one must glean this information, insofar as possible, from the translations into possible-world discourse of the modal-operator-discourse sentences in which the subsentential expressions figure. This feature of the standard formulations of counterpart theory is properly suited to a certain linguistic point of view concerning the synonymy of modal-operator-discourse sentences and the possible-world-discourse sentences giving the truth conditions of the former sentences, and the possibility of exhausting the semantics of modal-operator discourse merely by supplying possible-world-discourse sentential correlates. This point of view is disputable. Moreover, it is quite independent of the issues that separate standard possible-world theorists from counterpart theorists, and it is quite inessential to the main philosophical ideas and intuitions that motivate counterpart theory. If a standard modal theorist adopts this point of view, he or she may easily reformulate the standard modal semantics as a set of instructions for translation of modal-operator-discourse sentences into possible-world-discourse sentences, remaining silent with respect to the compositional nature of the semantics of sentences in terms of the semantics of subsentential expressions. In order to highlight the contrast with standard modal semantics, while clearing away the unimportant differences in what has come to be the usual sort of formulations of each, it is best to reformulate counterpart theory along lines that parallel as closely as possible, within the bounds of the spirit of the philosophical motivation for counterpart theory, the usual formulation of standard possible-world semantics.

I shall do this using the notion of a *counterpart assignment*. A counterpart assignment c_w (with respect to a kind k) for a world w is a function that assigns to any possible individual i (of kind k) a counterpart of i at w, if i has any counterparts at w, and assigns nothing otherwise. If there is no counterpart of i at w existing in w, then depending on the particular counterpart theory in question, the counterpart assignment may be undefined for i, as with Lewis's theory, or it may assign the individual i to itself as its own counterpart at w, as with Forbes's. On Forbes's theory, counterpart assignments are totally defined functions.

Let us call an ordered pair of a world and a counterpart assignment for that world a *world-assignment pair*. In counterpart theory with respect to kind k, reference and truth are relativized not merely to worlds but to world-assignment pairs.[9] Thus one speaks of the referent of a singular term *with respect to* a world-assignment pair $<w, c>$ under an assignment of values to variables s. Equivalently, one may speak of reference with respect to a world w *and* a counterpart assignment c for w, under an assignment of values to variables s. Similarly, one speaks of a sentence (open or closed) as being true, or as not being true, *with respect to* a world-assignment pair under an assignment of values to variables. As in standard possible-world semantics, predicate application is relativized only to worlds. The referents of simple singular terms with respect to world-assignment pairs will depend on whether the term has been assigned something of kind k. If α is an individual variable and s is an assignment of values to variables that assigns to α a possible individual not of kind k, then $Ref_{w,c,s}(\alpha) = s(\alpha)$. If α is an individual variable and s is an assignment that assigns to α a possible individual of kind k, then $Ref_{w,c,s}(\alpha) = c(s(\alpha))$. If α is a simple individual constant that refers to an actual individual x not of kind k, then $Ref_{w,c,s}(\alpha) = x$. If α is a simple individual constant that refers to an actual individual x of kind k, then $Ref_{w,c,s}(\alpha) = c(x)$. An atomic formula $\ulcorner \pi(\alpha_1, \alpha_2, \ldots, \alpha_n) \urcorner$ is $true_{w,c,s}$ if and only if π applies$_w$ to $<Ref_{w,c,s}(\alpha_1), Ref_{w,c,s}(\alpha_2), \ldots, Ref_{w,c,s}(\alpha_n)>$. A formula $\ulcorner \Box\phi \urcorner$ is $true_{w,c,s}$ if and only if ϕ is $true_{w',c',s}$, for every world w' and every counterpart assignment c' for w' (i.e., for every world-assignment pair $<w', c'>$). A formula $\ulcorner \Diamond\phi \urcorner$ is $true_{w,c,s}$ if and only if ϕ is $true_{w',c',s}$, for some world w' and some counterpart assignment c' for w' (i.e., for some world-assignment pair $<w', c'>$.)[10] Notice that the clause 'w' is accessible to w' has been deleted; counterpart theory avoids the need for an accessibility relational semantics. A sentence is *true (simpliciter)* if and only if it is $true_{\text{the actual world},c,s}$, for every counterpart assignment c for the actual world and every assignment of values to variables s.

The major difference between counterpart theory and standard possible-world semantics may be illustrated by means of a simple modal sentence from Chisholm's paradox,

$$\Diamond M(a, h_1).$$

On standard possible-world semantics, this sentence is true exactly on the condition that there is a possible world (determinately) accessible to the actual world in which table a—the very table a itself—is the only table formed according to such-and-such a plan from hunk h_1 (instead of from its actual original matter h). The counterpart theorist does not admit that this condition is fulfilled. Instead, typically the counterpart theorist denies that there is any such possible world. The counterpart theorist is still able to accommodate the truth of the displayed sentence. On counterpart theory with respect to artifacts, the sentence is true exactly on the condition that in some possible world, some counterpart of *a*—not necessarily the very table *a* itself—is the only table formed according to such-and-such a plan from hunk h_1. Counterpart theory with respect to artifacts thus assigns a different truth condition to the sentence, one whose fulfillment seems beyond doubt.

In effect, counterpart theory replaces the intransitive accessibility relation with an intransitive counterpart relation. There are glaring technical differences between the two types of solutions to the modal paradoxes, however. (There are glaring motivational differences as well. The motivation for counterpart theory, as a solution to the modal paradoxes, is discussed in section V below.) First, certain intuitively correct premises involved in Chisholm's Paradox are counted unequivocally true on the accessibility solution but cannot be thus accommodated on counterpart theory (as I have formulated it). Consider the argument (CP). Suppose again that in the sequence of hunks of matter, $h_1, h_2, \ldots, h_n$, some one hunk h_m is the first in the sequence to pass the threshold. Then on counterpart theory with respect to artifacts, the premise

$$(P_m) \quad \Box[M(a, h_{m-1}) \supset \Diamond M(a, h_m)]$$

will not be true, since there is a world w_{m-1} in which a counterpart of table *a* is formed from hunk h_{m-1}, whereas at any world in which a table is formed from hunk h_m, that table, though a counterpart of the counterpart of *a* at w_{m-1}, is not a counterpart of *a* itself. (A similar situation obtains if the threshold is vague and there is a range of hunks $h_k, h_{k+1}, \ldots, h_{m-1}$ for which it is indeterminate whether a possible table formed from one of these hunks is a counterpart of *a*.) Thus whereas the accessibility solution blocks (CP) by counting it logically invalid, counterpart theory with respect to artifacts (as I have formulated it) blocks (CP) by counting it logically valid but unsound.[11]

Another glaring difference between the two solutions to the modal paradoxes is brought out in their respective treatments of the Four Worlds Paradox. Although counterpart theory with respect to artifacts is able to accommodate *S5* modal propositional logic, in so doing it foregoes certain valid inferences of standard quantified *S5* modal logic. In particular, it is able to accommodate the truth of the necessitation of the modal principle (II), and of certain sorts of instances of it, like the fourth premise of the Four Worlds Paradox,

$$\Box(x)[M(x, h') \supset \Diamond M(x, h'')].$$

In standard quantified modal logic, it follows from this together with the result

$$\Diamond(\exists x)[x \neq a \,\&\, M(x, h')]$$

and the trivial truism

$$\Box(x)\Box[M(x, h'') \supset (\exists y)(y = x)]$$

that

$$\Diamond\Diamond(\exists x)[x \neq a \,\&\, M(x, h'')].$$

Counterpart theory with respect to artifacts invalidates this inference and thereby blocks the paradox. The accessibility solution, on the other hand, allows the inference, but invalidates further inference by possibility deletion. Thus both solutions to the Four Worlds Paradox count the argument of the paradox invalid, though on distinctly different grounds. Similarly, counterpart theory with respect to artifacts accommodates

$$\Box(x)[M(x, h_{m-1}) \supset \Diamond M(x, h_m)].$$

while blocking the inference from this together with '$\Box[M(a, h_{m-1}) \supset (\exists x)(x = a)]$' to the (CP) premise ($P_m$) displayed above

In the general case, if counterpart-theoretic possible-world semantics is devised in such a way as to preserve *S5* modal propositional logic together with the philosophical institutions that motivate the theory, it foregoes the following modal version of universal instantiation, valid in standard quantified modal logic:

$$\text{(MUI)} \qquad \frac{\Box(x)\phi_x}{\therefore \Box[(\exists x)(x = \alpha) \supset \phi_\alpha],}$$

where α is a simple individual constant or individual variable other than 'x', ϕ_α is just like ϕ_x except for having free occurrences of α wherever ϕ_x has free occurrences of 'x', and ϕ_x may contain occurrences of modal operators. This deviation from standard quantified modal logic prevents the derivation of paradoxical conclusions from the necessitation of (II).[12]

IV

Each of the necessitated conditional premises of the argument (CP) is equivalent in *S4* to an unnecessitated material conditional, so that the argument may be recast in *S4* into the standard form of a sorites argument in classical propositional logic:

$$\begin{array}{ll} \text{(CP)}' & \Diamond M(a, h) \\ & \Diamond M(a, h) \supset \Diamond M(a, h_1) \\ & \Diamond M(a, h_1) \supset \Diamond M(a, h_2) \\ & \quad \vdots \\ & \Diamond M(a, h_{n-1}) \supset \Diamond M(a, h_n) \\ \hline & \therefore\ \Diamond M(a, h_n). \end{array}$$

Forbes emphasizes this feature of Chisholm's Paradox and argues that the paradox should be treated in a manner exactly parallel, or as closely as possible, to a contemporary treatment of the standard propositional sorites paradox, such as the paradox of the short person:

Anyone only 5 ft. tall is short.
If anyone 5 ft. tall is short, then so is anyone 5 ft. $\frac{1}{1{,}000{,}000}$ in. tall.

⋮

If anyone 5 ft. 11 and $\frac{999{,}999}{1{,}000{,}000}$ in. tall is short, then so is anyone 6 ft. tall.

∴ Anyone 6 ft. tall is short.

Standard sorites paradoxes arise from vagueness in some key expression or concept. In the case of the paradox of the short person, the key term is the adjective 'short', which is clearly true of anyone (or at least, any adult human) only 5 feet tall, clearly false of anyone 6 feet tall, but neither clearly true nor clearly false with respect to a range of heights in between. Now one extremely plausible way of diagnosing the problem with this sorites argument is as follows. Assuming that the first premise of the argument is true and that the conclusion is false (its negation true), somewhere down the list of the 12 million conditional premises to the argument—in fact, at least twice, and most plausibly, a large number of times down the list—a conditional premise is neither true nor false. For somewhere down the list there is a conditional with a true antecedent but a consequent neither true nor false, followed by a sequence of conditionals with both antecedent and consequent neither true nor false, followed finally by a conditional with an antecedent neither true nor false and a false consequent. Each of these premises is itself neither true nor false. Thus the classical sorites argument in propositional logic is formally valid but unsound. Not all of its premises are true, even if none are strictly false.[13]

The solution to the modal paradoxes offered in Section II above allows for a treatment of (CP)′ exactly parallel to this. In particular, the critically vague term involved in (CP)′, if any (see note 3), is the accessibility predicate of possible-world discourse, and thereby the possibility operator '$\Diamond$' occurring throughout (CP)′. A sentence $\ulcorner \Diamond \phi \urcorner$ is true (*simpliciter*) if and only if ϕ is true with respect to some world determinately accessible to the actual world. The same sentence is false (*simpliciter*) if and only if ϕ is false with respect to every world determinately accessible to the actual world and untrue—either false or neither true nor false—with respect to every world neither determinately accessible nor determinately inaccessible to the actual world.[14] The intransitive accessibility account allows that there may be a hunk of matter h_k such that table a originates from it in some world neither determinately accessible nor determinately inaccessible to the actual world, but does not originate from it in any determinately accessible world. If this is so, $\ulcorner \Diamond M(a, h_k) \urcorner$ is neither true nor false. Hence at least two of the conditional premises of (CP)′ will be neither true nor false, just as in the paradox of the short person. Insofar as it is desirable for a solution to (CP)′ to parallel as closely as possible a contemporary solution to the classical propositional sorites paradox, the indeterminate accessibility solution does exactly what is desired.

More important than this, the accessibility solution severs the alleged equivalence between (CP) and (CP)′, and in fact, the original modal argument (CP) comes out differently in a very important respect from a standard sorites argument. Unlike the premise set of the propositional sorites argument (CP)′, all of the premises of the original argument (CP) are determinately true, whereas its conclusion is determinately false. This reflects a crucially important difference between Chisholm's Paradox and the standard sorites paradox. It is important to remember that Chisholm's Paradox, as well as the Four Worlds Paradox and others belonging to the same class, are paradoxes of *modality*. Chisholm's Paradox is not a paradox in classical propositional logic, but a paradox in modal logic. The key feature of Chisholm's Paradox—the feature of it that makes it a peculiarly modal paradox—is its essential use of *nested modalities*. It proceeds from the observation that the truth of the modal principle (II) is no accident but is a necessary truth, thus yielding the nesting of modal operators in the modal premises of (CP). The intransitive-accessibility solution to Chisholm's Paradox properly distinguishes between the original argument (CP) and the propositional recasting (CP)′, the latter being a familiarly valid but unsound argument in classical propositional logic and the former an interestingly invalid argument in modal logic. It is a critical defect in the counterpart-theoretic solution (as well as other rivals to the intransitive-accessibility solution) that it is blind to the crucial differences that separate the two cases. The modal paradoxes, as they naturally arise in pondering essentialist doctrines of the sort put forward in principle (III) (and as they did in fact arise in Chisholm's pioneering queries on the subject), are peculiarly modal in that they involve nested modality and depend upon the fallacy of possibility deletion, or equivalently, the presumption that accessibility between worlds is transitive. The counterpart-theoretic solution, in attempting to reduce the modal paradoxes to "the previous case" of standard sorites paradoxes such as the paradox of the short person, recommits the same fallacy and, in so doing, fails to recognize the rightful status, and consequently the proper lesson, of the modal paradoxes.

