Counterparts and Actuality

Michael Fara
Cornell University
michael.fara@cornell.edu

Timothy Williamson
Oxford University
timothy.williamson@philosophy.oxford.ac.uk

To appear in *Mind* **114**, 2005.

1. Introduction

David Lewis has argued that when it comes to interpreting claims about what might or must have been, we should use the resources of his counterpart theory, not those of quantified modal logic (Lewis 1968, Lewis 1986). Lewis has offered a host of reasons for this. The language of counterpart theory is purely extensional, and questions about logical relations between modal claims are reduced, via counterpart theory, to easily solvable questions about classical first-order logical consequence; and the language of counterpart theory has greater expressive power than the language of quantified modal logic—whatever can be expressed in the latter can be expressed in the former, but not vice versa. These two alleged benefits are not unique to counterpart theory, since they will be shared by any view of modality which takes as basic a non-modal first-order language, whether it mentions counterparts or not. But Lewis's list of reasons for insisting on counterpart theory continues: only with the machinery of counterpart theory can we adequately explain the "inconstancy of representation *de re*," the wavering back and forth of our intuitive judgments about essentialist claims; by appealing to counterpart theory we can solve problems of material constitution, without being committed either to the existence of distinct coincident objects or to the absurd view that identity is a relation that may hold only contingently between an object and itself; and counterpart theory permits us to endorse such haecceitistic-sounding claims as "I might have been Frank Sinatra while everything else is as it actually is," without being committed to haecceitism proper, the view that there are distinct but qualitatively identical possible worlds.

This is no small list of counterpart theory's virtues, and many philosophers have been persuaded for these reasons that counterpart theory is preferable to quantified modal logic. We will argue, however, that even if the supposed virtues are genuine,[1] they are not worth their price: there is very good reason to reject

[1] Whether the supposed virtues are to be regarded as virtues at all depends, of course, on one's views about haecceitism, material constitution, and so on. We take no stand on these issues here.

counterpart theory as a means of interpreting modal claims. Those attracted to *modalism*, the view that modal operators are linguistically or conceptually primitive, should welcome our arguments. But we hope equally to convince the anti-modalist that, however modality is to be understood, it is not to be understood in terms of counterparts.

Counterpart theory, as originally developed by Lewis (1968), is the theory which results from adding to classical first-order logic with identity eight postulates governing the interaction of four distinguished predicates, and taking the domain of quantification to be the class of all possible individuals (among them, the possible worlds). One of these four predicates is intended to hold of just the possible worlds; we will depart from Lewis in abandoning this predicate in favor of sorted variables: v, w are to range over possible worlds; x, y, z, x_1, x_2, $\ldots$ over other possible individuals. Nothing of substance will turn on this simplification. The remaining three primitive predicates of counterpart theory are these:

Ixw (x is in possible world w)
Ax (x is actual)
Cxy (x is a counterpart of y)

The first four postulates of counterpart theory ensure that if an individual is in anything at all, then it is in exactly one thing, a possible world; and whatever either is or has a counterpart is in something. Two more ensure that everything has exactly one counterpart at the world it is in: itself. The last two postulates ensure that there is a unique, non-empty possible world in which are all and only the actual individuals—this world is the actual world, @.[2]

What is a counterpart? According to Lewis,

> The counterpart relation is our substitute for identity between different things in different worlds. Where some would say that you are in several worlds, in which you have somewhat different properties and somewhat different things happen to you, I prefer to say that you are in the actual world and no other, but you have counterparts in several other worlds. Your counterparts resemble you closely in content and context in important respects. They resemble you more closely than do the other things in their worlds. (Lewis 1968, pp. 27–8)

[2] Formally, '@' abbreviates the description '$\iota x \forall y (Iyx \equiv Ay)$'.

Given this understanding of the counterpart relation, Lewis maintains, counterpart theory is superior to quantified modal logic as a means of understanding modal discourse in natural language, something that is achieved with the help of a systematic method of translating sentences of English into sentences of the language of counterpart theory.

The translation process takes two steps. First, translate an English sentence into the language of quantified modal logic; we follow Lewis in taking this step to be familiar. Then translate from there to the language of counterpart theory by means of the following translation scheme:

<u>Lewis's Translation Scheme</u>

• A sentence φ of the language of quantified modal logic is translated to the sentence $\varphi^{@}$ of the language of counterpart theory, to be read "φ holds at the actual world".[3]

• For any sentence φ of the language of quantified modal logic, and any world-term v, the sentence φ^v of the language of counterpart theory is defined recursively as follows:

(LA) φ^v is φ, if φ is atomic.

(L¬) $(\neg\varphi)^v$ is $\neg\varphi^v$.

(L&) $(\varphi \;\&\; \psi)^v$ is $\varphi^v \;\&\; \psi^v$.

(L∀) $(\forall x\varphi)^v$ is $\forall x(Ixv \supset \varphi^v)$.

(L∃) $(\exists x\varphi)^v$ is $\exists x(Ixv \;\&\; \varphi^v)$.

(L□) $(\Box\varphi)^v$, where the unbound terms in φ are $a_1, \ldots a_n$, is

$\forall w \forall x_1 \ldots \forall x_n((Ix_1w \;\&\; \ldots \;\&\; Ix_nw \;\&\; Cx_1a_1 \;\&\; \ldots \;\&\; Cx_na_n) \supset$

$\varphi^w(x_1, \ldots, x_n))$.

(L◇) $(\Diamond\varphi)^v$, where the unbound terms in φ are $a_1, \ldots a_n$, is

$\exists w \exists x_1 \ldots \exists x_n(Ix_1w \;\&\; \ldots \;\&\; Ix_nw \;\&\; Cx_1a_1 \;\&\; \ldots \;\&\; Cx_na_n \;\&\; \varphi^w(x_1, \ldots, x_n))$.

For example, according to Lewis's translation scheme the English sentence (1) gets translated into the counterpart theoretic sentence (3) via the sentence (2) of quantified modal logic:

[3] We are sloppy, throughout, about distinguishing use and mention of expressions, choosing (for the most part) to avoid writing quotation-marks of any kind.

(1) There are two things that might have been brothers.

(2) $\exists x \exists y(x \neq y \,\&\, \Diamond B(x, y))$.

(3) $\exists x(Ix@ \,\&\, \exists y(Iy@ \,\&\, [x \neq y \,\&\, \exists w \exists x_1 \exists x_2(Ix_1 w \,\&\, Ix_2 w \,\&\, Cx_1 x \,\&\, Cx_2 y \,\&\, B(x_1, x_2))]))$.

We began by listing some of counterpart theory's alleged virtues. One of these was its comparative expressive power: everything that can be said in the language of quantified modal logic can be said in the language of counterpart theory, but not vice versa. The expressive weakness of the language of quantified modal logic has long been known, however.[4] Consider, for example, the English sentence

(4) It might have been that everyone who is in fact rich was poor.

This sentence is not expressible in the language of first-order quantified modal logic in any way that reflects its semantic structure, in particular using quantification over a domain of persons and translations of 'rich' and 'poor' as the only nonlogical atomic predicates. The only hopes for expressing (4) in this way, without resorting to plural quantification, quantification over sets or the like, seem to be either (5) or (6):

(5) $\Diamond \forall x(Rx \supset Px)$.

(6) $\forall x(Rx \supset \Diamond Px)$.

Neither of these captures what is meant by the English sentence (4), however. Since no-one can be both rich and poor, (5) will be true just in case it might have been that no-one was rich, which is surely not what (4) intends. And (6) will be true as long as there is, for each rich person, a world in which she is poor; yet (4) is about the possibility of the actually rich people being poor *together*.

The standard response to this fact about the expressive weakness of modal languages has been to take it as motivating the enrichment of those languages with the addition of an actuality operator. Syntactically, this is a sentential operator ACT that, like $\Box$ and $\Diamond$, attaches to a sentence to yield a sentence. Semantically, the actuality operator can be characterized with the following simple clause:

- $ACT\varphi$ is true at a world w iff φ is true at the actual world.