V

The fundamental defect of the counterpart-theoretic solution to the modal paradoxes is revealed when considering the motivation for invoking counterpart theory in attempting to solve the paradoxes. If Chisholm's Paradox is to be regarded on the model of the paradox of the short person, one must ask what term or expression involved in the former plays the role of the crucially vague term 'short' involved in the latter.

It cannot be expression '*a*' itself. In fact, it is not in the least bit clear what it would mean to say that a proper name—or an individual constant such as '*a*', which functions as a proper name—is "vague," unless it means that '*a*' is ambiguous or nonreferring. We may pretend, for present purposes, that the name '*a*' unambiguously refers to a particular table. The paradoxes still arise. It is even less clear what it would mean to say that the table *a* itself is vague, unless it means that the table has a vague *boundary,* in the sense that with respect to certain molecules at the periphery of the table, it is vague—indeterminate, neither true nor false, there is no objective fact of the matter—whether they are or are not constituents of the table. But vagueness in the table's boundary is not at issue here; the modal paradoxes would arise even if tables came with sharp boundaries.

Nor is there any relevant vagueness in the term 'table', or in the property of being a table. No doubt there are things such that it is vague whether they are to count as tables (as opposed to, say, counters or chests), but we may take it that *a* itself is a clear and central case of a table. The paradox still arises.

Nor is there any relevant vagueness in the hunks of matter $h, h_1, \ldots, h_n$. We may suppose that these are precisely given, with an exact accounting of every molecule included and the exact configuration of their totality. The paradox still arises. Nor is there any relevant vagueness in the relational concept of a table x being originally formed from a hunk of matter y according to such-and-such a plan. Wherein, then, does the vagueness reside?

One might try looking at the matter thus: In the short person paradox, there is a sequence of heights, 5 ft., 5 ft. $\frac{1}{1{,}000{,}000}$ in., …, 6 ft., such that, though each height is precise and exact enough in itself, for some of these precisely delineated heights it is vague whether someone of that exact height counts as being *short*. Similarly, in the case of Chisholm's Paradox, we have a sequence of hunks of matter, $h_1, h_2, \ldots, h_n$, each precisely given, and a corresponding sequence of worlds, $w_1, w_2, \ldots, w_n$, such that in any world w_i there is a table a_i just like a except that it is originally formed from hunk h_i instead of from a's original matter, h. This sequence of possible tables, a_1, $a_2, \ldots, a_n$, plays the role analogous to that of the sequence of heights in the short person paradox. Each is precisely given, though for some it is vague whether the table still counts as being a or not. In the actual world, there is also a table just like a originally formed from hunk h. This table is a itself. In world w_n, the table a_n formed from hunk h_n is definitely not a, since by principle (III) there is no genuinely possible world in which a is formed from h_n. With respect to certain worlds w_k intermediate in the sequence between the actual world and w_n, it is vague—indeterminate, neither true nor false, there is no objective fact of the matter—whether the table a_k formed from hunk h_k in that world is or is not the very table a from the actual world. To use the contemporary vernacular, what is indeterminate is whether a_k has a's haecceity—the property of being identical with a—in w_k. Thus the vague concept involved in Chisholm's Paradox would appear to be that of *being identical with a in a possible world*, or more simply, *possibly being a*. More specifically, since the name 'a' is itself nonvague, the relevant vague concept involved would appear to be the relational concept of cross-world identity, or that of possible identity, expressed by '$\Diamond x = y$'. Evidently, this vagueness traces to vagueness in the very concept of identity itself. The ultimate source of the vagueness involved in Chisholm's Paradox thus appears to be the 'is' of identity.

Kripke, apparently having reasoned along lines similar to these, concludes that a counterpart-theoretic approach may be useful in dealing with the vagueness of identity in Chisholm's Paradox. He says that

> perhaps, . . . given certain counterfactual vicissitudes in the history of the molecule of a table, *T*, one may ask whether *T* would exist, in that situation, or whether a certain bunch of molecules, which in that situation would constitute a table, constitute the very same table *T*. . . . In concrete cases we may be able to answer whether a certain bunch of molecules would still constitute *T*, though in some cases the answer may be indeterminate. (*Naming and Necessity*, pp. 50–51)

In a footnote to this passage, he writes:

> There is some vagueness here. If a chip, or molecule, of a given table had been replaced by another one, we would be content to say that we have the same table. But if too many chips were different, we would seem to have a different one. . . . Where the identity relation is vague, it may seem intransitive; a chain of apparent identities may yield an apparent nonidentity. Some sort of 'counterpart' notion . . . may have some utility here. One could say that strict identity applies only to the particulars (the molecules), and the counterpart relation to the particulars 'composed' of them, the tables. The counterpart relation can then be declared to be vague and intransitive. . . . Logicians have not developed a logic of vagueness. (p. 51, note 18)

There are a number of difficulties with this motivation for the counterpart-theoretic approach. Kripke's idea seems to be that where (the characteristic function of) the concept of identity is undefined, it may facilitate a semantic investigation if the identity concept is represented in the metalanguage by means of a surrogate relation, counterparthood, which is vague and intransitive. Now it may indeed facilitate a semantic investigation into the logic of a vague term or predicate such as 'bald' to consider various regimented or sharpened surrogates or approximations to the vague concept, precisely defined—say, in terms of an exact number of strands of hair on the top portion of the head per square inch of surface area. One might thus verify the validity of the inference ⌜α has a full head of hair $\therefore$ α is not bald⌝. But Kripke is proposing that an allegedly vague concept, identity, be investigated in terms of *another vague concept*, counterparthood. It is difficult to see how there is anything to be gained in representing one vague concept by means of another. If our problem is that we lack a logic of vagueness, we can no more treat the latter than we can the former. If our purpose is to investigate the logic of identity among tables, surely we are better off sticking with genuine identity and doing the best we can, than turning our attention elsewhere only to find the same obstacles arise there.

Perhaps Kripke committed a slip of the pen here and meant to declare the counterpart relation to be *non*vague and intransitive—as opposed to genuine identity among tables, which it represents and which (we are to suppose) is vague but transitive. For example, one might define a relation of counterparthood in such a way that any possible table is a counterpart of itself, i.e., of a determinate self, whereas for any pair of possible tables a_i and a_j for which it is either false or vague (neither true nor false) that they are identical, neither counts as a counterpart of the other. This counterpart relation would thus play the facilitating role of a sharpened or regimented approximation to identity among tables and other artifacts.

Even when Kripke's proposal is modified in this way, it seems confused. It is quite unclear what it means to say that strict identity does not "apply" to tables. Suppose there is a possible table a_k such that a and a_k are neither determinately identical nor determinately distinct. Then on this interpretation of Kripke's proposal, a is a counterpart of a (itself) but not of a_k. It follows directly by Leibniz's Law, or the Indiscernibility of Identicals, that a and a_k are distinct, contradicting the hypothesis. (A similar argument applies if counterparthood is defined so that a and a_k are counterparts.)

The defender of this proposal may protest that within the counterpart-theoretic framework, one is barred from saying anything about the identity or distinctness of a and a_k. One can speak only about the cross-world similarity relations between a and a_k; one must settle for the weak claim that a and a_k are not counterparts. But the Leibniz's Law inference cries out to be drawn; if a is a counterpart of a but not of a_k, then a has a counterpart that a_k does not have. Whether we are allowed to say so or not, it follows that a and a_k *cannot be* one and the very same object and *must be* distinct. Our refraining or being prohibited from saying so does not make it any less true.

When the truth is spoken, incoherence is the result. Consider again the sequence of possible tables $a, a_1, a_2, \ldots, a_n$. Kripke's remarks concerning this sort of situation are highly compressed, and his exact intent is unclear. He says: "Where the identity relation is vague, it may *seem* intransitive; a chain of *apparent* identities may yield an *apparent* nonidentity" (emphasis added). Presumably, if "the identity relation is vague," then things that are *apparently* identical (or *apparently* distinct) need not be *determinately* identical (or *determinately* distinct). A pair of objects x and y may *appear* to be identical (or distinct) when in reality, there is no objective fact of the matter as to their identity (or their distinctness). Perhaps Kripke's view, then, is this: (i) any table in the sequence $a, a_1, a_2, \ldots, a_n$ appears to be identical to its immediate successor in the sequence; (ii) the initial table a and the final table a_n appear to be distinct; but (iii) in reality, for any pair of tables a_i and a_j where $i \neq j$, there is no objective fact of the matter concerning their identity or distinctness.

In that case, Kripke's view would involve rejection of both the modal principles (II) (since a and a_1 only *appear* identical) and (III) (since a and a_n only *appear* distinct). This is not a very satisfactory solution to the modal paradoxes. Both (II) and (III) are intuitively correct, even if it is vague what is to count as "sufficiently substantial overlap." In fact, if Kripke's view is that it is vague—or indeterminate, or neither true nor false, or there is no objective fact of the matter—whether tables a and a_n are distinct, then his view involves rejection of the modal principle (0), a principle that is both weaker than (III) and precisely formulated in a way that (III) is not. It would be difficult, if not impossible, to reconcile this consequence of Kripke's view with his attempt in the very same work to provide "something like proof" for principle (0), or a principle directly like it. (See note 2 above.)

Another possible view might be that in the sequence of possible tables $a, a_1, a_2, \ldots, a_n$, each element is determinately identical with its immediate successor, though there is some range of elements, $a_k, a_{k+1}, \ldots, a_{m-1}$, that are each neither determinately identical with nor determinately distinct from the initial element a, whereas the next element in the sequence, a_m, and all of its successors are determinately distinct from the initial element a. However, this is equally incoherent. If a and a_{k-1} are determinately identical, then they are *one and the very same*, and if a_{k-1} and a_k are determinately identical, then they are also one and the very same. But then there is only one table here. Which table? Well a, aka a_k. Tables a and a_k are one and the very same after all; they are determinately identical. Conversely, if a_m is determinately distinct from a, yet determinately one and the very same table as a_{m-1}, then a_{m-1} must be determinately distinct from a after all. Moreover, if each element in the sequence and its immediate successor are one and the very same, then what we have is simply an n-ary sequence of table a taken n times in a row. It is quite literally impossible for some element in this sequence to be distinct from a. Conversely, if any table in the sequence fails to be determinately identical with the initial table a, then the sequence is not simply the n-ary sequence of a taken n times in a row. Hence it is impossible for each element in the sequence to be one and the very same as its immediate successor.

The idea that the identity relation is vague, in the sense that its characteristic function is undefined for certain pairs of concrete objects like tables, is itself incoherent. In fact, it is provable that the identity concept, or the 'is' of identity, is totally defined for every pair of individuals. The proof, which was foreshadowed in the arguments just given, goes as follows: Suppose, on the contrary, that there is a pair of individuals, x and y, for which the 'is' of identity is undefined—a pair to which neither the predicate 'are one and the very same' nor its negation 'are not one and the very same' correctly applies. Then this pair $<x, y>$ is quite definitely *not* the same pair as the reflexive pair $<x, x>$, since the 'is' of identity—or the predicate 'are one and the very same'—does correctly apply to the latter. That is, the pair $<x, x>$ is an element of the extension of the 'is' of identity (the class of ordered pairs of which the predicate is determinately true), whereas the pair $<x, y>$ is not; hence, they are distinct. It follows by standard ZF set theory that $x \neq y$. But then, contrary to the hypothesis, the 'is' of identity is defined for the pair $<x, y>$. The 'is' of identity is determinately false of the pair; its negation correctly applies. The general form of this argument can be applied to a variety of philosophical issues concerning identity.[15]

In fact, this brief argument also proves that the concepts of identity within a possible world, i.e., intra-world identity, and of cross-world identity (and by analogy, identity at a time and identity over time) are also totally defined. For each is definable in terms of absolute, unrelativized identity as follows:

$x =_w y =_{def.} x = y.$
x in w_1 is identical with y in w_2 $=_{def.}$ x exists in
w_1 & y exists in w_2 & $x =_{w_1} (\iota z)\,[y =_{w_2} z]$.

Perhaps most important, Chisholm's Paradox and the other modal paradoxes do not even involve the concept or relation of identity. The paradoxes *can* be formulated in terms of possible identity or cross-world identity, but they can just as easily be formulated without identity. In fact, the 'is' of identity does not occur in either (CP) or (CP)′—not once, not anywhere. If (CP) and (CP)′ constitute a paradox of vagueness, the vagueness must reside in one or more of the terms actually used in the formulation. Since the identity predicate does not even occur, if there is any vagueness, it must reside elsewhere. It is a mistake to see Chisholm's Paradox as stemming from vagueness in identity.

Forbes's motivation for his counterpart-theoretic approach to Chisholm's Paradox is somewhat different from Kripke's, though he seems to mislocate the vagueness in the same place. He writes:

> [There] is no sharp distinction between those sums [of matter] which could, and those which could not, constitute [the table *a*]. Given that there is no fuzziness in the boundaries of particular sums of wood or in the constitution relation, it seems that this vagueness must arise from an underlying vagueness in the concept of possibly being identical to [*a*]; however, in standard [possible-world] semantics, such vagueness could only be represented by vagueness in [*a*' s cross-world] identity conditions, and a solution of the paradox in which we think of identity as vague would be rather unappealing. But [it] does make sense to think of *similarity* as being vague, in the sense of admitting *degrees*. . . . [The] counterpart relation is fixed by similarity considerations—in the present context, similarity of design and constituting matter. (''Two Solutions to Chisholm's Paradox,'' p. 174)

Forbes's overall argument appears to be this: The orginal argument (CP) is equivalent in *S5* to (CP)′, a standard propositional sorites-type argument; hence it is simply a special case of a general and familiar sort of paradox of vagueness. Since the vagueness crucially involved in (CP)′ does not reside in the hunks of matter h_1, h_2 …, h_n or in the relation of being a table formed from such-and-such matter, it must reside in the concept of possibly being *a*. On the standard possible-world semantic analysis of modal-operator discourse, this would mean that there is vagueness in the identity relation itself. But the idea that identity is vague is ''rather unappealing'' as a solution to Chisholm's Paradox. Counterpart theory provides an alternative possible-world semantic analysis of modal-operator discourse in which the vagueness of possibly being *a* is derived not from vagueness in identity, but from vagueness in a relation of similarity, the relation of counterparthood. Therefore, a counterpart-theoretic approach should afford a superior solution to Chisholm's Paradox.