(In a formal Kripke-style model theory, where models come with triples $< W, R, w^* >$, with W a non-empty set, R a relation on W and w^* a designated member of W, $ACT\varphi$ is true at a member w of W in

[4]See, for example, Hazen (1976), Crossley & Humberstone (1977), Davies (1981), Cresswell (1990).

$< W, R, w^* >$ iff φ is true at w^* in $< W, R, w^* >$.) Quantified modal logic, enriched with the operator ACT, now apparently has the resources to express (4), namely as

(7) $\Diamond\forall x(ACTRx \supset Px)$.

Interpreting $\Diamond$ as an existential quantifier over possible worlds, (7) seems to say that there is a world w such that everything that is rich in the actual world is poor in w, exactly what is intended by (4).

It is not often noted, however, that the question of whether (7) is the correct formalization of (4) is a delicate one. On a standard variable-domained Kripke semantics for quantified modal logic (Kripke (1963)), on which the modal operators are interpreted as quantifiers ranging over possible worlds and different objects may exist at different worlds, (7) can be true in virtue of the existence of a world none of whose inhabitants exists at the actual world. But the existence of such a world seems insufficient for the truth of (4), at least on one of its natural readings.

So much the worse, perhaps, for variable-domained Kripke semantics. For this problem will not arise in a semantics according to which the range of the quantifiers, whether inside or outside the scope of modal operators, is invariant.[5] Indeed, the problem will not arise in a semantics which validates the Converse Barcan formulas

(CBF) $\Box\forall x\varphi(x) \supset \forall x\Box\varphi(x)$,

or, in existential form,

(ECBF) $\exists x\Diamond\varphi(x) \supset \Diamond\exists x\varphi(x)$.

For the validity of (CBF) in Kripke semantics amounts to the requirement that for any worlds u and v in a model, if v is accessible from u then the domain of quantification at u is a subset of the domain of quantification at v.[6] Such a semantics might be motivated by possibilism, the view that the quantifiers range with respect to a world over every possible thing, even those things that don't exist at that world. Or it might be taken to motivate the belief that everything exists necessarily (where existing is simply being

[5]The problem can also be avoided while retaining variable-domained semantics by enriching the language of quantified modal logic, for example by adding the *actuality quantifiers* of Hazen (1990), or the *Vlach operators* described by Forbes (1989, p. 27).

[6]In fact, the problem being discussed will be avoided provided only that (CBF) is *weakly* validated by the semantics, true at the actual world of every model, where this amounts to the requirement that everything in the domain of the actual world is also in the domain of every world accessible from the actual world.

something)—for given that it is necessary, possibilism aside, that everything exists, an instance of (CBF) *requires* that everything exists necessarily (let $\varphi(x)$ be $\exists y(x = y)$.)

Many philosophers find it hard to accept the apparent consequence of possibilism that there are things that don't exist, and equally hard to accept that everything exists necessarily, and so choose to reject any semantics which validates the Converse Barcan formulas.[7] But it seems at least to be a point in favor of such semantics that they allow us to endorse (7) as a translation of (4). For (7) does seem to be just a symbolic representation of the clumsy English equivalent of (4):

(8) It is possible that everything actually rich is poor.

In any case, our concern here is not to argue for one semantics for quantified modal logic over another; we are instead motivating the enrichment of the language of quantified modal logic with an actuality operator. Note that those who don't want to endorse the Converse Barcan formulas typically are equally reluctant to accept the Barcan formulas:

(BF) $\forall x \Box \varphi(x) \supset \Box \forall x \varphi(x)$,

or, in existential form,

(EBF) $\Diamond \exists x \varphi(x) \supset \exists x \Diamond \varphi(x)$.

Indeed, the Barcan formulas and their converses seem to stand or fall together, since they are interderivable even in fairly weak systems of modal logic. Any normal modal logic with the B axiom

(B) $\varphi \supset \Box \Diamond \varphi$,

corresponding semantically to the symmetry of accessibility, yields the equivalence of (BF) and (CBF). The standard system for metaphysical modality, S5, is an example. But one must reject the Barcan schema in order to endorse, for example,

(9) There could be something that doesn't actually exist.

Yet this very sentence is inexpressible in a modal language without an actuality operator (Hodes 1984*b*).

[7]But see Williamson (1998), Williamson (2002) and Linsky & Zalta (1994) for arguments that such philosophers are mistaken. See also Prior (1956) for an argument showing that (CBF) is derivable in any quantified modal logic with the K axiom: $\Box(\varphi \supset \psi) \supset (\Box\varphi \supset \Box\psi)$. Kripke (1963) shows how to reformulate quantified modal logic in order to block Prior's argument.

We will assume, then, that the operator ACT is available in the language of quantified modal logic. Given this assumption, Lewis's translation scheme is incomplete. We need to add to his recursive definition of φ^v a clause telling us how to translate sentences beginning with the ACT operator. Problems for providing such a clause were first raised by Allen Hazen (1979), and have since been explored in depth by Graeme Forbes and Murali Ramachandran. Forbes and Ramachandran have responded by proposing modified versions of counterpart theory intended to solve these problems.[8] We will argue in this paper, however, by generalizing some of Hazen's observations and raising some hitherto unnoticed problems, that there can be no plausible, systematic translation of sentences involving the operator ACT into sentences of the language of counterpart theory.

2. The Problem of No Counterparts

How might counterpart theory be revised to accommodate sentences that make use of the operator ACT? What is needed, it seems, is an extra clause to supplement Lewis's translation scheme. A natural suggestion is this: since $\Diamond Fx$ is interpreted by counterpart theory as meaning that there is a world in which a counterpart of x is F, $ACTFx$ should be interpreted as meaning that the *actual* world is a world in which a counterpart of x is F. Thus, for formulas with occurrences of just one term, for example,[9] we might try:

(L1) $(ACT\varphi(a))^v$ is $\exists x(Ix@ \;\&\; Cxa \;\&\; \varphi^@(x))$.

"φ holds, at the actual world, of some actual-world counterpart of a."

Tempting as it may seem, however, the translation clause (L1) is hopeless. For on the assumption that there are possible objects with no counterparts in the actual world—an assumption that would be endorsed by Lewis[10]—(L1) translates inconsistent sentences of the language of quantified modal logic to sentences that are true in some models of counterpart theory.

Consider, for example,

[8]Hazen in effect responded by refusing to treat counterpart theory as an alternative to quantified modal logic. Instead, on Hazen's view, there is room for the notion of a counterpart in the *semantics* for quantified modal logic, where these semantics do not face the problem of transworld identity that allegedly plagues Kripke semantics. As Hazen proves, however, the semantics he gives are equivalent—in the sense of validating the same sentences—to a version of Kripke's semantics for quantified modal logic (the version given in Kripke (1963) together with the assumption, described by Kripke in footnote 11, that the extension of a predicate at a world contains only objects that exist at that world.) According to Hazen, then, we should reject counterpart *theory* as a means of understanding modal claims of natural language, although we might appeal to a relation of similarity between possible objects in giving a semantics for quantified modal logic.

[9]This is just for sake of legibility; the generalization to n-place formulas is obvious.

[10]See the discussion in Lewis (1968, pp. 28–9).

(10) $Fa \mathbin{\&} \neg ACTFa$.

(10) is inconsistent, in the sense that its denial is a theorem of a standard quantified modal logic with an actuality operator.[11] Yet (10) translates, via (L1), to

(11) $Fa \mathbin{\&} \neg\exists x(Ix@ \mathbin{\&} Cxa \mathbin{\&} Fx)$.

(11) is true in models of counterpart theory in which a is assigned an object, in the extension assigned to F, which has no counterparts in the actual world. Since (L1) therefore translates inconsistencies to truths, it is to be rejected.

Some philosophers may balk here at our talk of an assignment to a of an object that has no counterparts in the actual world, since, given that every object has itself as a counterpart, this requires assigning to the name a an object that does not exist in the actual world. When choosing the denotations of names, it may be insisted, we have only *actually existing* objects available to choose among.[12] So it might be thought that the sentence (11) is *not* true in any model of counterpart theory, since there are no models in which a is assigned a non-actual object as its denotation, and hence no models in which a is assigned an object with no actual-world counterparts.