This motivation for the counterpart-theoretic solution, though apparently different from Kripke's, is defective in a related way. As I have already noted, neither the argument (CP) nor its alleged equivalent (CP)′ involves the concept of identity, and hence neither involves the concept of possibly being identical with *a*. If (CP)′ constitutes a paradox of vagueness, the vagueness must reside elswhere, in some concept essentially involved in the argument.

In fact, despite Forbes's motivational remarks, in his formal treatment the vagueness is indeed located elsewhere. Specifically, by invoking a counterpart theory in which the counterpart relation is vague, Forbes formally locates the vagueness involved in (CP) in a certain second-order modal concept: the concept of a property's being such that *a* might have had it. Formally, the crucially vague expression involved in (CP)′, according to Forbes's formal treatment, is ⌜◇ . . . *a* . . .⌝, or ⌜it might have been that *a* . . .⌝; the crucially vague concept is that designated by '$\lambda F \Diamond F(a)$'.

Forbes's formal treatment may be correct in imputing vagueness to this modal locution, for if there is any vagueness relevantly involved in Chisholm's Paradox, it can be only in such locutions as this. However, it is not at all true that standard possible-world semantics can accommodate the vagueness of this locution only by treating identity as vague. In fact, even if identity is (incoherently) regarded as vague, that would not be sufficient to impute vagueness to the locution in question, since this locution does not involve the identity predicate. It involves only the sentential possibility operator and the proper name (individual constant) 'a'. We have already seen that the name 'a' is not a source of vagueness. Hence, if there is any vagueness relevantly involved in the modal paradoxes, it resides in the modal operators themselves, and the modal operators are precisely where Forbes's formal treatment ultimately locates the vagueness upon which the paradoxes turn.

We have also already seen that nothing so radical as a departure from standard possible-world semantics in favor of a counterpart-theoretic semantics is called for in order to accommodate vagueness in the modal operators. Standard possible-world semantics can accommodate the relevant vagueness in the modal operators in precisely the way I have suggested: one should treat the accessibility relation between worlds as itself vague (its characteristic function partially defined), so that certain pairs of worlds are neither determinately mutually accessible nor determinately mutually inaccessible. When fully worked out, this involves intransitivity in the accessibility relation via a region of indeterminacy, and hence an abandonment of *S4* modal logic in favor of something weaker or independent (such as the modal system *B*). This approach affords a solution to the modal paradoxes that accommodates vagueness precisely where it must arise, if anywhere, and it does so within the framework of standard possible-world semantics without resorting to the entirely unnecessary and unjustified tack of invoking counterparts in place of cross-world identities. This approach also recognizes a crucial difference between the modal paradoxes and the standard paradoxes of vagueness: the former turn on a fallacy special to modal logic—the fallacy of possibility deletion, or equivalently, the fallacy of necessity iteration.

The counterpart-theoretic approach is not merely unnecessary and unjustified. It is positively misleading, and logically distinctly counterintuitive. I shall develop these criticisms each in turn.

VI

Counterpart theory appears to provide an alternative to standard possible-world semantics that is able to accommodate modal principles like (0), (I), (II), and (III), and their multiple necessitations, within an *S5* framework (i.e., maintaining an equivalence accessibility relation) without generating the paradoxes. Yet as it is typically intended, counterpart theory with respect to artifacts accommodates precisely the opposite of (II): if a wooden table x is originally formed from a hunk of matter y, and y' is any hunk of matter distinct from y, then even if y' substantially overlaps y and is otherwise just like y, x is such that it *could not* have been originally formed from y' instead of from y. The reason for this is that, as it is typically intended, counterpart theory with respect to artifacts includes the basic tenet that possible artifacts formed in their respective possible worlds from distinct (even if substantially overlapping) hunks of matter are always themselves distinct (though they may be mutual counterparts). Thus if x is a wooden table originally formed from a hunk of matter y, and y' is a hunk of matter even only one atom or molecule different from y, the counterpart theorist with respect to artifacts would typically deny that there is a genuinely metaphysically possible scenario, a genuinely possible world, in which the one and only very table x—that very table and no other—is formed from y' instead of from y. The counterpart theorist will insist that, strictly speaking, if we are ever to have one and the very same table x—that very table and no other—existing in a counterfactual scenario that might have obtained, x must be originally formed in that scenario from exactly the same matter, atom for atom, quark for quark, right down to the tiniest of subatomic material components. For this reason, counterpart theory with respect to artifacts is, at bottom, a particularly inflexible brand of essentialism. The counterpart theorist with respect to artifacts can mouth the words 'x might have been formed from y' instead of from y', thereby seeming to advocate (II). But in counting this remark true and therefore assertible, the counterpart theorist means to be committed to nothing more than the availability of a possible scenario in which some table or other sufficiently similar to x—not necessarily x itself—is formed from y'. The counterpart theorist thus says one thing and means another.[16]

Forbes has responded to this objection by claiming that

> whether or not [counterpart] theory admits contingency [of the table x's original matter] . . . turns only on whether or not it [counterpart theory] is consistent with the truth of [the sentence 'x is formed from y, and might have existed without being formed from y'], and by this criterion, counterpart theory admits contingency beyond all question. (''Two Solutions to Chisholm's Paradox,'' p. 179)

This response involves a confusion—or perhaps an equivocation—between two distinct senses in which a theory may be said to ''admit'' or accommodate a principle or proposition.[17] A theory accommodates a proposition p in the primary sense if the theory embraces p itself, that is, if the proposition p is included as a part of the theory (or at least as a logical consequence of the theory in combination with uncontroversial premises). A theory may be said to accommodate a proposition p in a secondary sense if the theory (or the theory in combination with uncontroversial premises) logically entails the metatheoretic proposition that some particular sentence ϕ is true, where ϕ is in fact a formulation of, or expresses, the proposition p. These two kinds of accommodation should be sharply distinguished. Counterpart theory with respect to artifacts can indeed accommodate modal principles like (0), (I), (II), and (III) in the secondary sense. But this sort of accommodation is deceptive, since as it is typically intended, counterpart theory with respect to artifacts fails to accommodate the critical principle (II) in the primary sense. Consider, by analogy, the following simple theory: (1) A table's exact original matter is always an essential feature of the table; (2) snow is white; and (3) the sentence 'Any particular wooden table might have been formed from metal instead of wood' means in English that snow is white. Call this theory 'T'. (To dispel the appearance of inconsistency, imagine the theory T being formulated in Chinese.) The theory T can hardly be said to admit contingency of original matter in any relevant sense, though it does accommodate (II) in the secondary sense. Like counterpart theory with respect to artifacts, the theory T avoids the modal paradoxes by *rejecting* the modal principle (II)—not the *formulation* of (II) given above, but the proposition (II) itself. It may not be entirely futile, but it would be a difficult matter indeed to argue the merits of the doctrine of contingency of original matter with a proponent of T. The advocate of T will apparently join in singing the praises of (II), but the agreement is merely verbal.

Forbes argues that to see counterpart theory on this model, as an inflexible essentialist theory that misrepresents the meanings of modal-operator-discourse formulations of principles such as (II), is

> to think of the extensional sentences of [possible-world discourse] as having some meaning given independently [of modal-operator discourse]. . . . But this conception of [the relation between the two types of discourse] is not very plausible. . . . The threat is that . . . we would have to . . . identify possible worlds with logical constructions of actual entities; and . . . [this identification] has recently been shown [by Alan McMichael] to be problematic. It seems better to think of the meanings of [sentences of possible-world discourse] as being given by those of the modal [-operator-discourse sentences] themselves (so far as this is possible). . . . [In giving the meanings of sentences of possible-world discourse by means of sentences of modal-operator discourse] it would be up to the theorist himself to decide just how to proceed, given his purposes. . . . But from this starting point, one cannot think of the sentences of either [standard or counterpart-theoretic possible world] semantics as yielding perspicuous representations of the 'real' meanings of the modal [-operator-discourse] sentences. . . . Yet Salmon's criticism makes sense only if we think of [possible-world discourse] in these unlikely ways. (ibid., pp. 179–80)

Forbes's conception of the nature and content of possible-world semantics raises large issues concerning the enterprise of semantics generally, issues too broad in scope to be debated adequately in the present forum. It is worth noting, though, that Forbes's conception of the nature of possible-world semantics is distinctly inplausible when extended to temporal semantics for tensed discourse, though Forbes has also suggested that some sort of temporal-counterpart theory may be useful in solving temporal paradoxes analogous to the modal paradoxes.[18] Semantics for tensed discourse usually employs the notion of a time t—perhaps a moment of time or an interval of time—and the relation of earlier-later between times. A semantics for tensed discourse can also be developed using the idea of an *instantaneous total state of the cosmos,* or what I shall call an *i.s.*, and the relation of temporal precedence between successive instantaneous states (assuming no instantaneous state of the cosmos is ever repeated). Using instantaneous states of the cosmos in place of times better emphasizes the analogy between temporal semantics and possible-world semantics. In *i.s.*

semantics for tensed discourse, the semantic attributes of reference, application (of a predicate), and truth value are relativized to *i.s.*'s. On the natural semantic development, a sentence of the form ⌜It has been the case that ϕ⌝ is true with respect to an *i.s.* *i* if and only if ϕ is true with respect to some *i.s.*, or some succession of consecutive *i.s.*'s, that precede *i*. Similar clauses may be given for other temporal operators ('it is going to be the case that', 'it has always been the case that', and so on). Now perhaps the meaning of the phrase 'instantaneous total state of the cosmos' is such that it can be explained, or is in fact learned, only by means of tense or other temporal operators; perhaps not. In either case, the phrase has a relatively clear meaning, and contrary to the spirit of Forbes's remarks, this meaning determines the correct correspondence between a sentence of temporal-operator discourse and the expression of its truth condition in *i.s.*-discourse, not vice versa. It is not the prerogative of the semanticist to devise whatever semantic clauses suit his or her philosophical interests and temperament.

Consider, for example, the sentence 'Bill has been baptized'. On the natural semantic development, this sentence is true in English with respect to the present *i.s.* if and only if Bill is baptized in some prior succession of consecutive *i.s.*'s. It is quite incredible to suppose that a philosopher particularly fond of the idea of cross-time resemblance is free to select some other truth condition for this sentence more to his or her liking. A temporal-counterpart theorist might tell us that on his or her theory, the tensed sentence 'Bill has been baptized' is translated into the following sentence of *i.s.*-discourse:

> In some succession of consecutive instantaneous total states of the cosmos that precede the present instantaneous total state, someone bearing such-and-such a resemblance to Bill, as he presently is, is baptized.

The claim that this sentence means simply that Bill has been baptized is bizarre. Even if Bill is now remarkably like his great-grandfather used to be, the fact that his great-grandfather was baptized has no bearing semantically on the truth in English of 'Bill has been baptized'. Of course, one could decide to use the *i.s.*-discourse sentence displayed above in such a way that it is, in effect, a semantically unstructured *idiom,* one that means simply that Bill has been baptized (in the way that the phrase 'kick the bucket' means to die), but such a decision involves a misleading and radical departure from English. The point of introducing such misleading idioms into semantics would be utterly mysterious. Why not use the original straightforward formulations in temporal-operator discourse?

The fact is that sentences of *i.s.*-discourse do not function in *i.s.* semantics as unstructured idioms, whether standard *i.s.* semantics or temporal-counterpart-theoretic *i.s.* semantics. On the contrary, it is the very internal semantic structure of *i.s.*-discourse sentences that makes *i.s.*-discourse suitable for the enterprise of doing a systematic semantics for a tensed language. In fact, the very existence of the theory of instantaneous states and cross-time counterparthood offered by the temporal-counterpart theorist gives the lie to the claim that the meaning of an *i.s.*-discourse sentence (such as the one displayed above) is fixed by its alleged analogue in tensed discourse

('Bill has been baptized'). Rather, the meaning of an *i.s.*-discourse sentence is fixed in the usual way, by the meanings of its components—including the meanings of 'instantaneous total state of the cosmos', 'precede', 'present', and 'resemblance', as they arise in formulating the temporal-counterpart theory. Thus the *i.s.*-discourse sentence displayed above has a meaning that involves the temporal- counterpart theorist's concept of cross-time resemblance. It cannot mean the same thing as the tensed discourse sentence 'Bill has been baptized', for the proposition that Bill has been baptized involves no concept of resemblance, and hence it does not involve the particular resemblance concept given in the temporal-counterpart theory and expressed in the proposed *i.s.*- discourse translation. The same is true if reference to persisting objects is replaced with reference to temporal stages of persisting objects, and if cross-time resemblance is replaced with a notion of spatiotemporal continuity.

Consider now a contemporary follower of Heraclitus who holds that one cannot step into the same river in the same spot twice—i.e., in two different instantaneous states of the cosmos—because new water is continuously flowing through. The contemporary Heraclite (perhaps unlike Heraclitus himself) believes that, in general, the matter that constitutes an object (e.g., the water in a river) is a permanent and unchanging feature of the object. A contemporary Heraclite may devise an elaborate temporal-counterpart theory with respect to material objects to make it possible to "speak with the vulgar"—to utter sentences like 'The Mississippi River once had cleaner water flowing through it than it now has'—but then this clever Heraclite does not mean by this sentence what the rest of us mean, or what the sentence itself means. Any such pronouncement in tensed discourse by this philosopher is merely a verbal camouflage. When the Heraclite says 'The Mississippi is the same river today as yesterday', he or she does not mean the word 'same' in the "strict and philosophic sense," but rather in what he or she believes is a "loose and popular sense," i.e., as a word for temporal counterparthood.[19]

The phrase 'possible world' may not be as clear in meaning as the phrase 'instantaneous total state of the cosmos', but there are a number of conceptions of possible worlds presently in vogue, each of which is clear enough to substantiate my labeling of counterpart theory as a particularly inflexible brand of essentialism. Possible worlds are variously construed as maximal compossible sets of propositions (Robert Adams), total histories the world might have had (Saul Kripke), maximal states of affairs that might have obtained (Alvin Plantinga), total states the cosmos might have been in (Saul Kripke, Robert Stalnaker), total scenarios that might have obtained (myself). For present purposes, these need not be regarded as competing conceptions of possible worlds. If the phrase 'possible world' is unclear in meaning, any of these clearer phrases may be substituted.[20] It is of course true that each of these explications of what a possible world is involves notions from modal-operator discourse: *possibile*, *compossibile*, or *might have*. The notion of a possible world is defined in terms of concepts like *might have*, rather than vice versa. In this sense, the meanings of sentences of possible-world discourse are not "given independently" of modal-operator discourse. But they do have meaning, and just as in the case of tensed and *i.s.*-discourse, the meanings of sentences in possible-world discourse determine the semantic clauses for modal-operator discourse, and not the other way around. It is not the prerogative of the semanticist to stipulate whatever semantic clauses suit his or her philosophical interests and temperament. The sentence 'Bill might have been a robot' is true if and only if there is a possible scenario—or a possible history, or a possible state of affairs, or a possible state of the cosmos—in which Bill is a robot. The availability of a possible scenario in which not Bill but something rather like Bill in such-and-such respects is a robot is entirely irrelevant.