This response on behalf of the counterpart theorist needs to be put carefully. A model of counterpart theory is simply an interpretation of classical first-order logic in which all the postulates of counterpart theory are true. And there *are* models of counterpart theory, in this sense, in which sentence (11) is true. The response, if it is to be pressed, should thus be put like this: We should conceive of counterpart theory as being strengthened by additional postulates which ensure that names are never assigned non-actual objects as their denotations. For example, for each name n of the language, we could add the postulate $In@$. And there are no models of counterpart theory in *this* sense in which sentence (11) is true.[13]

There is little point in pressing the response in this context, however. For the objection being given—that an inconsistent sentence translates via (L1) to a sentence which is true in some models of counterpart theory—can as easily be put using bound variables instead of names, as witness the satisfiability of (13), the

[11] For example, the logic axiomatized in Hodes (1984*a*). We use Hodes' axiomatization here only as a codification of standard principles about the logic of "actually"; in particular, we are not appealing to a non-counterpart-theoretic model theory. We are unimpressed by rejections of standard logic that are not accompanied by a properly worked out alternative.

[12] See, for example, Forbes (1990).

[13] In his original presentation of counterpart theory, Lewis required that names be replaced by definite descriptions before sentences containing them are translated into the language of counterpart theory. Since those descriptions will be given wide scope with respect to modal operators, however, this point has no bearing on the objection being considered here.

translation under (L1) of the inconsistent (12):

(12) $\Diamond\exists x(ACTFx \equiv ACT\neg Fx)$.

(13) $\exists w\exists x(Ixw \,\&\, [\exists y(Iy@ \,\&\, Cyx \,\&\, Fy) \equiv \exists y(Iy@ \,\&\, Cyx \,\&\, \neg Fy)])$.

(13) is true in models of counterpart theory, provided there are possible objects with no counterparts in the actual world.

Note that not only is the denial of (12) a theorem of standard modal logics with the actuality operator, but the necessitation of that denial is a theorem as well:

(14) $\Box\neg\Diamond\exists x(ACTFx \equiv ACT\neg Fx)$.

Unlike the previous sentence (10), then, (12) is inconsistent even in logics with the rule of necessitation, i. e. those whose theorems are deemed to hold necessarily, not just actually. Thus the version of the objection being made now with (12) is more robust with respect to the logic of actuality than the version made earlier with sentence (10). Similar remarks, although they are omitted, would be appropriate at various points below.

If the clause (L1) will not work as an extension to Lewis's translation scheme, then what might be offered in its place? Notice that (L1) interpreted ACT in the manner of $\Diamond$: For φ to be actually true of x, φ must hold true of *some* counterpart of x in the actual world. A natural alternative to (L1) might then be to interpret ACT in the manner of $\Box$, and require that φ hold true of *every* counterpart of x in the actual world if φ is to be actually true of x:

(L2) $(ACT\varphi(a))^v$ is $\forall x[(Ix@ \,\&\, Cxa) \supset \varphi^@(x)]$.
"φ holds, at the actual world, of every actual-world counterpart of a."

Yet (L2) fails for a reason parallel to that levelled against (L1): it translates inconsistent sentences of quantified modal logic into sentences that are true in some models of counterpart theory. Consider the inconsistent

(15) $Fa \,\&\, ACT\neg Fa$,

which, given (L2), is translated to

(16) $Fa \,\&\, \forall x[(Ix@ \,\&\, Cxa) \supset \neg Fx]$.

But (16) is true in models of counterpart theory in which a is assigned an object, in the extension of F, which has no actual-world counterparts. Since (L2) thus translates inconsistencies to truths, it must be rejected.

As before, for those who dislike assignments of non-actuals as the denotations of names, the point can be made instead with the inconsistent sentence we considered earlier,

(12) $\Diamond\exists x(ACTFx \equiv ACT\neg Fx)$,

which translates, this time, to the satisfiable

(17) $\exists w\exists x(Ixw \mathbin{\&} [(\forall y(Iy@ \mathbin{\&} Cyx) \supset Fy) \equiv (\forall y(Iy@ \mathbin{\&} Cyx) \supset \neg Fy)])$.

We have so far considered two ways of extending counterpart theory to accommodate sentences of quantified modal logic involving the operator ACT. The first, (L1), effectively treated ACT as an existential quantifier over actual-world counterparts; the second, (L2), treated it as a universal quantifier over actual-world counterparts. Might we not then try treating ACT as some quantifier over actual-world counterparts *other* than an existential or a universal? Perhaps, for example, φ should count as actually true of x if φ holds true, at the actual world, of *most*, or *at least two*, or *all but three* actual-world counterparts of x. We can put this proposal schematically:

(L3) $(ACT\varphi(a))^v$ is $[Qx : Ix@ \mathbin{\&} Cxa](\varphi^@(x))$.
"Q actual-world counterparts of a are actually φ".

We will assume that the candidate quantifiers to be substituted for Q in (L3) are the *logical* quantifiers, in the terminology of Westerståhl (1989). Intuitively, a quantifier Q is logical if, in saying that Q As are Bs, it is irrelevant what the non-As are like: for the sake of interpreting the quantified sentence, we may as well suppose that everything is an A. Moreover, all that matters is *how many* As are Bs and *how many* As are not Bs. In that sense, a logical quantifier expresses a purely structural higher-order relation, one invariant under permutations and bijections of individuals. Formally, we take a logical quantifier to be any that meets the following condition:[14]

> If f is a bijection of the satisfiers of $\varphi(x)$ in a model M onto the satisfiers of $\varphi^*(x)$ in a model M^* that maps those of the former that satisfy $\psi(x)$ in M onto those of the latter that satisfy $\psi^*(x)$ in M^*, then $[Qx : \varphi(x)](\psi(x))$ is true in M iff $[Qx : \varphi^*(x)](\psi^*(x))$ is true in M^*.

[14] Logicality is typically stated using set-theoretic apparatus, but this is inessential, as our statement demonstrates. Even the function variable f can be interpreted as second-order rather than set-theoretic.

One consequence of this characterization, which we make use of below, is that for any logical quantifier Q there is a relation R_Q such that, if the satisfiers of $\varphi(x)$ in a model M form a set, then $[Qx : \varphi(x)](\psi(x))$ is true in M iff the number of satisfiers of $\varphi(x)$ & $\psi(x)$ in M bears R_Q to the number of satisfiers of $\varphi(x)$ & $\neg\psi(x)$ in M. Intuitively, the point of requiring the quantifier Q in (L3) to be logical is that since it is introduced into the semantics to handle ACT, and ACT is purely logical in meaning, the quantifier should be equally logical. The modal logician thinks that whether a formula $ACT\varphi$ is true of a possible object depends on whether φ is true of that object at the actual world. The counterpart theorist should think that it depends, instead, on facts about the *counterparts* of that object at the actual world. But it shouldn't depend also on anything *other* than facts about the counterparts of that object at the actual world, and on how many of those counterparts φ is true of—which it would on a version of (L3) with a non-logical quantifier.

Unfortunately for the counterpart theorist, there is *no* logical quantifier Q for which the corresponding instance of (L3) yields a plausible extension of Lewis's translation scheme, on the assumption, once again, that there could be objects with no actual-world counterparts. Consider, for example, the sentence (18), which translates via (L3) to (19):

(18) $Fa \ \& \ (ACT\neg Fa \vee \neg ACTFa)$.

(19) $Fa \ \& \ ([Qx : Ix@ \ \& \ Cxa](\neg Fx) \vee \neg[Qx : Ix@ \ \& \ Cxa](Fx))$.

(18) is inconsistent, but its translation (19) is true, as before, in models of counterpart theory that assign to a an object which has no actual-world counterparts and to F an extension containing that object, as can be seen by noticing the following fact, which holds for any logical quantifier Q:

> For any formula $\varphi(x)$, if no object satisfies $\varphi(x)$, then $[Qx : \varphi(x)](\psi(x))$ is either *true* for all formulas $\psi(x)$ or *false* for all formulas $\psi(x)$.