As in the case of tensed and *i.s.*-discourse, the very existence of the counterpart theorist's theory of possible worlds and counterparthood as a relation of cross-world similarity gives the lie to Forbes's claim about what fixes the meanings of the possible-world-discourse sentences that allegedly give the truth conditions of sentences in modal-operator discourse. The meanings of possible-world-discourse sentences are fixed in the usual way, by the meanings of their grammatical components—including the word 'counterpart', as it arises in the counterpart-theorist's formulation of his or her theory. Another indication of this is the fact that Forbes relies on possible-world discourse, rather than on untutored modal-operator-discourse intuition, as the court of last arbitration to determine the fine detail of which inferences in modal-operator discourse are to count as a valid and which are to count as invalid. The very enterprise of a systematic possible-world semantics for modal logic would be impossible if the sentence giving the truth-in-a-model condition for a particular object language sentence has its meaning fixed by the object language sentence itself.

The fact that counterpart-theoretic semantics misinterprets modal-operator discourse is made evident by the logic the former imposes on the latter. We have already seen that if counterpart theory is devised in such a way as to preserve *S5* modal propositional logic, it typically invalidates certain special cases of an intuitively valid modal variant of universal instantiation, (MUI), which permits the inference from $\ulcorner$Necessarily, everything is $\phi\urcorner$ to $\ulcorner$Necessarily, if α exists, then it is $\phi\urcorner$, where α is a simple singular term. (See note 12.) The misinterpretation of modal-operator discourse is made even more plain if a predicate for the intra-world analogue of counterparthood is added to the latter. For if the counterpart theory includes the usual constraint that all of a possible individual's existing counterparts within a given world are counterparts at that world of anything that the object is itself a counterpart of at that world, then the theory validates the intuitively fallacious inference from

$$\Diamond(\exists x)\,[x \text{ is a counterpart of } a \;\&\; F(x)]$$

to

$$\Diamond F(a).$$

The validity of this inference in counterpart-theoretic modal logic illustrates the weak interpretation placed on simple possibility sentences such as 'Bill might have been a robot'. Normally, if someone were to utter this sentence, he or she would mean something considerably stronger than, if not entirely independent of, whatever may be entailed by the claim that there might have been a robot counterpart of Bill.[21]

VII

The various explications of possible worlds given in the preceding section support the legitimacy of the idea of an impossible world, as well as the intransitive-accessibility account of the modal paradoxes. Just as there are such things as maximal compossible sets of propositions, there are also such things as maximal consistent but not compossible sets of propositions. If there are such things as total histories the world might have had, maximal states of affairs that might have obtained, and total states the cosmos might have been in, then there are also such things as total histories the world could not have had, maximal states of affairs that could not have obtained, and total states the cosmos could not have been in. Some of these impossible worlds are such that they might have been possible worlds instead of impossible worlds; their modal status as possible or impossible is a contingent feature of them. In fact, among the impossible worlds are those that might have been possible, those that could not have been possible but might have been such that they might have been possible, those that could not have been such that they might have been such that they might have been possible but might have been such that they might have been, and so on, perhaps to infinity. In any case, for some fairly large finite number n, there are worlds that are not possible in the nth degree (not possibly possibly . . . (n - 1 times) . . . possible), but that might have been, i.e., they are possible in the (n + 1)th degree. Consider, for example, the possible total scenario (history of the world, and so on) w_1 in which everything is just as it actually is except that the table a is formed from the hunk of matter h_1 instead of from hunk h (and whatever other differences are required by this difference in order to ensure genuine possibility). The total scenario (history, and so on) w_2, that is just like w_1 except that table a is formed from hunk h_2 instead of from h_1, is a possible scenario relative to scenario w_1. That is, in scenario w_1, scenario w_2 is a possible scenario. Eventually, there is a total scenario w_m that is not possible relative to the actual total scenario, i.e., that is not a genuinely possible scenario, but that might have been. That is, w_m is possible in the second degree, but not in the first. Similarly, as we have seen, the total scenario w_{2m}, in which table a is formed from hunk h_{2m}, is possible in the third degree, but not in the second, and hence not in the first. Even the total scenario w_n, in which table a is formed from entirely different matter, is possible in some sufficiently large degree, though presumably it is not possible in only the second or third degree.

Thus far I have ignored the fact that certain sentences may be neither true nor false, perhaps in virtue of a false presupposition, as with the occurrence of a nonreferring definite description (e.g., Russell's 'The present king of France is bald'), or in virtue of the occurrence of a vague predicate (e.g., 'Louis is bald', where Louis has enough hair on his head so that he is not determinately bald but not enough hair so that he is determinately not bald). When we take note of this fact, it emerges that possible worlds are not *maximal* or *total* in the ordinary sense. For example, the proposition that the present king of France is bald is arguably neither true nor false in the actual world, so that the set of true propositions includes neither this proposition nor its negation (the proposition that the present king of France is not bald). If the actual world is just the set of true propositions, then a possible world may be a compossible set of propositions that falls short of being genuinely maximal. Similarly, if the actual world is the true history of the world, or the total state the cosmos is in, and so on, then since the true history of the world, and the total state the cosmos is in, include nothing that determines that the present king of France is bald and also nothing that determines that the present king of France is not bald, a possible world may fall short of being *total* in the sense of deciding 'yes' or 'no' on every possible question of fact. Still, of course, a possible world must approach maximality or totality as closely as possible. A possible world must be *maximal* or *total* in the weaker sense that for any proposition or question of fact p left undecided, there must be enough propositions or questions of fact decided (e.g., that there is no present king of France, or that the number of hairs on the top portion of Louis's head per square inch of surface area is n) to determine that there is no objective fact of the matter concerning p.

This observation supports the feasibility of the indeterminate-accessibility account of the modal paradoxes sketched in section II above. A total (in the weak sense) scenario w' is accessible to a total scenario w if and only if it is a fact in w that w' might have obtained. If the notion of possibility is itself vague, there will be total (in the weak sense) scenarios w_k such that the actual total scenario includes nothing about whether w_k might have obtained or not. World w_k would thus be neither determinately accessible nor determinately inaccessible to the actual world, in the same way that some people are neither determinately bald nor determinately not bald in the actual world.

Forbes objects to these conceptions of what possible worlds are by endorsing a criticism, due to Alan McMichael,[22] that a theory of such entities as maximal compossible sets of propositions or maximal states of affairs that confines itself to things that actually exist—an *actualist* theory of such entities—is problematic. McMichael's criticism, very briefly, is this. The following sentence involving nested modalities is true:

> S: It might have been the case that there exists someone who: (a) does not actually exist; (b) is bald; and (c) might have existed without being bald.

Following the standard approach rather than the counterpart-theoretic approach, S is true if and only if there is a possible world w in which there exists an individual x such that (a) x does not exist in the actual world; (b) x is bald in w; and (c) there is a world w' accessible to w in which x exists but is not bald. McMichael argues that this truth condition apparently cannot be fulfilled within an actualist theory of possible worlds. Suppose for example that possible worlds are identified with maximal compossible sets of states of affairs. Then in order for S's truth condition to be fulfilled, it seems there would have to be one such set w that includes the state of affairs of there existing an individual x who does not actually exist and who is bald, and another such set w' that includes the states of affairs of x's existing and x's not being bald. But since x does not actually exist, there are no such states of affairs as x's existing or x's not being bald, and hence no such set as w'.

The argument here is fallacious, though exposing the fallacy is a delicate matter. No such set as w' is required to exist for the truth of S. Exactly what is required is the existence of a maximal compossible set w of states of affairs that includes the complex state of affairs of there existing an individual x such that: (a) the state of affairs of x' s existing does not actually obtain; (b) x is bald, and (c) there is a maximal compossible set of states of affairs w' that includes the states of affairs of x' s existing and x' s not being bald. This in turn requires the existence, and the possibly obtaining, of the state of affairs of there existing some individual or other who is bald, whose existence does not actually obtain, and whose existence while not being bald might have obtained. But it in no way requires the existence of either the state of affairs of this nonactual individual's existence or of his or her not being bald.

We may put the matter this way: Suppose that possible worlds are maximal compossible sets of propositions. Now it has been observed by a number of philosophers, including McMichael, that within an actualist framework, a set of possibly true propositions may be maximal (in either the strong or weak sense) and yet may include some particular existential generalization without including any singular instance of it. This occurs when the existential generalization is such that no actual entity yields an instance that is possibly true. For example the proposition expressed by '($\exists x$) [x does not actually exist]', though false, is such that it might have been true. Since there is no actual entity that can serve as the relevant constituent of a possibly true singular instance of this existential generalization, however, there is no singular instance that is possibly true. Now in order for S to be true, there must be a maximal (in the weak sense) compossible set w of propositions that includes the proposition expressed by the sentence:

($\exists x$) [x does not actually exist & x is bald & ($\exists w'$) (w' is possible & the proposition *that x exists and is not bald* $\in w'$)].

As was just indicated, w will include no singular instance of this proposition, since there are none to be included. More importantly, however, the sentence displayed above is equivalent to the following:

($\exists w'$) [w' is possible & ($\exists x$) (x does not actually exist & x is bald & the proposition *that x exists and is not bald* $\in w'$)].

This sentence also expresses precisely the sort of existential proposition that is possibly true but has no possibly true singular instance. What the truth of S requires is the existence, and the possible truth, of this existential proposition; it does not require the existence of any singular instance of it.

VIII

My criticisms of counterpart theory are independent of the logic of vagueness that may be supplied to supplement the theory. In fact, the logic of vagueness is all but irrelevant to the main idea behind a counterpart-theoretic approach to the modal paradoxes. Forbes proposes treating the counterpart relation as itself vague and a matter of degree. Essentially the same account results from speaking of *determinate counterparts* in place of counterparts *simpliciter*—where, if it is indeterminate to some degree whether x is a counterpart of y, then it is determinately true that x is not a determinate counterpart of y.

Following J. A. Goguen,[23] Forbes proposes to treat the sort of vagueness found in concepts like that of being short or that of being similar by means of infinitely many *degrees of truth and falsehood* in place of the conventional all-or-nothing dichotomy of truth and falsehood. Accordingly, on Forbes's theory, a sentence containing a vague term may be wholly true, almost wholly true, more true than false, equally as true as false, more false than true, almost wholly false, or wholly false. Degrees of truth and falsehood are represented by means of the real numbers between 0 and 1, inclusive, where 1 represents complete truth, 0 represents complete falsehood, and the sum of the degree of truth of a sentence and its degree of falsehood (the degree of truth of its negation) is 1.

Many find the idea of a sentence being (unambiguously) partly true and partly false grating. Truth and falsehood appear to be mutually exclusive absolutes; nothing "partly false" is genuinely and literally true in the ordinary sense. But it would be a mistake to conclude that the concept of degrees of truth and falsehood is utterly without merit in the logic of vagueness. To illustrate: suppose there are two men, Smith and Jones, for whom it is vague—indeterminate, neither true nor false, there is no objective fact of the matter—whether either is bald. Ordinarily, though neither of the two men has little enough hair to qualify as determinately bald, one of the two, say Smith, will be "balder" than the other, in the sense that Smith has proportionately less hair on the top portion of his head per square inch of surface area than does Jones. Neither is determinately bald, but Smith is "closer" to being determinately bald than Jones is. Although the adjective 'bald' is neither true nor false of both men, it is closer to being true of Smith than it is to being true of Jones. The sentence 'Smith is bald' is closer to being true than is the sentence 'Jones is bald', though neither sentence is true (and neither is false). One may decide to put this another way by saying that both sentences partake of a certain "degree of truth" less than the "maximal degree," and that the first is "more true" than the second. It does not follow, of course, that the first sentence is true *simpliciter*—any more than Smith's being taller than Jones entails that Smith is tall.

Similarly, though the proportion of hair on Smith's head does not fall squarely into either the category *bald* or the category *not bald*, in all likelihood it is closer to one end of the scale than to the other. Suppose that Smith is such that if he were to lose just a very few more strands of hair, he would become determinately bald rather than indeterminate with respect to baldness, whereas if he grew as many strands of hair, he would remain indeterminate with respect to baldness. Then the sentence 'Smith is bald', though neither true nor false, is closer than its negation to being true; it is closer to being true than it is to being false. One might put this by saying that it is "more true than false," though strictly speaking, of course, it is neither. This interpretation provides significance to the notion of "degrees of truth" in the logic of vagueness.

The important point about this construal of a degrees-of-truth approach should not be obscured by the somewhat misleading jargon of a sentence being "more true than false" or "more false than true." A sentence that is true *simpliciter* is now being said to be "wholly" or "completely" true, or true "to the maximum degree," and a sentence that is false *simpliciter* is now being said to be "wholly" or "completely" false, or false "to the maximum degree." A sentence said to be only "partly true," or "less than but almost wholly true," is not true at all, and a sentence said to be "less than but almost wholly false" is not false at all. On the construal I am suggesting of the degrees-of-truth approach, the classical three-way division among *true, false,* and *neither true nor false* is built into that approach—as maximal truth, maximal falsehood, and everything in between. The range of degrees of truth between maximal falsehood and maximal truth, exclusive, are nothing more than gradations of the traditional category of *neither true nor false,* so that classical three-valued logics emerge as subtheories of analogous degrees-of-truth approaches. If a sentence is said to be "more true than false," or "almost but not quite wholly true," it is neither true nor false, though in the sense sketched above it is closer to being true than to being false.