(That this is a fact can be seen by noticing that, if no object satisfies $\varphi(x)$, then $[Qx : \varphi(x)](\psi(x))$ is true iff $0R_Q0$ for any $\psi(x)$, where R_Q is the relation, mentioned above, associated with the quantifier Q.) Since (L3) translates inconsistent sentences into sentences true in some models of counterpart theory, it should be rejected. And the point can be made, once again, by considering bound variables instead of names. This time, the inconsistent (12) translates under (L3) to the satisfiable (20):

(12) $\Diamond\exists x(ACTFx \equiv ACT\neg Fx)$.

(20) $\exists w\exists x(Ixw \;\&\; ([Qy : Iy@ \;\&\; Cyx](Fy) \equiv [Qy : Iy@ \;\&\; Cyx](\neg Fy)))$.

Given the fact about logical quantifiers recently mentioned, (20) is true in models of counterpart theory, provided there are possible objects with no counterparts in the actual world.

Our discussion so far has shown that, on the assumption that there could be objects without any actual-world counterparts, Lewis's translation scheme cannot plausibly be extended to accommodate the actuality operator. Some counterpart theorists may take this to motivate rejecting this assumption, supplementing counterpart theory with a requirement to the effect that every possible object have at least one counterpart in the actual world. It is worth noting, however, that adopting this requirement in conjunction with either of the clauses (L1) or (L2) will yield the validity of

(21) $\Box\forall xACT\exists y(x = y)$,

the claim that it is necessary that everything actually exists, a result with which the counterpart theorist may not be content.[15] (21) will also be validated under any instance of (L3) for which a sufficient condition for $[Qx : \varphi(x)](\psi(x))$ to be true is that something satisfies $\varphi(x)$ and that whatever satisfies $\varphi(x)$ satisfies $\psi(x)$.[16] It is hard to see how this could not be a sufficient condition; if a possible object has actual-world counterparts, all of which are F, how could it fail to be actually F on the counterpart-theoretic approach? We return to these issues in the last section of the paper. For now, though, we turn to consider different objections to the translation clauses we have considered, objections which arise if we assume, as Lewis does, that some possible objects have *multiple* actual-world counterparts.

3. The Problem of Multiple Counterparts

Recall one of the alleged benefits of counterpart theory listed at the outset: It allows us to solve problems of material constitution. Consider a world W in which a lump of clay is at the moment of its creation formed into the shape of a statue and remains in that shape until its eventual destruction.[17] The problem of material constitution raised here is the problem of defending an answer to the question, Is the lump identical with

[15] (21) is in fact valid under (L2) independently of any assumptions.

[16] If $[Qx : \varphi(x)](\psi(x))$ is false when nothing satisfies $\varphi(x)$ then, under (L3), (21) requires every possible object to have an actual-world counterpart, whereas (BF) in effect requires every possible object to be the counterpart of something actual. In this paper we do not discuss whether counterparthood is symmetric; see Lewis (1968, pp. 28–9) for arguments that it is not.

[17] This kind of case is considered by Gibbard (1975).

the statue? According to defenders of counterpart theory, the solution to this problem lies in recognizing that the following claim is true at W: "The lump is identical with the statue, but it might not have been identical with the statue". It is true at W that the lump is identical with the statue, since to claim that they are distinct would be to admit the possibility of distinct but forever spatially coincident objects, which is allegedly implausible. But the sentence "The lump might not have been identical with the statue" is also true at W, since it might have been that the lump, but not the statue, survived flattening.

Lewis's preferred way of invoking counterpart theory to make sense of this, without being committed to the absurdity that the lump of clay bears the relation of identity to itself only contingently, is to posit a multiplicity of counterpart relations. There is a lumplike-counterpart that survives flattening, and a statuelike-counterpart that does not, and this, according to Lewis, is how the lump and the statue might have been distinct even though they are, at W, identical. But even restricting herself to just a single counterpart relation, the counterpart theorist apparently has the resources to explain how the lump and the statue, although identical at W, might have been distinct. She need only take this to mean that there is a world in which one object—the lump, which in W is the statue—has two counterparts.[18]

Let us suppose that this is a plausible solution to the problem of material constitution. Then it is equally plausible to suppose that among the worlds in which the lump has two counterparts is the actual world. So let us suppose this too; and let us call the two actual counterparts of the lump 'Lumpl' and 'Goliath'. Recall now the translation clause (L1), which interpreted ACT as an existential quantifier over actual-world counterparts. If objects, like our statue at W, can have more than one actual-world counterpart, then again (L1) translates inconsistent sentences of the language of quantified modal logic into sentences that are true in some models of counterpart theory. A case in point is

(22) $a = b \mathbin{\&} ACT(a \neq b)$,

which, although inconsistent, translates via (L1) to

(23) $a = b \mathbin{\&} \exists x \exists y (Ix@ \mathbin{\&} Iy@ \mathbin{\&} Cxa \mathbin{\&} Cyb \mathbin{\&} x \neq y)$.

(23) is true on any interpretation which assigns to both a and b an object that is not in the actual world but has a pair of distinct actual-world counterparts. For example, the statue at W might be assigned to a and b,

[18] Lewis describes this as a case where an identity pair has a non-identity pair as a counterpart, leaving open the question whether the *members* of the pairs need be counterparts as well. See Lewis (1986, §4.5). Multiple counterparts at a world are explicitly invoked by Lewis to account for sentences like "I might have been one of a pair of twins". Our discussion could be adapted to fit such examples.

with Lumpl and Goliath as its distinct actual-world counterparts. Since there are such interpretations, (23) is true in some models of counterpart theory; since (22) is inconsistent, (23) is not the correct translation of (22); and so (L1) should be rejected.

As before, for those philosophers who begrudge our talk of assigning non-actual objects as the denotations of names, the same point can be made using bound variables instead. Consider, for example, the inconsistent sentence (24), a variant of one offered by Hazen (1979):

(24) $\Diamond\exists x(ACTFx \;\&\; ACT\neg Fx)$,

where F is some atomic formula. The clause (L1), applied to (24), yields

(25) $\exists w\exists x[Ixw \;\&\; \exists y(Iy@ \;\&\; Cyx \;\&\; Fy) \;\&\; \exists y(Iy@ \;\&\; Cyx \;\&\; \neg Fy)]$.

"In some world, some object has an F actual counterpart and a not-F actual counterpart."

Again, there are models of counterpart theory in which (25) is true: just consider a model in which x is assigned the statue at W, with its two actual-world counterparts Lumpl and Goliath, exactly one of which is statue-shaped.

The translation clause (L2), which treated ACT as a universal quantifier, fails for analogous reasons, provided an object can have more than one counterpart at the actual world. For (L2) renders satisfiable the inconsistent

(26) $a = b \;\&\; \neg ACT(a = b)$,

translating it to

(27) $a = b \;\&\; \neg\forall x\forall y[(Ix@ \;\&\; Iy@ \;\&\; Cxa \;\&\; Cyb) \supset x = y]$,

which we already know to be true in some models of counterpart theory since it is equivalent in first-order logic to the satisfiable sentence (23), the translation under (L1) of (22). And, as before, for those who resist talk of names with non-actual denotations, the same point can be made using just quantifiers and variables: (L2) translates the inconsistent sentence[19]

(28) $\Diamond\exists x(\neg ACTFx \;\&\; \neg ACT\neg Fx)$,

[19] Again, this is a variant of a sentence offered by Hazen (1979).

where F is some atomic formula, to the sentence

(29) $\exists w \exists x[Ixw \,\&\, (\neg\forall z[(Iz@ \,\&\, Czx) \supset Fz] \,\&\, \neg\forall z[(Iz@ \,\&\, Czx) \supset \neg Fz])]$,

which is true if there is an object in some world with a pair of actual-world counterparts, exactly one of which is F. Once again, the statue at W is such an object, with its pair of actual-world counterparts Lumpl and Goliath.

So both (L1) and (L2) fail to yield a plausible interpretation of sentences involving the actuality operator, on the assumption that there could be objects with multiple actual-world counterparts. What of the schematic clause, (L3), capturing the thought that ACT functions as some kind of quantifier other than an existential or a universal? This clause fares no better. For let Lumpl(x) be a predicate true of Lumpl but false of Goliath. Then the sentence of quantified modal logic

(30) $\Diamond\exists x[ACT(\text{Lumpl}(x)) \equiv ACT(\neg\text{Lumpl}(x))]$.

is inconsistent. Yet it translates under (L3) to

(31) $\exists w \exists x[Ixw \,\&([Qy : Iy@ \,\&\, Cyx](\text{Lumpl}(y)) \equiv [Qy : Iy@ \,\&\, Cyx](\neg\text{Lumpl}(y)))]$.

And as we will shortly see, (31) is true, given the existence of the world W containing a statue with the two actual-world counterparts Lumpl and Goliath. Given this backdrop, if we let s denote the statue in W, then to show that (31) is true it will suffice to establish the following biconditional:

(32) $[Qx : Ix@ \,\&\, Cxs](\text{Lumpl}(x))$ is true iff $[Qx : Ix@ \,\&\, Cxs](\neg\text{Lumpl}(x))$ is true.

Now recall that, since Q is a logical quantifier, $[Qx : \varphi(x)](\psi(x))$ is true if and only if $|\{d : d$ satisfies $\varphi(x)$ and $\psi(x)\}|$, the number of things that satisfy $\varphi(x)$ and $\psi(x)$ (as values of x), has the relation R_Q to $|\{d : d$ satisfies $\varphi(x)$ but not $\psi(x)\}|$, the number of things that satisfy $\varphi(x)$ but not $\psi(x)$ (as values of x), where, as here, the things in question are not too many to be numbered. Given this fact, we can establish the biconditional (32), and hence the satisfiability in counterpart theory of (31), as follows.

$[Qx : Ix@ \;\&\; Cxs](\text{Lumpl}(x))$ is true iff

$|\{d : d$ satisfies $(Ix@ \;\&\; Cxs)$ and $(\text{Lumpl}(x))\}|\ R_Q$

$|\{d : d$ satisfies $(Ix@ \;\&\; Cxs)$ but not $(\text{Lumpl}(x))\}|$ iff

$|\{\text{Lumpl}\}|\ R_Q\ |\{\text{Goliath}\}|$ iff

$1 R_Q 1$ iff

$|\{\text{Goliath}\}|\ R_Q\ |\{\text{Lumpl}\}|$ iff

$|\{d : d$ satisfies $(Ix@ \;\&\; Cxs)$ and $(\neg\text{Lumpl}(x))\}|\ R_Q$

$|\{d : d$ satisfies $(Ix@ \;\&\; Cxs)$ but not $(\neg\text{Lumpl}(x))\}|$ iff

$[Qx : Ix@ \;\&\; Cxs](\neg\text{Lumpl}(x))$ is true. ■

So (32) holds, in the situation described, for any logical quantifier Q. It follows that the sentence (31) is true; and since (31) is the translation under (L3) of the inconsistent sentence (30), we should reject (L3).

Given the structure of the inconsistent sentence (30), the argument just given shows that, if (L3) is to be accepted, then ACT does not commute with negation, on the assumption that at some world there is an object with two counterparts at the actual world.[20] But it is evident that ACT *does* commute with negation, and more generally with every truth-functional connective—these facts are part of the most elementary logic of ‘actually’. Therefore, since the assumption of the argument is precisely the assumption that enables counterpart theory to solve problems of material constitution (and, more generally, to explain the alleged plausibility of contingent identity statements), (L3) must be rejected. But if (L3) is rejected, then there is no way of extending Lewis's counterpart theory to accommodate sentences of quantified modal logic involving the operator ACT.

4. Revisions of Counterpart Theory

We take the observations made so far to show that there is no plausible way of extending Lewis's translation scheme by adding a clause that treats the actuality operator as a quantifier ranging over actual-world counterparts. But what if the counterpart theorist, instead of *extending* Lewis's translation scheme, tried *replacing* it with another scheme? We turn now to address this question.

[20] The argument can be generalized so that it rests only on the assumption that an object at some world has an even or infinite number of counterparts at the actual world: replace ‘Lumpl(x)’ by a formula satisfied and dissatisfied by an equal number of counterparts.

One philosopher to have tried this is Graeme Forbes.[21] Like Lewis, Forbes provides us with a scheme for translating sentences of the language of quantified modal logic into a first-order extensional language. In essence, Forbes's translation scheme is the following:[22]

Forbes's Translation Scheme

• A sentence φ of the language of quantified modal logic is translated to the sentence Rel$[\varphi, @]$ of the language of counterpart theory, the *relativization* of φ to the actual world.

• For any sentence φ of the language of quantified modal logic, and any world-term v, the sentence Rel$[\varphi, v]$ of the language of counterpart theory is defined recursively as follows:

(FA) If $\varphi(a_1, \ldots, a_n)$ is atomic, where $a_1, \ldots, a_n$ are distinct token occurrences of terms, then Rel$[\varphi(a_1, \ldots, a_n), v]$ is
$\exists x_1 \ldots \exists x_n (Ix_1 v \,\&\, \ldots \,\&\, Ix_n v \,\&\, Cx_1 a_1 \,\&\, \ldots \,\&\, Cx_n a_n \,\&\, \varphi(x_1, \ldots, x_n))$.

(F$\neg$) Rel$[\neg\varphi, v]$ is $\neg$Rel$[\varphi, v]$.

(F&) Rel$[\varphi \,\&\, \psi, v]$ is Rel$[\varphi, v]$ & Rel$[\psi, v]$.

(F$\forall$) Rel$[\forall x\varphi, v]$ is $\forall x(Ixv \supset$ Rel$[\varphi, v])$.

(F$\exists$) Rel$[\exists x\varphi, v]$ is $\exists x(Ixv \,\&$ Rel$[\varphi, v])$.

(F$\Box$) Rel$[\Box\varphi, v]$ is
$\forall w \forall x_1 \ldots \forall x_n ([Ix_1 w \,\&\, \ldots \,\&\, Ix_n w \,\&\, Cx_1 a_1 \,\&\, \ldots \,\&\, Cx_n a_n] \supset$ Rel$[\varphi(a_i/x_i), w])$, if $a_1, \ldots, a_n$ are the occurrences of unbound terms in φ not within the scope of any modal operator in φ; Rel$[\Box\varphi, v]$ is $\forall w$Rel$[\varphi, w]$ if there are no such terms.

(F$\Diamond$) Rel$[\Diamond\varphi, v]$ is
$\exists w \exists x_1 \ldots \exists x_n (Ix_1 w \,\&\, \ldots \,\&\, Ix_n w \,\&\, Cx_1 a_1 \,\&\, \ldots \,\&\, Cx_n a_n \,\&$ Rel$[\varphi(a_i/x_i), w])$, if $a_1, \ldots, a_n$ are the occurrences of unbound terms

[21] See Forbes (1982), Forbes (1985) and Forbes (1990).

[22] For reasons that won't matter to us, Forbes views the counterpart relation as a three-place, not a two-place, relation, and he allows objects to have as counterparts at a world things which are not in that world; we ignore these features in our presentation. The translation scheme given here is a version of the one given in Forbes (1990), containing a modification of Forbes's earlier scheme, given in response to some objections raised by Ramachandran (1989).

in φ not within the scope of any modal operator in φ; Rel$[\Diamond\varphi, v]$ is $\exists w$Rel$[\varphi, w]$ if there are no such terms.

(F-ACT) Rel$[ACT\varphi, v]$ is Rel$[\varphi, @]$.

Forbes's translation scheme avoids the objection we raised to extensions of Lewis's scheme, the objection that inconsistent sentences are translated to sentences true in some models of counterpart theory. For example, one objectionable sentence we considered was

(18) $Fa \;\&\; (ACT\neg Fa \vee \neg ACT Fa)$.