This interpretation of the degrees-of-truth machinery evidently clashes with Forbes's intent. First, Forbes has denounced the traditional three-way division among *true, false,* and *neither* as arbitrary, whereas on the construal suggested this division is embedded in the degrees-of-truth approach.[24] More important, Forbes's definition of *validity* in the logic of vagueness does not accord well with the suggested construal of the nature of the truth value status represented by real numbers between 0 and 1. Forbes calls an argument or inference pattern 'valid' roughly, if in any model, its conclusion is at least as true as the least true of its premises (more accurately, if in any model, the degree of truth of the conclusion is at least as great as the greatest lower bound of the degrees of truth of the premises). This leads him to brand *modus ponens* an invalid inference pattern, since in his logic of vagueness, a conditional that is neither (wholly) true nor (wholly) false may be closer to being true (have a "greater degree of truth") than either its antecedent or its consequent taken individually. Forbes calls the inference pattern of *modus ponens* 'the fallacy of detachment' and blames the standard sorites paradoxes on this alleged fallacy. He sees the choice between the accessibility solution to the modal paradoxes and the counterpart-theoretic approach as a choice between rejecting *S5* modal logic while consequently treating the two arguments (CP) and (CP)′ differently, on the one hand, and rejecting *modus ponens* while treating the two arguments equivalently, on the other. Since *modus ponens* must be rejected in any case, Forbes argues, the counterpart-theoretic approach is superior to the accessibility approach. It retains *S5* modal logic while allegedly reducing (CP) to a familiar paradox of vagueness in classical propositional logic.

Can it be that *modus ponens* is a fallacious inference pattern and that this is the fallacy involved in the traditional sorites paradoxes, such as the paradox of the short person? I can think of no inference pattern whose validity is more obvious than *modus ponens*. Rather than place myself in the hopeless position of Achilles, though, I will say here only that the validity of *modus ponens* is certainly more intuitively obvious and compelling than the alleged validity of the *S4* axiom of modal logic, or equivalently, the inference pattern of necessity iteration (possibility deletion). If the choice were as Forbes sees it, the accessibility approach should be the winner beyond all question!

In fact, though, Forbes has posed a false dichotomy. An inference pattern is properly *valid* if and only if it is truth-preserving. i.e., if and only if for every instance, its conclusion is true in every model in which its premises are true. This is the proper notion of validity even in the logic of vagueness. In a degrees-of-truth logic of vagueness, as I have proposed construing it, an inference pattern is valid (properly so-called) if and only if it preserves "complete truth" or "truth to the maximum degree." By this criterion, *modus ponens* is unquestionably valid even on the degrees-of-truth account. Why place the blame for paradoxes of vagueness on *modus ponens*? In fact, the traditional sorites argument in classical propositional logic is perfectly valid. What goes wrong in a standard sorites paradox, such as the paradox of the short person, is not that the argument is invalid, but that it is unsound. Not all of the conditional premises are (wholly) true. At least two are neither true nor false. In the terminology of the degrees-of-truth approach, at least two conditional premises are "partly true and partly false," or "less than wholly true but true to some degree." The paradox of the short person is dissolved by noting that one should not attempt to establish conclusions by reasoning from premises that are untrue—even if they may be said to be "almost wholly true" in the sense sketched above. Almost is simply not good enough.

The intransitive-accessibility account rejects *S4* and accommodates both of the modal principles (II) and (III) in both the primary and secondary senses, whereas the counterpart-theoretic approach, as I have devised it, retains *S5* but fails to accommodate (II) in the primary sense. This is the real choice. The paraphernalia of degrees-of-truth, the alleged loose and popular sense of 'identity', cross-world counterparts, so-called identity from the point of view of a particular world, and the rest, tend to obscure the point.

IX

On the intransitive-accessibility account that I advocate, there are distinct yet purely qualitatively identical worlds in which the very same matter exists in exactly the same configuration, and in which all matter undergoes exactly the same physical processes, down to the finest detail, throughout all of time. This is not quite the same as the apparent conclusion ($C2$) of the second version of the Four Worlds Paradox, which is surely unacceptable. It is open for the accessibility theorist to argue (as I have elsewhere) that any two distinct such worlds are mutually inaccessible and are not both (determinately) accessible to the actual world. In modal-operator discourse, the accessibility account yields the following conclusion, where p is a (the) materially complete proposition that would have been true if table a had been formed from hunk h'' $[=h_{m-1}]$ rather than from hunk h:

$$(C3) \qquad \Diamond[p \,\&\, M(a, h'')] \,\&\, \Diamond\Diamond[p \,\&\, {\sim}M(a, h'')].$$

Some philosophers have objected to this conclusion on the basis of a principle of the identity of materially indiscernible worlds, i.e., worlds in which the same materially complete proposition is true.[25] Conclusion ($C3$) is in fact perfectly compatible with any reasonable version of the Identity of Indiscernibles. Moreover it can be modified to show that the principle of the identity of materially indiscernible worlds in fact contradicts classical Indiscernibility of Identicals. This can be seen through consideration of another conclusion correctly obtainable from the assumptions of the Four Worlds Paradox:

$\Diamond\Diamond M(a, h')$.

If table a had been formed from hunk h'' instead of from hunk h, then it *would have been* possible for it to have been formed from hunk h' instead of from hunk h''. Let p' be a (the) materially complete proposition that would have been such that it would have been true if table a had been formed from hunk h', if only table a had been formed from hunk h'' instead of from hunk h. Take care here. Since a could not have been formed from h', it is arguable that any proposition, and hence any materially complete proposition, would have been true if a had been formed from h', or alternatively, that no proposition, and hence no materially complete proposition, would have been true if a had been formed from h'. Though it is not *in fact* possible for a to have been formed from h', it *might have been* possible, and indeed it would have been possible if only a had been formed from h''. Proposition p' is a (the) materially complete proposition such that: if a had been formed from h'', then it would have been the case that if a had been formed from h', then p' would have been true. Then we have:

($C4$) $\Diamond[p' \mathbin{\&} {\sim}M(a, h')] \mathbin{\&} \Diamond\Diamond[p' \mathbin{\&} M(a, h')] \mathbin{\&} {\sim}\Diamond[p' \mathbin{\&} M(a, h')]$.

More intuitively there is a world w accessible to the actual world in which a table distinct from a is formed from hunk h'. There is also a world w' accessible to some world accessible to the actual world (through none accessible to the actual world itself) which is exactly like w in every detail concerning the very matter it contains, with its exact configuration and causal interconnections throughout time, atom for atom, quark for quark, but in which a is the table formed from hunk h'. Worlds w and w' are materially, and hence also purely qualitatively, indistinguishable. Exactly the same material facts obtain in both. Though they are materially indiscernible, they differ in their accessibility relations. World w is accessible to the actual world, whereas world w' is not. Hence, by the Indiscernibility of Identicals, the two worlds are distinct.

An unbridled principle of the identity of materially indiscernible worlds is refuted by the example of the worlds w and w'. Though materially indiscernible, the worlds w and w' are indeed discernible, and not merely by their accessibility relations to the actual world. They also differ as regards which facts obtain in them. World w' includes the fact that a is the table formed from hunk h', whereas world w excludes this. In w, some table distinct from a is the table formed from hunk h'. It follows again by the Indiscernibility of Identicals that the worlds w and w' are distinct.

The temptation to identify the worlds w and w' may stem, in part, from misconceiving possible worlds as material objects, or as entities made solely of matter.[26] Possible worlds are abstract entities whose structure comes from the facts that obtain in them. We saw in section VI that worlds may be conceived as maximal (in the weak sense) consistent sets of propositions, or total (in the weak sense) histories or states of the cosmos, or maximal states of affairs, or total scenarios, and so on. Consider the first conception: worlds as maximal consistent sets of propositions. Then w and w' are maximal consistent sets that both include the materially complete proposition p' as an element. The set w' includes the further proposition that the table formed from hunk h' is a, whereas the set w includes the further proposition that the table formed from hunk h' is some table distinct from a. Both sets are maximal consistent. Thus both are equally legitimate as worlds *per se*. Through they are not disjoint, they are unquestionably distinct sets.

Similar remarks may be made with respect to any of the alternative conceptions of the worlds w and w'. In fact, these various conceptions of worlds strongly suggest an alternative to simple material indiscernibility as a criterion for identity between worlds. They suggest a principle of the identity of *factually* indiscernible worlds, worlds in which the very same facts obtain. One might also endorse an independent principle of the identity of mutually accessible materially indiscernible worlds (a version of the supervenience thesis mentioned at the end of section I above) or a principle of the identity of materially indiscernible worlds accessible to the actual world. On any of the conceptions of worlds mentioned here, an unrestricted principle of the identity of simply materially indiscernible worlds is straightforwardly false.

X

Forbes has raised a second sort of objection to the intransitive-accessibility solution to the modal paradoxes. He argues that if we consider essentialist principles like (III), "we see that there is a conceptual character to such claims," and that metaphysical necessity is "fundamentally an *a priori* matter, to do with the content of our concepts [for example, our concepts of a table and of original matter], even though with the addition of *a posteriori* information, necessary *a posteriori* truths can be inferred."[27] Furthermore, "any *a posteriori* truth p necessary at the actual world is so by being true at the actual world and by some conceptual [*a priori*] truth's entailing that p's truth makes it necessary."[28] Since metaphysical necessity is thus the product of conceptual *a priority*, Forbes argues, every instance of the *S4* axiom schema is indeed true. For if it is conceptually *a priori*, and consequently necessary, that p, then it is also conceptually *a priori* that it is conceptually *a priori* that p. And if it is necessary but *a posteriori* that p, then it is nevertheless conceptually *a priori*, and consequently necessary, that if p then it is necessary that p. From this it follows (in even the weak system T of modal propositional logic) that if it is necessary but *a posteriori* that p, then it is still necessary that it is necessary that p.

It may be true that conceptual *a priority* entails metaphysical necessity, in the sense that (with somewhat rare, and for present purposes irrelevant, exceptions) anything that is conceptually *a priori* is generally *ipso facto* metaphysically necessary. Probably something like this accounts for the fact that (II) is not only necessarily true, but it is also necessary that it is necessarily true, and it is necessary that it is necessary that it is necessarily true, and so on. As Forbes acknowledges in presenting his argument, there are examples—coming primarily from the work of Kripke—of propositions that are metaphysically necessary yet conceptually *a posteriori*. With respect to these examples, the argument that *a priori* necessity iterates—the argument that if it is necessary, because *a priori*, that *p*, then it is also necessary that it is necessary that *p*, and so on—is inapplicable. The argument is inapplicable precisely because the examples in question, though necessary, are not *a priori*, and hence not necessary-by-virtue-of-being-*a-priori*.

The propositions that the intransitive-accessibility account holds to be necessary but not doubly necessary (for example, the proposition that table *a* is not originally formed from hunk h_m) are precisely certain *a posteriori* propositions whose necessity is derived by means of *a priori* modal principles like (III) taken together with certain further information, at least some of which is not *a priori*. That is, the propositions that the intransitive-accessibility account holds to be necessary but not doubly necessary are propositions of precisely the sort that Kripke cites as necessary yet *a posteriori*. The *a priori* principle (III) might be used to establish the necessity of table *a*' s not originating from hunk h_m, but the fact that *a* does not thus originate is itself unquestionably empirical and not *a priori*.

In fact, not even the conditional 'If table *a* is not originally formed from hunk h_m, then it is necessary that *a* is not originally formed from h_m' is *a priori*. For all that is known *a priori*, table *a* may have originated from hunk h_1, in which case *a* would still not have originated from hunk h_m, although it would then be possible for *a* to have thus originated.[29] The necessary *a posteriori* truth that table *a* is not formed from hunk h_m is thus a counterexample to Forbes's claim concerning the source of necessary *a posteriori* truths. Since the conditional proposition that if *a* is not formed from h_m then *a* is necessarily not thus formed is not *a priori*, it cannot be entailed by any conceptual *a priori* truth.

That *a*' s not originating from h_m is in fact necessary yields no reason to suppose that it must also be doubly necessary, triply necessary, and so on. Indeed, the fact *a* does not originate from h_m might not have been necessary at all.

The accessibility account rejects the *S4* axiom in its unrestricted form, but the account allows that there may be interesting special cases of necessity iteration that are logically valid. For example, it may be that, as Forbes's argument suggests, necessity iteration is legitimate whenever the proposition in question is necessary by virtue of being *a priori*. Certainly necessity iteration is legitimate with respect to purely mathematical propositions and (classical) logical truths. Maybe here is a legitimate restricted version of the *S4* axiom schema, or the rule of necessity iteration. Are there others? Necessity iteration is fallacious with respect to certain *a posteriori* propositions, but are there any necessary *a posteriori* propositions with respect to which necessity iteration is a legitimate logical inference? For example, Kripke and Putnam have argued that it is necessary, even though *a posteriori*, that cats are animals, or at least that cats are not robots. Presumably, they would argue that it is even necessary that it is necessary that cats are not robots. Does the latter modal fact follow logically from the former, taken together with certain information concerning the nature of the proposition that cats are not robots? If so, what is it about the proposition that cats are not robots that allows for necessity iteration as a logical inference, whereas necessity iteration with respect to other *a posteriori* propositions is fallacious?

The questions raised here seem to be worthy of further research. These and other challenging philosophical questions arise directly from the modal paradoxes. This alone makes the paradoxes deserving of our attention.