But (18) translates, via Forbes's scheme, to an unsatisfiable sentence. Forbes's scheme manages this, since, as Forbes himself notes, unlike the extensions of Lewis's scheme that we have considered, it respects the fact that the operator ACT commutes with every truth-functional connective.

Be that as it may, problems remain with Forbes's translation scheme which show it to be unacceptable. Already fatal is the fact that identity is rendered non-transitive, in the sense that Forbes's scheme translates

(33) $\exists x \exists z \Diamond \exists y ACT(x = y \;\&\; y = z \;\&\; x \neq z)$

to a sentence true in some models of counterpart theory.[23] To see this, suppose a non-actual object Yolanda has a pair of actual-world counterparts, Alfred and Agnes. (Like Lewis, Forbes allows objects to have more than one counterpart at a world.) Then an actual-world object (Alfred) shares an actual counterpart with Yolanda, who shares an actual counterpart with another actual-world object (Agnes). Yet these two actual-world objects, Alfred and Agnes, have no counterparts in common at the actual world, since each is their own sole counterpart there.

This problem can be generalized. Not only does Forbes's scheme render *identity* non-transitive, in fact almost *any* intuitively transitive, reflexive relation is rendered non-transitive. Let R be an intuitively transitive, reflexive relation, let Yolanda, Alfred and Agnes be as above, and suppose Alfred does not bear R to Agnes. Then, according to Forbes's translation scheme (and assuming that 'x bears R to y' is an atomic formula), 'Alfred bears R to Yolanda' will be true, 'Yolanda bears R to Agnes' will be true, yet 'Alfred

[23] A similar point is made by Ramachandran (1990*b*), although, as he acknowledges, his way of putting it requires the language to contain names of actually non-existent objects. We note that Forbes (1982) was working with a background of fuzzy logic, on which there are degrees of satisfiability. Given this background, and the assumption that the degrees to which some objects are counterparts of something cannot sum to more than one, (33) translates to a sentence that is only partially satisfiable.

bears R to Agnes' will be false. Dispensing with the names of non-actuals, the relation R is rendered non-transitive by Forbes in the sense that his scheme translates

(34) $\exists x \exists z \Diamond \exists y ACT(Rxy \,\&\, Ryz \,\&\, \neg Rxz)$

to a truth.

Forbes's supposition that there could be an object with distinct actual-world counterparts also makes trouble when his translation scheme is applied to monadic predicates. For given an object with distinct actual-world counterparts there will be properties F and G such that one counterpart instantiates F, the other counterpart instantiates G, and, at the actual world, no object instantiates both F and G. If the counterparts are Alfred and Agnes, for example, we could take F and G to be the essence-specifying properties *being Alfred* and *being Agnes* respectively. Or, for those suspicious of such properties, we could suppose that one of our object's counterparts is spherical and one is cuboid, and take F and G to be the incompatible properties *being spherical* and *being cuboid*. But given this assumption—that there could be an object with one actual-world counterpart which is F, and one actual-world counterpart which is G, where nothing at the actual world is both F and G—if Forbes's translation scheme were correct, it would follow, given his clause (FA), and the fact that Forbes requires every object to be its own sole counterpart at the world it is in, that there could be an object which was actually *both* F and G, even though, actually, *nothing* is both F and G! Forbes's translation scheme thus renders satisfiable the manifestly inconsistent sentence

(35) $\Diamond \exists x ACT[\exists y(x = y) \,\&\, Fx \,\&\, Gx \,\&\, \forall y \neg(Fy \,\&\, Gy)]$.

Forbes's Canonical Counterpart Theory therefore fares no better than extensions of Lewis's counterpart theory when it comes to sentences involving the operator ACT. But Forbes is not the only philosopher to have proposed variants of counterpart theory. Over a series of papers, Murali Ramachandran has developed the theory he calls CT*, which is essentially Lewis's counterpart theory together with a new translation scheme.[24]

[24] See Ramachandran (1989), Ramachandran (1990*a*) and Ramachandran (1990*b*). The translation scheme we give here is that of Ramachandran (1989); in later versions, Ramachandran adds a special clause governing identity sentences. This modification has no bearing on our discussion.

Ramachandran's Translation Scheme

• A sentence φ of the language of quantified modal logic is translated to the sentence $[\varphi]^@$ of the language of counterpart theory, the *relativization to the actual world* of the *atomic translation* of φ.

• For any sentence φ of the language of quantified modal logic, its atomic translation $[\varphi]$ is the result of replacing every atomic constituent S of φ that contains *modally-free* occurrences of terms $t_1, \ldots, t_n$ with

$$\exists x_1 Cx_1t_1 \mathbin{\&} \ldots \mathbin{\&} \exists x_n Cx_nt_n \mathbin{\&} \forall x_1 \ldots \forall x_n((Cx_1t_1 \mathbin{\&} \ldots \mathbin{\&} Cx_nt_n) \supset S[x_i/t_i]),$$

where $S[x_i/t_i]$ is the result of replacing, for each i, every occurrence of t_i in S with x_i. An occurrence of a term in a sentence is modally-free if it is free in the largest substring of the sentence containing the occurrence of the term but no modal operators.

• For any sentence φ of the language of quantified modal logic (where that language is taken to include a two-place predicate C), and any world-term v, the sentence φ^v of the language of counterpart theory is defined recursively, exactly as in Lewis's translation scheme but with an additional clause for the actuality operator, and modified clauses for the other modal operators $\Box$ and $\Diamond$:

(R-ACT) $(ACT\varphi)^v$ is $\varphi^@$.

(R$\Box$) $(\Box\varphi)^v$ is $\forall w \varphi^w$.

(R$\Diamond$) $(\Diamond\varphi)^v$ is $\exists w \varphi^w$.

Like Forbes's Canonical Counterpart Theory, Ramachandran's translation scheme avoids the objections to extensions of Lewis's scheme, by respecting the fact that ACT commutes with all truth-functional connectives. And it improves on Forbes's theory by preserving the transitivity of identity and other transitive relations. Moreover, Ramachandran's translation scheme correctly renders unsatisfiable the sentence that was fatal to Forbes's theory:

(35) $\Diamond \exists x ACT[\exists y(x = y) \mathbin{\&} Fx \mathbin{\&} Gx \mathbin{\&} \forall y \neg(Fy \mathbin{\&} Gy)]$.

Ramachandran's scheme manages this by requiring that if there could be an object which is actually F and G, it is not enough that the object have some actual-world counterpart which is F and some actual-world counterpart which is G: in addition *all* of the object's actual-world counterparts must be both F and G—and this is ruled out by the last conjunct of (35).

Unfortunately, avoiding this problem in this way creates another just as bad. Consider the inconsistent sentence

(36) $\Diamond\exists x ACT[Ex \,\&\, \neg(Fx \vee Gx) \,\&\, \forall y(Fy \vee Gy)]$,

where E is an atomic existence-predicate, true at a world of just those things that have at least one counterpart at that world.[25] It is compatible with Ramachandran's CT* that there could be an object with two actual-world counterparts, one of which was F but not G, the other of which was G but not F. But then not all of this object's actual-world counterparts are F, nor are all of them G. We need then only suppose, in addition, that *every* object in the actual world is either F or G to see that (36) translates to a truth—even though it is inconsistent.

Ramachandran nowhere says whether he follows Lewis[26] and Forbes in requiring that, at any world, an object has itself as its sole counterpart in that world. If he does not impose this requirement, then his scheme renders satisfiable the inconsistent

(37) $ACT\exists x(Fx \,\&\, ACT\neg Fx)$,

since it translates to

(38) $\exists x(Ix@ \,\&\, Fx \,\&\, \neg[\exists y(Iy@ \,\&\, Cyx) \,\&\, \forall y((Iy@ \,\&\, Cyx) \supset Fy)])$,

which will be true if there is an actual object which is F but which has an actual-world counterpart which is not F.

Further problems for Ramachandran's translation scheme concern its treatment of some binary relations between objects. Consider, for instance, *connected* relations, where a relation R is connected over a domain iff, for any two objects in the domain, either the first bears R to the second or the second bears R to the first. So, for example, the relation of *being at least as tall as* is connected over the domain of people; the relation of *being less than or equal to* is connected over the domain of ordinals. But now consider a consequence of a relation R's being (actually) connected (over some domain—we leave the domain implicit):

[25] This existence-predicate is Ramachandran's, introduced to allow objects to count as existing at worlds at which they have more than one counterpart. At the atomic stage of translation, Et is to be translated as $\exists xCxt$. See Ramachandran (1989, p. 136).

[26] The early Lewis, that is. Lewis later retracted this requirement to enable counterpart theory to respect some haecceitist intuitions without being committed to haecceitism. See Lewis (1986, p. 232, n. 22).

(39) $\Box\forall x\forall y ACT[(Ex \,\&\, Ey) \supset (Rxy \vee Ryx)]$.

The translation of (39) under Ramachandran's scheme is complex, so we omit it here, but it is enough to note that this translation will be true only if, for every pair of possible objects, each of the actual-world counterparts of one of the pair bears R to all the actual-world counterparts of the other. But this need not be so, even if R is a connected relation. For example, taking the implicit domain to be the class of people, a person may have two actual-world counterparts, one shorter and one taller than someone else's only actual-world counterpart; it should not follow, as it would if Ramachandran's translation scheme were truth-preserving, that the relation of *being at least as tall as* is not connected over the domain of people.[27]

Ramachandran's revision of counterpart theory, therefore, is no more successful than Forbes's, and neither of them fares better than extensions of Lewis's translation scheme. This fact by itself, of course, does not show that *no* translation scheme for counterpart theory could succeed. But there is a quite general reason that we have not yet mentioned for thinking that there cannot be a plausible translation scheme for any version of counterpart theory that allows for possible objects with multiple actual-world counterparts.

Assume only that there is a possible object *symmetrically* related to its two actual-world counterparts, in the sense that neither counterpart is any more of a counterpart of the object than the other. To fix ideas, let the lump of clay in the world W that we considered earlier be such a possible object, symmetrically related to its two actual-world counterparts Lumpl and Goliath. Given the symmetry, the lump should at W satisfy

(40) $ACT(\text{Lumpl} = x) \equiv ACT(\text{Goliath} = x)$.

[27] We note also that the sensitivity on Ramachandran's translation scheme to the question of whether an occurrence of a term is "modally-free" in a sentence has some bizarre consequences. For example, the second occurrence of the variable x is modally-free in (i) but not in (ii):

(i) $\exists x(\Box(p \supset p) \,\&\, Fx)$.
(ii) $\exists x Fx$.

Since $\Box(p \supset p)$ is a theorem of any sensible modal logic, (i) and (ii) are equivalent. Yet if, at the actual world, there is an object in the extension of F with a counterpart at the actual world which is *not* in the extension of F, then the translation under Ramachandran's scheme of (i) will be false, while that of (ii) will be true. Any translation scheme that assigns different truth-values to the translations of logically equivalent sentences is unacceptable. And even if Ramachandran disallows multiple worldmate counterparts of objects, the difference in semantic treatment between (i) and (ii) is wholly implausible: (ii) translates to itself under Ramachandran's scheme, while (i) translates to

(iii) $\exists x(Ix@ \,\&\, \forall w((p \supset p) \,\&\, \exists y(Iyw \,\&\, Cyx) \,\&\, \forall y(Iyw \supset (Cyx \supset Fy))))$.

Ramachandran informs us (p.c.) that, instead of the definition of "modally-free" that he gives in Ramachandran (1989), he now prefers the following:

> The occurrence of a term-token t in a sentence is *modally free* iff *either* t is not bound by a quantifier at all, *or* it is bound by a quantifier but also occurs within the scope of a modal operator that occurs between the quantifier and t.

Since Lumpl and Goliath are the only actual-world counterparts of the lump of clay at W and both satisfy $(\text{Lumpl} = x) \vee (\text{Goliath} = x)$ as values of x, the lump should at W satisfy

(41) $ACT[(\text{Lumpl} = x) \vee (\text{Goliath} = x)]$.

And since Lumpl and Goliath are (actually) distinct, neither satisfies $(\text{Lumpl} = x)$ & $(\text{Goliath} = x)$ as a value of x, so the lump should at W satisfy

(42) $\neg ACT[(\text{Lumpl} = x)$ & $(\text{Goliath} = x)]$.

But if the lump satisfies, at W, each of (40), (41) and (42), then at W it satisfies their conjunction. So if counterpart theory could accommodate sentences involving ACT, it would predict the satisfiability of

(43) $\Diamond \exists x([ACT(\text{Lumpl} = x) \equiv ACT(\text{Goliath} = x)]\ \&$

$ACT[(\text{Lumpl} = x) \vee (\text{Goliath} = x)]\ \&$

$\neg ACT[(\text{Lumpl} = x)\ \&\ (\text{Goliath} = x)])$.

But since ACT commutes with both conjunction and disjunction (even in the scope of $\Diamond$), and since $(p \equiv q)\ \&\ (p \vee q)\ \&\ \neg(p\ \& q)$ is truth-functionally inconsistent, (43) implies a contradiction.[28]

5. Concluding Remarks

We have argued that there is no coherent way to extend Lewis's scheme for translation from the language of quantified modal logic to the language of counterpart theory, if quantified modal logic is regarded, as it should be, as containing an actuality operator. We have argued, too, that the revisions of Lewis's translation scheme offered by Forbes and Ramachandran are also incoherent. And we have just given a general reason for thinking that the prospects for *any* coherent translation scheme are dim.

As we noted along the way, however, the problems that we raised for the various translation schemes made essential use of the fact that, according to counterpart theory, objects could have more or less than one counterpart at the actual world. Our arguments could therefore be resisted if the counterpart theorist made the assumption that every possible object has exactly one actual-world counterpart. If this assumption is made, however, then many of the alleged benefits of counterpart theory, and so its motivations, disappear.

[28]This argument can be generalized to any case where there are more than two counterparts symmetrically related to a possible object, including the case where an object has infinitely many counterparts at the actual world.

First, if the assumption is made, then although counterpart theory could still make sense of some contingent-identity statements, it could not make sense of others. In particular, given the assumption, (44) would translate to an unsatisfiable sentence, even though (45) would translate to a satisfiable one:

(44) $\Diamond\exists x\exists y(x = y \,\&\, ACTx \neq y)$.
"There could have been an identical x and y that are actually distinct."

(45) $\exists x\exists y(x = y \,\&\, \Diamond x \neq y)$.
"There are an identical x and y that could have been distinct."

Consequently, some versions of problems of material constitution cannot be solved by the counterpart theorist who adopts the assumption. Lumpl and Goliath are identical and forever spatially coincident, but could have been different; but Lumpl and Goliath couldn't have been identical and forever spatially coincident while being *actually* different. This seems an odd combination of views.

Second, recall that counterpart theory was supposed to allow us to endorse claims like "I might have been Frank Sinatra while everything else is as it actually is", without being committed to haecceitism.[29] The trick here is to allow that, in the actual world, I have two counterparts, Frank Sinatra and myself.[30] But this trick is unavailable once the counterpart theorist makes the assumption that every possible object, including me, has exactly one actual-world counterpart.

Third, with this assumption in place, counterpart theory is committed to the claim that it is necessary that everything actually exists.[31] For if every possible object has exactly one actual-world counterpart, then for every possible object that counterpart serves to represent it as actually existing.[32] Some philosophers may welcome the claim that it is necessary that everything actually exists, on the ground that it sits well with their extreme actualist tendencies. Others may welcome it because they view it as a consequence of a Barcan formula, and hence as a theorem of the simplest quantified modal logic. Still others may be attracted to the claim because, recalling our discussion in the Introduction, it facilitates them in retaining Kripke semantics for quantified modal logic while at the same time endorsing (7) as the formalization of (4):

[29] Of course, not every counterpart theorist need find this consideration motivating.