APPENDIX: THE DETERMINACY OF IDENTITY

The proof that identity is nonvague and either determinately true or determinately false for any pair of objects of any kind whatsoever proceeds from the observation that if there is a pair of objects, *x* and *y*, of which the 'is' of identity is neither determinately true nor determinately false (i.e., there is no objective, determinate fact of the matter whether *x* and *y* are numerically identical), then since the 'is' of identity is absolutely determinately true of the pair $<x, x>$, the two pairs must be different pairs of objects. It follows that the objects *x* and *y* are themselves distinct. In that case, the 'is' of identity is, contrary to the hypothesis, defined as determinately false for the pair $<x, y>$. Therefore, there is no pair of objects of any kind for which the question of their identity is metaphysically indeterminate.[30]

Although this proof is very convincing—in fact, to my mind, conclusive—I have found that (like most arguments against firmly entrenched philosophical views) it does not always convince. By far the most common objections I have encountered are based on the contention that the proof relies on principles of classical reasoning, whereas the view it purports to refute demands some special nonclassical logic of vagueness. Hence, it is worth emphasizing that the proof does not illegitimately assume or presuppose classical two-valued logic. To assume that every identity proposition is either true or false would certainly be eristically illegitimate, since the argument is advanced against a view that requires a nonclassical, nonbivalent logic. The critical move in the proof is a simple Leibniz's Law inference from an assumption of the form $\ulcorner\alpha$ has a property F that β does not have$\urcorner$ to its trivial consequence $\ulcorner\alpha \neq \beta\urcorner$. Even on the view being disputed, any inference from something assumed to be true is legitimate if the inference pattern is such as to preserve truth (or such as to preserve "determinate truth," or "complete truth," or "truth to the maximum degree," and so on). Analogously, the term 'bald' (in the sense of 'nearly absolutely bald') is unquestionably vague, in that there are (or at least there could be) individuals who have very little hair on their heads but just enough so that it is neither true nor false (vague, indeterminate, there is no objective fact of the matter) that they are bald. It would be illegitimate to assume that every proposition concerning whether someone is bald is either true or false. A nonclassical, nonbivalent logic is needed in order to reason properly with respect to such propositions. Despite this feature of the term 'bald', if one assumes for the sake of argument (e.g., for a *reductio* argument or for a conditional proof) that Harry has a full head of hair, it is perfectly legitimate to infer that Harry is not bald, for we have assumed as determinately true information that is such that if it is determinately true, then so is the proposition that Harry is not bald. There can be no question but that the Leibniz's Law inference invoked in the proof of the determinacy of identity is likewise such as to preserve determinate truth, and is therefore likewise legitimate. Whatever x and y may be, they are not one thing if they differ in any way, since any one thing has every property it has. Nothing could be more trivial.[31]

The critical premise involved in the proof is the assumption that the 'is' of identity is determinately true of any object and itself, and determinately false of any pair of determinately distinct objects. Lest anyone wish to challenge this assumption, it is important to recall Kripke's powerful '*schmidentity*' method of philosophical argument (which ironically applies virtually unchanged to the present case):[32] We may invent a new sense of 'is'—the 'is' of schmidentity—such that our assumption is true solely by stipulation for this new sense. Then we may prove, by the now familiar argument, that schmidentity is fully defined and determinate for any pair of objects. Yet this allegedly new sense of 'is' is precisely one that gives rise to the very sorts of problems for which the theory of indeterminate identity was introduced in the first place. Who cares about any other alleged sense of 'is' when our concern is with a question of identity? What's so important about x and y being neither determinately "identical" nor determinately "distinct" in some other sense if they are determinately not one and the very same but two, determinately not schmidentical? Nothing. Where one's concern is with a question of numerical identity, *almost* doesn't count. In fact, it doesn't even make sense.[33]

There is an alternative way of constructing the proof, one which applies the Leibniz's Law inference directly to the objects x and y: Suppose again that it is indeterminate whether x and y are identical. Then x and y differ in that x is determinately identical with x, whereas y is not. That is, y does not have x's property of being such that the 'is' of identity is determinately true of the ordered pair of x together with it. Hence, contrary to the hypothesis, x and y are determinately distinct.

This alternative construction reveals that the general form of the argument is essentially that used in proving the necessity of identity as a theorem of quantified modal logic: for every x and every y, if $x = y$, then it is necessary that $x = y$.[34] More analogously, the argument parallels the proof of the contrapositive of the necessity of identity, a theorem that Alonzo Church has called 'Murphy's Law of Modality': For every x and every y, if it is possible that $x \neq y$, then (since y does not have x's property of being necessarily identical with x) $x \neq y$.

Church has recently used this general form of argument to argue that if quantification into propositional attitude contexts is accepted as meaningful together with the usual laws of classical logic, then it is very likely that for every x and every y, if someone believes that $x \neq y$, then $x \neq y$.[35] The general argument can also be used to establish—or at least to argue compellingly for—a number of other philosophically interesting and highly controversial (in some cases, nearly universally denied) theses, such as the following:

T1: For every x and every y, if $x = y$, then whenever x exists, $x = y$;

T2: For every x and every y, if $x = y$, then if one believes anything at all involving x, one knows that $x = y$;

T3: For every x and every y, the question of whether $x = y$ is not a matter of decision, convention, or convenience, nor of elegance, simplicity, or uniformity of theory;

T4: For every x and every y, if $x = y$, then x is the only possible individual that could possibly have any metaphysically relevant "claim" or "title" to be y;

T5: For every x and every y, the question of whether $x = y$ does not turn on any fact concerning anything other than x and y;

T6: For every x and every y, if $x = y$, then the fact that $x = y$ does not require any "criteria of identity" for things of x's sort or kind;

T7: For every x and every y, if $x = y$, then the fact that $x = y$ is not grounded in, or reducible to, qualitative nonidentity facts about x and y other than x's existence, such as facts concerning material origins, bodily continuity, or memory;

T8: For every x and every y, if $x = y$, then the fact that $x = y$ obtains by virtue of x's existence, and not at all by virtue of any other qualitative nonidentity facts about x and y, such as facts concerning material origins, bodily continuity, or memory;

T9: For every x and every y, if one knows that $x = y$, then one knows this (primarily) solely by logic and by one's acquaintance with x, and not by knowing qualitative nonidentity facts about x and y, such as facts concerning continuity, location, or qualitative persistence or similarity.[36]

Each of these theses is diametrically opposed to the views, theories, or presuppositions of some major segment of the contemporary analytic philosophical community. Much of the literature on cross-time identity (and especially on personal identity), for example, presupposes the opposite of one or more of theses T6, T7, and T8. Many of the most widely held theories in this literature involve denying several (and in some cases all) of the remaining theses. Nearly the same is true of much of the literature on cross-world identity. In particular, that cross-world and cross-time identity facts are grounded in nonidentity facts is a recurrent theme in Forbes's work.[37] Although it is evidently not widely recognized, each of the theses mentioned is in fact, despite its unpopularity, a virtual consequence of Leibniz's Law together with some trivial feature of the reflexive law of identity. The trick (if there is any) is to extract the right property of x from the relevant trivial feature of the law (or proposition or fact) that $x = x$.

Consider thesis T7: Whatever x may be, the trivial fact that $x = x$ is not at all grounded in, or reducible to, any facts about x like those concerning x's material origins, x's bodily continuity through time, or x's memory of past experiences. If the fact that $x = x$ is grounded in any other fact about x, it is only grounded in the mere fact that x exists. Thus x has the complex property of being such that the fact that x is identical with it is not grounded in any qualitative nonidentity facts about x other than x's existence. Hence, by Leibniz's Law, for every y, if x and y are one and the very same, then y also has this complex property. Thus, if $x = y$, then the fact that $x = y$ is not grounded in any qualitative nonidentity facts about x (which are also facts about y) other than x's existence. Indeed, since x and y are one and the very same, the fact that $x = y$ is just the fact that $x = x$. Consequently, the fact that $x = y$ must have the property of the fact that $x = x$ that it is not grounded in any qualitative nonidentity facts about x (which are also facts about y) other than x's existence—*QED*. What a trivial and yet wonderful thing is Leibniz's Law!

The original proof that identity is determinate and fully defined for every pair of objects incidentally yields a persuasive reason for believing that the threshold for the amount of different original matter possible in the construction of an artifact consists in a sharp cutoff point rather than in a range of indeterminacy. A simple thought experiment shows that the threshold for the amount of different matter possible in the *reconstruction* of an artifact at some time after its disassembly does indeed consist in a sharp cutoff point. Recall that the number of molecules in the original matter of table a is n. Suppose that at time t_1, n distinct tables, $a_1, a_2, \ldots, a_n$, each qualitatively identical to a, are constructed from exactly n molecules apiece. At a later time t_2, each table is completely dismantled. At a still later time t_3, n tables, $a_1', a_2', \ldots, a_n'$, are constructed according to the same plan in the following way: a_1' is formed from all of the original molecules of a_1 except for the replacement of one molecule by a qualitative duplicate; a_2' is formed from all of the original molecules of a_2 except for the replacement of two molecules by qualitative duplicates of each; and so on, up to a_n', which is formed from entirely new matter. Clearly $a_1 = a_1'$ whereas $a_n \neq a_n'$. In fact, the construction of the second sequence of tables may be such that, for any i, if $a_i = a_i'$, then $a_{i-1} = a_{i-1}'$, and if $a_i \neq a_i'$, then $a_{i+1} \neq a_{i+1}'$. By the proof of the determinacy of identity, for any i, it is either determinately true that $a_i = a_i'$, or else it is determinately true that $a_i \neq a_i'$. Therefore, there must be some precise amount of different matter that first passes the threshold for the amount of different matter possible in the reconstruction of table a, i.e., there must be some m such that $a_m \neq a_m'$ but $a_{m-1} = a_{m-1}'$. On certain natural assumptions, this yields an excellent reason for supposing that in the sequence of hunks of matter $h_1, h_2, \ldots, h_n$, there is a hunk h_m that is the first to pass the threshold for the amount of different original matter possible in the construction of table a, i.e., that this threshold also consists in a precise cutoff point.

It was noted in section X above that it is *a posteriori*, even though it is necessary, that table a is not originally formed from hunk of matter h_m. Hence, it cannot be *a priori* that a is *necessarily* not originally formed from h_m. Is it then *a posteriori*? It is difficult to imagine establishing, by philosophical argument or otherwise, exactly what number m is, i.e., precisely how many molecules of difference from the actual original matter of table a would first result in a new and different table. It seems likely that it is *unknowable* that table a is necessarily not originally formed from hunk of matter h_m. That is, although it is knowable *a posteriori* that a is not in fact originally formed from h_m, and it is knowable (perhaps even knowable *a priori*) that there is some number m such that a difference of original matter of fewer than m molecules would still result in the same table though a difference of m or greater would result in a different table, it seems unlikely that one could know (*a priori* or *a posteriori*) *of* the relevant number *m*, whatever it is, that it is the threshold number of molecules of difference for the potential construction of table *a*. Whatever number *m* is, the fact that *a* is necessarily not originally formed from precisely that many different molecules from *a*' s actual original matter would appear to be a fact that is neither knowable *a priori* nor knowable *a posteriori*, since it appears not to be knowable at all.[38]

Notes

1. I have discussed some of these paradoxes in ''How *Not* to Derive Essentialism from the Theory of Reference,'' *The Journal of Philosophy* 76 (December 1979): 703–25, at 722–25; ''Fregean Theory and the Four Worlds Paradox: A Reply to David Over,'' *Philosophical Books* 25, no. 1 (January 1984): 7–11, at 9–11; ''Impossible Worlds,'' *Analysis* 44, no. 3 (June 1984): 114–17; and in more detail in *Reference and Essence* (Princeton, N.J., 1981), appendix I, 219–52.

A version of one of the paradoxes was apparently first noted by Saul Kripke in *Naming and Necessity* (Cambridge, Mass., 1980, originally published 1972), at 51, n. 18, where it is briefly discussed. Something directly akin to this paradox was also noted and discussed by Roderick Chisholm in ''Parts as Essential to their Wholes,'' *Review of Metaphysics* 26 (1973): 581–603, at 584–86, and again in *Person and Object* (La Salle, 1976), appendix B, 148–49. This paradox is highly reminiscent of Chisholm's paradoxical queries concerning cross-world identity in his seminal ''Identity Through Possible Worlds: Some Questions,'' *Noûs* 1 (March 1967): 1–8. I follow Graeme Forbes in calling this paradox 'Chisholm's Paradox', though I am uncertain as to the propriety of the epithet. See also Quine, ''Worlds Away,'' *The Journal of Philosophy* 73 (1976): 859–63, at 861; and N. L. Wilson, ''Substances Without Substrata,'' *Review of Metaphysics* 12 (June 1959): 521–39. The general solution to Chisholm's Paradox that I advocate in *Reference and Essence* and defend here was first proposed by Hugh Chandler in ''Plantinga and the Contingently Possible,'' *Analysis* 36 (January 1976): 106–9.

The modal paradoxes have been discussed by other writers, most notably Forbes in ''Thisness and Vagueness,'' *Synthese* 54, no. 2 (February 1983): 235–59; ''Two Solutions to Chisholm's Paradox,'' *Philosophical Studies* 46 (1984): 171–87; and *The Metaphysics of Modality* (Oxford, 1985), 160–90 and *passim;* and Anil Gupta in *The Logic of Common Nouns* (New Haven, 1980), 94–107. See also William R. Carter, ''Salmon on Artifact Origins and Lost Possibilities,'' *The Philosophical Review* 92, no. 2 (April 1983): 223–31; Forbes, ''Canonical Counterpart Theory,'' *Analysis* 42 (1982): 33–37, and ''On the Philosophical Basis of Essentialist Theories,'' *The Journal of Philosophical Logic* 10 (February 1981): 73–99; David Over, ''The Consequences of Direct Reference,'' *Philosophical Books* 25, no. 1 (January 1984): 1–7, at 4–7; and Robert Stalnaker, ''Counterparts and Identity,'' this volume.

2. The argument here is derived from one given by Kripke in *Naming and Necessity*, at 114, n. 56. See also p. 1. The argument is analyzed in *Reference and Essence*, chapter 7, 196–216. For similar arguments, see Forbes, ''Origin and Identity,'' *Philosophical Studies* 37 (1980): 353–62, and ''On the Philosophical Basis of Essentialist Theories''; and Colin McGinn, ''On the Necessity of Origin,'' *The Journal of Philosophy* 73 (March 11, 1976): 127–35, at 132. See also Richard Sharvy, ''Why a Class Can't Change its Members,'' *Noûs* 2 (1968): 303–14, and ''Individuation, Essence and Plenitude,'' *Philosophical Studies* 44 (1983): 61–70. (I should mention that I am here merely citing these works for further reference, and not endorsing the arguments or theses put forward in them.)

3. In fact, I believe that the threshold may indeed consist in a sharp cutoff point, though my approach to the modal paradoxes does not depend on this. An argument for a sharp cutoff is presented in the appendix below.

4. See ''How *Not* to Derive Essentialism from the Theory of Reference,'' 723–25, n. 22, and *Reference and Essence*, 251–52, n. 31, for further details concerning the argument of the paradox.

5. If the indeterminate accessibility account sketched here is correct, principle (I) must be regarded as untrue. For suppose that *z* is a hunk of matter for which it is vague or indeterminate whether a particular actual table *a* might have been originally formed from it. Then *a* is not formed from *z* in any world determinately accessible to the actual world, though it is so formed according to some plan *P* in some world w_k

neither determinately accessible or determinately inaccessible to the actual world. Now it is determinately possible for a table—some table or other—to be the only table originally formed from hunk z according to plan P. Hence there is a determinately accessible world w in which some table x is formed from hunk z according to plan P. Since w is determinately accessible to the actual world, and a is not formed from z in any determinately accessible world, it follows that tables x and a must be distinct. If (I) were true, it would follow by *modus ponens* from the existence of w that in every world not determinately inaccessible to the actual world (in every world either determinately accessible or neither determinately accessible nor determinately inaccessible), no table distinct from x is the only table formed from hunk z according to plan P. (See note 14 and the appendix, below.) Yet w_k is precisely such a world in which a table distinct from x, viz, a, is the only table formed from hunk z according to plan P. Hence if there is such a world as w_k, then (I) is not true—though (depending on the deails of the three-valued logic) it need not be false, since there need not be any *determinately accessible* world in which a table distinct from x is the only table formed from hunk z according to plan P. (But see note 3 above.)

Even if (I) is untrue for these reasons, it can be maintained that the following weakened version of (I) is necessarily true:

> (I′) If a table x is the only table originally formed from a hunk of matter z according to a certain plan P, then there could not be a table that is distinct from x and the only table originally formed from hunk z according to plan P.

The necessitation of (I′) is equivalent in *S4* to the necessitation of (I), though in the independent modal propositional logic B, the necessitation of (I′) is not sufficient for the derivation of principle (0). Of course, (0) may be true nevertheless.

6. See the works cited in note 1. Differences in terminology and emphasis, as well as certain theoretical differences, tend to obscure the overall fundamental similarity among the theories advocated or suggested by these writers. In place of Lewis's terminology of 'counterparts', Chisholm employs an alleged distinction, due to Joseph Butler, between "identity in the strict and philosophic sense" and "identity in the loose and popular sense," where artifacts made in different possible worlds from different constituent molecules are according to Chisholm never numerically one and the very same ("in the strict and philosophic sense"), though they may be said to be "the same" in the alleged loose and popular sense. The major difference between Chisholm and the other counterpart theorists is that Chisholm does not propose to replace the standard possible-world semantic analysis of formulations in modal-operator discourse of principles like (II) by an interpretation in terms of his counterpart relation, "identity in the loose and popular sense." Thus Chisholm dissents from the formulation of (II), whereas the other counterpart theorists assent to it. Perhaps this difference is enough to disqualify Chisholm as a genuine counterpart theorist, properly so-called, but I maintain that this difference is merely verbal and masks a basic agreement as to the facts (and that in this respect Chisholm is more perspicuous than the others).

Gupta occasionally uses the term 'counterpart' (at 105), but generally prefers to speak, somewhat misleadingly, of 'transworld identity relative to a world'. On Gupta's scheme for handling the modal paradoxes, an artifact x from one world and an artifact y from another world may be said to be "identical relative to" one world w and yet not "identical relative to" another world w'. Since x is "identical" with x (itself) relative to w' (assuming x exists in w'), however, it trivially follows by Leibniz's Law, or the Indiscernibility of Identicals, that x and y are not genuinely identical—they are not one and the very same—but are two distinct artifacts. (The sort of argument just given is discussed in the appendix below.) At most, then, x and y are merely counterparts at w, and it is at best misleading to call them *identical relative to* w. (Gupta's terminology is even more misleading than this, since he gives the title 'absolute identity' to the relation that obtains between a pair of objects when there is a world at which they are counterparts. This prompts him to make the astonishing claim that "absolute identity" is not transitive.)

I received a copy of Robert Stalnaker's "Counterparts and Identity" after the typescript of the present essay was submitted. Although I have not had the opportunity to study Stalnaker's essay carefully, several aspects of his theory seem similar to Gupta's. Stalnaker rejects my argument that his ternary, world-relative notion of "identity" is not genuine identity, but mere counterparthood, basing this rejection on the contention that there is no absolute (non-world-relative), binary notion of identity. This contention coupled with the rest of his theory, however, involves a number of serious difficulties, which I can only outline here.

First, Stalnaker claims (as part of the argument that there is no absolute notion of identity) that absolute truth is truth in the actual world. This gives us a notion of absolute identity: possible individuals are *absolutely identical* if and only if they are identical relative to the actual world. This notion coincides exactly with my intended notion of absolute identity if we assume that every possible individual x is such that, actually, $x = x$. Otherwise, it is better to substitute 'some world' for 'the actual world'. One way or another, absolute identity is definable in terms of world-relative identity, and Stalnaker's theory would thus allow for my notion of absolute identity if only it admitted the notion of intraworld identity. Unfortunately, the notion that Stalnaker calls 'intraworld identity' is not identity at all. This is made clear by Stalnaker's claim (which is essential to the point of his theory) that a single individual a in the actual world (e.g., Theseus's ship) can be two distinct individuals b and c in another possible world w. He defends this claim against the charge of violating the transitivity of identity, in part, by claiming that even though in w, $b \neq c$, in the actual world, $b = c$. But in w, $c = c$. This is inconsistent with (intraworld) Leibniz's Law (which Stalnaker claims to accept), since b does not (actually) have c's (actual) property of *being identical with c in w*. If b and c (actually) differ in this respect, then whatever else they are, they are not (actually) one and the very same object—since one object cannot differ from itself in any repect. In what sense are b and c (actually) "identical," then, except in the highly misleading sense of (actually) being *distinct* counterparts of a? Given that the intraworld relation that Stalnaker calls 'identity' (actually) holds between discernible objects b and c, this relation is not, in fact, genuine intraworld identity but is merely a counterpart relation. Stalnaker purports to explain this relation as genuine identity by saying that it is the binary relation whose extension, in any possible world w, is the set of pairs $<d, d>$ such that d is in the domain of w. The phrase 'the set of pairs $<d, d>$ such that', understood in its standard set-theoretic sense, ultimately involves the notion of identity. Indeed, in fixing the extension (with respect to any possible world) of a genuine identity predicate, the notion of identity is typically invoked in order to exclude pairs of distinct objects from the extension. If Stalnaker's purported explanation thus invokes genuine identity (as it seems to), it is inconsistent (via Leibniz's Law) with his claim that the discernible objects b and c (actually) stand in the relation. If, on the other hand, the purported explanation invokes, instead, Stalnaker's world-relative notion of what he calls 'identity' relative to w, the purported explanation is highly misleading. Moreover, it is circular and does not actually fix the metaphysical intension of the relation in question.

7. See especially "Thisness and Vagueness"; "Two Solutions to Chisholm's Paradox"; and *The Metaphysics of Modality*, chapters 3 and 7, the appendix, and *passim*.

8. A term is a *persistent* (or *persistently rigid*) *designator* if and only if it designates the same thing with respect to every possible world in which that thing exists and designates nothing or something else with respect to all other worlds. A term is an *obstinate* (or *obstinately rigid*) *designator* if and only if it designates the same thing with respect to every possible world, whether that thing exists there or not. For more on this distinction between two types of rigid designators, see *Reference and Essence*, 32–41.

9. A notion very similar to that of a counterpart assignment was apparently first introduced by Allen Hazen in his Ph.D. dissertation, "The Foundations of Modal Logic" (Pittsburgh, 1977). See his "Counterpart Theoretic Semantics for Modal Logic," *The Journal of Philosophy* 76, no. 6 (June 1979): 319–38, at 333–34, where analogues of counterpart assignments (there called 'representative functions') and world-assignment pairs ('stipulational worlds') are put to a use very similar to (though not exactly the same as) their use here.

10. The truth theoretic analysis that I am formulating here by means of counterpart assignments yields some significant differences in truth value assignments to particular modal-operator-discourse sentences from Lewis's own scheme. Specifically, the following clauses for the modal operators accord better with Lewis's actual scheme:

$\ulcorner \Box\phi \urcorner$ is $true_{w,c,s}$ iff ϕ is $true_{w',c' \circ c,s'}$ for every world w' and every counterpart assignment c' for w';
$\ulcorner \Diamond\phi \urcorner$ is $true_{w,c,s}$ iff ϕ is $true_{w',c' \circ c,s'}$ for some world w' and some counterpart assignment c' for w';

where $c' \circ c$ is the composite of the assignments c' and c, i.e., the function that assigns to any possible individual x, $c'(c(x))$.

Following Forbes, I am devising counterpart-theoretic possible-world semantics in such a way that '$\Diamond\Diamond F(a)$' is true exactly on the condition that a has an F counterpart at some world, so that '$\Diamond\Diamond F(a)$' is equivalent to '$\Diamond F(a)$', thus preserving *S4* modal logic. Lewis's original scheme has '$\Diamond\Diamond F(a)$' true exactly on the condition that a has a counterpart that itself has an F counterpart at some world. Since counterparthood is not transitive, this condition may be fulfilled though a itself has no F counterpart at any world. Lewis's scheme thus fails to preserve *S4*, since '$\Diamond\Diamond F(a)$' is weaker than '$\Diamond F(a)$'. This separates Lewis motivationally from theorists such as Forbes, who invoke counterpart theory precisely to retain *S4* modal logic in the face of the modal paradoxes. This does not mean, however, that Lewis himself blocks the Four Worlds Paradox in the same way as the accessibility solution. See note 20.

11. On Lewis's original scheme, (P_m) comes out true. See note 10 above.

12. Specifically, counterpart-theoretic possible-world semantics, as I have devised it, invalidates the inference

$$\frac{\Box(x)[G(x) \supset \Diamond F(x)]}{\therefore \Box(Exists(a) \supset [G(a) \supset \Diamond F(a)]),}$$

since it may be that every possible individual that is G in its own world has an F counterpart at some world, and that a has an existing G counterpart at some world, though a itself (as opposed to its G counterpart) has no F counterpart at any world. The trouble with this instance of (MUI) arises from the nesting of modalities in the conclusion. Lewis's original scheme validates this instance of (MUI), but as noted in note 10 above, it does not preserve *S4* modal logic. In a sense, then, the counterpart theorist is faced with a choice between *S4* and such instances of (MUI). Forbes chooses the former, Lewis the latter. Standard quantified *S5* modal logic validates both *S4* and (MUI). (Neither version of counterpart theory validates all instances of (MUI).)

Of course, counterpart-theoretic possible-world semantics can be artificially made to capture as much standard quantified *S5* modal logic as desired by placing further constraints on the counterpart relation. Standard modal logic emerges as the special case of counterpart-theoretic modal logic where counterparthood is identity. It is the philosophical motivation for counterpart theory, and the consequent explication of counterparthood in terms of sufficient cross-world similarity in certain respects, that requires the nontransitive and one-many nature of counterparthood.

13. Here and throughout this essay I am ignoring the possibility that semantic terms like 'true' and 'false' might themselves be vague or have partially defined semantic characteristic functions. If 'true' and 'false' are themelves vague or otherwise partially defined, a simple atomic sentence may suffer from second-order vagueness or second-order failure of truth value, in that the sentence may be, say, determinately untrue though it is indeterminate (vague, neither meta-true nor meta-false, there is no objective fact of the matter) whether the sentence is false or not. Similarly, a sentence that is determinately not false may be neither determinately true nor determinately untrue. The possibility of higher-order vagueness does not directly affect the main points I wish to make concerning the modal paradoxes and sorites paradoxes, and further discussion of this phenomenon in the present essay would introduce unnecessary complications. Notice that it is still reasonable to count a classical propositional sorites argument unsound, even if the sequence of sentences making up the antecedents and consequents of the conditional premises (e.g., sentences of the form ⌜Anyone of height h_i is short⌝) run the full gamut from determinately true to determinately not false but neither determinately true nor determinately untrue, to determinately neither true nor false, to determinately untrue but neither determinately false nor determinately not false, and finally to determinately false. Some of the conditional premises would have to be counted determinately not false while neither detrerminately true nor determinately untrue, but still others should be counted determinately not true. For it would be most reasonable to count a conditional determinately neither true nor false, and hence determinately untrue, whenever its antecedent is determinately not false but its consquent is determinately neither true nor false, and similarly whenever its antecedent is determinately neither true nor false and its consquent is determinately untrue.

14. Here I assume the three-valued modal semantics I put forward in "Impossible Worlds," 114, n. 2, rather than that of *Reference and Essence*, 248, n. 27.

15. The proof just presented that identity is nonvague is discussed further in the appendix.

16. Cf. *Reference and Essence*, 232–38.

17. I am concerned here with theories in the ordinary sense of the word. A theory in this sense is not merely a set of expressions closed under a special syntactic relation, but something more along the lines of a set of fully *interpreted* sentences, or a set of propositions, closed under genuine logical consequence.

18. "Thisness and Vagueness," 252. See also 258, n. 27.

19. Cf. Chisholm's doctrine of *ens successiva* in *Person and Object*, 89–113, 145–58.

20. One might also employ a conception, due to David Lewis, of possible worlds as *way things might have been*. My reason for not including this in the list is that Lewis himself (usually) takes a way-things-might-have-been-but-are-not to be something like an immense concrete object someplace "far, far away," in another dimension of the total cosmos, rather than a way the cosmos might have been, i.e., a possible state of the cosmos. If Lewis insists on this conception of a possible world, strictly speaking his version of counterpart theory is not a brand of essentialism at all, nor is it even relevant to modality in general. It is a fantastic cosmological theory.

21. Cf. *Reference and Essence*, 234–35. Forbes has responded to this objection by claiming, in effect, that the alleged validity of the inference in question does not violate any relevant logical intuition. See "Two Solutions to Chisholm's Paradox," 182; and *The Metaphysics of Modality*, 180. This might be taken as an indication that Forbes does not mean what the rest of us mean by 'Bill might have been a robot'. Most of us understand this sentence in such a way that it is true if and only if there is a scenario that might have obtained—or a history the world might have had, or a state of affairs that might have obtained, or a state the cosmos might have been in—in which Bill himself is a robot. Contrary to Forbes, we have a strong logical intuition that the proposition that Bill himself might have been a robot, whether true or false, is no logical consequence of any proposition to the effect that there might have been a robot counterpart of Bill—unless counterparthood is just identity.

Forbes's reply is based partly on his contention (disputed here) that possible-world-discourse sentences are, in effect, essentially idioms whose meanings are stipulated and fixed by the theorist's proposed translations in modal-operator discourse and not generated from the meanings of their grammatical components.

22. "A Problem for Actualism About Possible Worlds," *The Philosophical Review* 92, no. 1 (January 1983): 49–66.

23. "The Logic of Inexact Concepts," *Synthese* 19, no. 3/4 (April 1969): 325–73.

24. "Two Solutions to Chisholm's Paradox," 177. On the other hand, Forbes is willing to identify his notions of maximal truth and maximal falsehood with the traditional truth values of classical two-valued logic (*The Metaphysics of Modality*, 170).

The possibility of vagueness of infinite order (see footnote 13 above) suggests an alternative interpretation of the degrees-of-truth semantics for vagueness. On my suggested construal of the degrees-of-truth approach, second-order vagueness can be accommodated by allowing that a sentence may be determinately greater than 0 (or determinately less than 1) in truth value status while it is indeterminate whether the sentence takes on the value 1 (or 0) rather than some real between 0 and 1. As with first-order vagueness, the degrees-of-truth approach with indeterminacy allows for finer distinctions than the simple three-valued approach with indeterminacy. Still, on this construal, the latter approach is completely embedded within the former.

25. See for example W. R. Carter, "Salmon on Artifact Origins and Lost Possibilities," 228–29; Forbes, "Two Solutions to Chisholm's Paradox," 182–84, and *The Metaphysics of Modality*, 165n.

Forbes in particular has objected that the acceptance of this conclusion is incompatible with a general metaphysical principle concerning identity facts and the concept of identity—a principle that entails an extreme, cross-world version of the Identity of Indiscernibles (from which the identity of indiscernible worlds is derivable), and that, according to Forbes, provides the ultimate justification for the essentialist principle (III). This is the reductionist or supervenience principle that all facts about the numerical identity or distinctness of a pair of objects, *x* and *y*—including facts of cross-time and cross-world identity and distinctness—are metaphysically "grounded in," and "consist in," nonidentity facts about *x* and *y*, so that such identity facts do not obtain independently and solely by their own hook but only *in virtue of* nonidentity facts.

Of course, this formulation does not make clear the exact import of the intended principle. Forbes's intent can be gleaned to a certain extent by noting what he takes the principle to entail. An argument purporting to disprove the principle (whatever its precise import) is given in the appendix below.

26. This is not the only likely source of the temptation. Another possible source stems from the natural and plausible reductionist principle that a table is "nothing over and above" its matter, in the sense that a complete accounting of all of the matter in a genuinely possible world, with its exact configuration throughout time, must determine all of the remaining facts about the material objects, like tables, and everything else, present in the world. This immediately yields the supervenience thesis mentioned at the end of section I in connection with the fallaciously obtained false conclusion ($C2$). Within an *S4* framework, the reductionist principle renders the worlds w and w' exactly alike in all of the facts that obtain in each and in their accessibility relations. Cf. *Reference and Essence*, 237–38.

It should be noted that the conclusion ($C4$), which I claim to be true, is contradictory in *S4* modal logic.

27. "Two Solutions to Chisholm's Paradox," 185.

28. *The Metaphysics of Modality*, 237, n. 26. I have replaced Forbes's letter 'T' by 'p' for perspicuity.

29. The epistemological status of such propositions as these is discussed further in the appendix below.

30. The proof is elaborated and defended in *Reference and Essence*, 243–45. The general form of the argument was first given by Gareth Evans in "Can There Be Vague Objects?" *Analysis* 38 (1978): 208, although the argument had occurred to me independently. For further discussion, see John Broome, "Indefiniteness and Identity," *Analysis* 44, no. 1 (January 1984): 6–12; H. W. Noonan, "Vague Objects," *Analysis* 42, no. 1 (January 1982): 3–6; H. W. Noonan, "Indefinite Identity: A Reply to Broome," *Analysis* 44, no. 3 (June 1984): 117–21; David Over, "The Consequences of Direct Reference," 6; Richmond Thomason, "Identity and Vagueness," *Philosophical Studies* 42 (1982): 329–32; and my "Fregean Theory and the Four Worlds Paradox: A Reply to David Over," 10.

31. Cf. *Reference and Essence*, 244n.

32. *Naming and Necessity*, 108.

33. Cf. *Reference and Essence*, 244–45.

34. See Ruth Barcan, "The Identity of Individuals in a Strict Functional Calculus of Second Order," *The Journal of Symbolic Logic* 12 (1947): 12–15.

35. Church concludes from this argument and from the provability of the necessity of identity in quantified modal logic that there are compelling reasons to reject the meaningfulness of quantification into either modal or propositional attitude contexts. See his "A Remark Concerning Quine's Paradox About Modality," in *Propositions and Attitudes*, edited by N. Salmon and S. Soames, forthcoming. For a response, see my "Reflexivity," *The Notre Dame Journal of Formal Logic* 27 (July 1986): 401–29.

36. For a general defense of thesis T9, see my *Frege's Puzzle* (Cambridge, Mass., 1986).

37. See, for example, "Origin and Identity," "Thisness and Vagueness," and *The Metaphysics of Modality*, 126–31 and *passim*.

38. The present essay has benefited from the helpful comments and suggestions of Pascal Engel, Graeme Forbes, David Lewis, and John Pollock.

References

Adams, R. M. 1974. "Theories of Actuality." *Nous* 8: 211–31; also in Loux, 190–209.

Adams, R. M. 1979. "Primitive Thisness and Primitive Identity." *The Journal of Philosophy* 76, no. 1 (January): 5–26.

Adams, R. M. 1981. "Actualism and Thisness." *Synthese* 49: 3–41.

Barcan, R. 1947. "The Identity of Individuals in a Strict Functional Calculus of Second Order." *The Journal of Symbolic Logic* 12: 12–15.

Broome, J. 1984. "Indefiniteness and Identity." *Analysis* 44, no. 1 (January): 6–12.

Burke, M. B. 1983. "Essentialism and the Identity of Indiscernibles." *Philosophy Research Archives* 9: 223–43.

Carter, W. R. 1983. "Salmon on Artifact Origins and Lost Possibilities." *Philosophical Review* 92, No. 2 (April): 223–31.

Chandler, H. 1976. "Plantinga and the Contingently Possible." *Analysis* 36 (January): 106–9.

Chisholm, R. 1967. "Identity Through Possible Worlds: Some Questions." *Noûs* 1 (March): 1–8; also in Loux, 80–87.

Chisholm, R. 1973. "Parts as Essential to their Wholes." *Review of Metaphysics* 26: 581–603.

Chisholm, R. 1975. "Mereological Essentialism: Some Further Considerations." *Review of Metaphysics* 28: 477–84.

Chisholm, R. 1976. *Person and Object*. La Salle.
Church, A. Forthcoming. "A Remark Concerning Quine's Paradox About Modality." In *Propositions and Attitudes*, edited by N. Salmon and S. Soames.
Evans, G. 1978. "Can There Be Vague Objects?" *Analysis* 38: 208.
Forbes, G. 1980. "Origin and Identity." *Philosophical Studies* 37: 353–62.
Forbes, G. 1981. "On the Philosophical Basis of Essentialist Theories." *The Journal of Philosophical Logic* 10 (February): 73–99.
Forbes, G. 1982. "Canonical Counterpart Theory." *Analysis* 42: 33–37.
Forbes, G. 1983. "Thisness and Vagueness." *Synthese* 54, no. 2 (February): 235–59.
Forbes, G. 1984. "Two Solutions to Chisholm's Paradox." *Philosophical Studies* 46, no. 2 (September): 171–87.
Forbes, G. 1985. *The Metaphysics of Modality*. Oxford.
Fumerton, R. A. 1978. "Chandler on the Contingently Possible." *Analysis* 38: 39–40.
Goguen, A. 1969. "The Logic of Inexact Concepts." *Synthese* 19, no. 3/4 (April): 325–73.
Gupta, A. 1980. *The Logic of Common Nouns*. New Haven.
Hazen, A. 1977. "The Foundations of Modal Logic." Ph.D. dissertation. Pittsburgh.
Hazen, A. 1979. "Counterpart Theoretic Semantics for Modal Logic." *The Journal of Philosophy* 76, no. 6 (June): 319–38.
Kaplan, D. 1979. "Transworld Heir Lines." In Loux, 88--109.
Kaplan, D. 1975 "How to Russell a Frege-Church." *The Journal of Philosophy* 72 (November 6): 716–29; also in Loux, 210–24.
Kripke, S. 1971. "Identity and Necessity." In *Identity and Individuation*, edited by M. Munitz, 135–64. New York. Also in *Naming, Necessity, and Natural Kinds*, 1977, edited by S. Schwartz, 66–101. Ithaca, N.Y.
Kripke, S. 1972. *Naming and Necessity*. Cambridge, Mass., and Oxford.
Lewis, D. 1968. "Counterpart Theory and Quantified Modal Logic." *The Journal of Philosophy* 65: 113–26; also in Loux, 110–28; also in Lewis, 1983, with substantive postscripts, 26–46.
Lewis, D. 1973. *Counterfactuals*. Cambridge, Mass.
Lewis, D. 1983. *Philosophical Papers I*: Oxford.
Loux, M., ed. 1979. *The Possible and the Actual*. Ithaca, N.Y.
Marcus, R. B.: See Barcan.
McGinn, C. 1976. "On the Necessity of Origin." *The Journal of Philosophy* 73 (March 11): 127–35.
McMichael, A. 1983. "A Problem for Actualism About Possible Worlds." *The Philosophical Review* 92, no. 1 (January): 49–66.
Noonan, H. W. 1983. "The Necessity of Origin." *Mind* 92, no. 365 (January): 1–20.
Noonan, H. W. 1985. "Indefinite Identity: A Reply to Broome." *Analysis* 44, no. 3 (June): 117–21.
Odegard, D. 1976 "On A Priori Contingency." *Analysis* 36 (June): 201–3.
Over, D. 1984. "The Consequences of Direct Reference." *Philosophical Books* 25, no. 1 (January): 1–7.
Plantinga, A. 1975. "On Mereological Essentialism." *Review of Metaphysics* 28: 468–76.
Plantinga, A. 1974. *The Nature of Necessity*. Oxford.
Quine, W. V. O. 1976. "Worlds Away." *The Journal of Philosophy* 73: 859–63.
Salmon, N. 1979. "How *Not* to Derive Essentialism from the Theory of Reference." *The Journal of Philosophy* 76 (December): 703–25.
Salmon, N. 1981. *Reference and Essence*. Princeton and Oxford.
Salmon, N. 1984. "Fregean Theory and the Four Worlds Paradox: A Reply to David Over." *Philosophical Books* 25, no. 1 (January): 7–11.
Salmon, N. 1984. "Impossible Worlds." *Analysis* 44, no. 3 (June): 114-17.
Salmon, N. 1986. *Frege's Puzzle*. Cambridge, Mass.
Salmon, N. 1986. "Reflexivity." *The Notre Dame Journal of Formal Logic* 27, no. 3 (July): 401–29.
Sharvy, R. 1968. "Why a Class Can't Change its Members." *Noûs* 2 (November): 303–14.
Sharvy, R. 1983. "Individuation, Essence and Plenitude." *Philosophical Studies* 44: 61-70.
Stalnaker, R. 1976 "Possible Worlds." *Noûs* 10: 65–75; also in Loux, 225–34.
Stalnaker, R. 1986. "Counterparts and Identity." In this volume.
Thomason, R. 1982. "Identity and Vagueness." *Philosophical Studies* 42: 329–32.
Wilson, N. L. 1959. "Substances Without Substrata." *Review of Metaphysics* 12 (June): 521–39.

Princeton University

Department of Philosophy
1879 Hall
Princeton, New Jersey 08544 U.S.A.

February 3, 1987

Professor Nathan Salmon
Department of Philosophy
University of California, Santa Barbara
Santa Barbara, CA 93106

Dear Nathan,

Enclosed are two papers, including the one in the Linsky volume too. The treatment of the relation _R_ is on pp. 69-70 (section 2.1) of "Semantical Analysis" and on p. 64 of "Semantical Considerations". The same characterization of _R_ is given in both places. As you see, there is no suggestion that S5 is basic and the weaker systems come from some restricted conception. _R_ is characterized in terms of truth and possibility of propositions in worlds. Notice also the discussion of the reduction axioms on p. 70 of "Semantical Analysis", and in particular of transitivity and S4.

One thing I do is, I now think, somewhat misleading. I should have stressed that the use of _R_ does not make "possible" (as applied to worlds) into a two-place predicate, any more than, as you say, "is bald" is. Probably I only noticed this afterwards. Also, I should have stressed that strictly speaking, many of the worlds are not "possible" but only "possibly possible", and so on, unless we have S4.

By the time I gave the seminar I talked to you about I had definitely thought these points through, having seriously considered whether the conventional presupposition that the basic modal logic is S5 is justified.

I am getting closer to thinking that your treatment of the ship is the correct solution. Certainly it is a very good piece of work. I am sorry if almost everyone is unable to see its virtues (you don't say quite that in the paper). As far as I can see, their counterarguments, as presented, are confused or circular. It was good talking to you. Talk to you about Russell, etc., some time.

Best,

Saul

Saul Kripke

Enc.