[30] The trick therefore requires, as Lewis (1986) notes, rejecting one of Lewis's postulates of counterpart theory, the postulate that ensures that every object has only itself as a counterpart in the world at which it exists.

[31] We note, though, that Forbes, in treating the counterpart relation as a three-place relation, avoids this commitment by allowing objects to have counterparts *at* the actual world without having any counterparts *existing in* the actual world.

[32] This commitment could be avoided if, for example, the counterpart theorist required (somewhat bizarrely) that for an object to be represented as existing at a world it must have more than one counterpart at that world. But then, with the assumption being considered here, that counterpart theorist would be committed to the unacceptable claim that it is necessary that *nothing* actually exists.

(4) It might have been that everyone who is in fact rich was poor.

(7) $\Diamond\forall x(ACTRx \supset Px)$.

But the claim that it is necessary that everything actually exists is surely something that a defender of counterpart theory would be loathe to embrace.

Fourth, the assumption that every possible object has exactly one counterpart in the *actual* world seems arbitrary. If this assumption is to be adopted, why not adopt the stronger assumption that every possible object has exactly one counterpart in *every* possible world? Refusing to privilege the actual world in this way would seem to fit with the plausible view (shared by Lewis) that it is a contingent matter which world is actual. Adopting this stronger assumption, however, would effectively collapse counterpart theory, yielding the necessity of identity, the validity of the claim that, necessarily, everything exists necessarily, and along with that the falsity of just about any unconditional non-trivial essentialist claim that may be proposed—so much for "inconstancy of representation *de re*".[33] Lewis himself is quite explicit:

> It would not have been plausible to postulate that nothing in any world had more than one counterpart in any other world. . . [and] it would not have been plausible to postulate that, for any two worlds, anything in one had some counterpart in the other. (Lewis 1968, p. 29)

Finally, it seems that the whole notion of what a counterpart is supposed to be militates against making even the weaker assumption, that every possible object has exactly one *actual*-world counterpart. Recall Lewis's explanation of the counterpart relation:

> Your counterparts resemble you closely in content and context in important respects. They resemble you more closely than do the other things in their worlds.

It is presumably possible that there be something utterly unlike anything that actually exists. If so, that possible object—call it Ghost—does not resemble any actually existing object in *any* important respect. So Ghost, if we are to understand the counterpart relation as Lewis instructs us, should have no actual-world counterparts. And even if we could be convinced that there is *some* respect in which Ghost resembles things

[33]The counterpart theorist could still cling to the contingency of distinctness, on the assumption that distinct objects can share a counterpart. But given the necessity of identity, for distinctness to be contingent would be odd. Worse, the necessity of distinctness is derivable, in modal logics which contain the B axiom (for example the standard system S5), from the necessity of identity; see Prior (1955, pp. 206–7) and Kripke (1980, p. 114). And it is derivable even without the B axiom in the logic of ACT; see Williamson (1996).

in the actual world, how could we be convinced that there is a *single* thing in the actual world which Ghost resembles *more closely* than it does anything else in the actual world?

For these reasons, the only way for counterpart theory to enjoy superiority over quantified modal logic as a means of formalizing modal discourse would be for it to reject the assumption that every possible object has exactly one actual-world counterpart. But, as we have argued, without this assumption there is no plausible, systematic translation from the language of quantified modal logic with an actuality operator to the language of counterpart theory.

This leaves two options for formalizing modal discourse. One option is to embrace modalism, to take the intensional language of quantified modal logic as basic, either not in need of or not amenable to further explication. The other option is to pursue anti-modalism, and attempt to explain the semantic content of modal operators in extensional terms—but not by invoking a counterpart relation. Indeed, there is already a familiar and elegant anti-modalist semantics: orthodox Kripke semantics, on which the modal operators are interpreted as quantifiers over worlds, and on which a single individual may be in the domain of many worlds.

We turn, finally, to consider an objection to our methodology. We have been following Lewis (1968) in supposing that if counterpart theory is to provide a means of formalizing modal claims of natural language, then it must be supplemented with a systematic method of translation from those claims, via the language of quantified modal logic, to the language of counterpart theory. But what if this supposition were rejected, and instead it were insisted that we can do without a systematic method of translation?[34] Can we not just work out how to formalize ordinary modal claims in the language of counterpart theory case by case? If so, then the fact that there can be no systematic translation method does not refute counterpart theory.

This objection can be taken in two ways. On the one hand, it can be taken as a claim about the semantics of natural language, to the effect that occurrences of expressions which function as actuality operators are to be represented semantically as quantifiers ranging over actual-world counterparts, even though (given the examples we have considered) some of these occurrences are to be represented (say) as existential quantifiers while others are to be represented as universal quantifiers. On this way of taking the objection, however, it requires accepting that the semantics for natural language is radically non-compositional. This is an extremely strong commitment, one which seems insufficiently motivated by the present considerations.

[34] This strategy was later recommended by Lewis (1986, pp. 12-13).

On the other hand, the objection could be taken as a metaphysical claim, to the effect that modal claims of natural language are made true by facts about objects' counterparts, and that even without a translation scheme we can in principle determine, for any modal claim involving an actuality operator, which facts about an object's actual-world counterparts are its truth-makers. Taken this way, however, the objection seems implausible. What reason is there to suppose, independently of a translation scheme that *demonstrates* it, that *de re* modal claims have anything to do with an object's counterparts? In the absence of a translation scheme, Kripke's famous objection to counterpart theory seems wholly appropriate. When we say that Humphrey might have won the election, it seems that we are talking about what might have happened to *Humphrey*, not about what does happen to someone else (Kripke 1980, p. 45). The standard response to Kripke, given by Hazen (1979) and many others, is to point out that counterpart theory is being offered as a semantic account of modal sentences, and that according to that account when we talk about what might have happened to Humphrey we *are* talking about what happens to someone else, his counterpart. But without a translation scheme, the counterpart theorist is not offering a semantic account of modal sentences, and so this standard response is unavailable to her. Unless there is a plausible translation scheme to suggest otherwise, which we have argued there is not, we have every reason to think that the meaning of modal sentences has nothing to do with counterparts; and since the meaning of a sentence determines its truth-conditions, and thereby constrains the range of its potential truth-makers, there is every reason to think that modal sentences are not made true by facts about objects' counterparts.

In any case, in either its semantic or its metaphysical version the objection being considered seems ill-placed to give an adequate account of the sentence (43) that we considered at the end of the last section:

(43) $\Diamond\exists x([ACT(\text{Lumpl} = x) \equiv ACT(\text{Goliath} = x)]\ \&$

$ACT[(\text{Lumpl} = x) \vee (\text{Goliath} = x)]\ \&$

$\neg ACT[(\text{Lumpl} = x)\ \&\ (\text{Goliath} = x)])$.

Since the actuality operator commutes with both conjunction and disjunction, (43) is inconsistent. Yet we showed that it will emerge as a truth given any plausible interpretation of ACT as a quantifier ranging over objects' actual-world counterparts, provided only that there is a possible object symmetrically related to its counterparts Lumpl and Goliath. Independently of considerations of systematicity, therefore, we have in (43) a *de re* modal sentence that should not be understood in terms of counterparts. Considerations of uniformity now mandate that *no de re* modal sentence should be understood in terms of counterparts.[35]

[35]We would like to thank Josh Dever, Cian Dorr, Graeme Forbes, Delia Graff, Allen Hazen, James Joyce, Josh Parsons, Murali Ramachandran, Wolfgang Schwarz, Theodore Sider, Yannis Stephanou, Zoltán Szabó and Richmond Thomason, as well as audiences at University of Michigan, University of Rochester, and University of Texas at Austin, for helpful comments on earlier versions of this work.

References

Cresswell, M. (1990), *Entities and Indices*, Kluwer Academic Publishers.

Crossley, J. & Humberstone, L. (1977), 'The logic of 'actually'', *Reports on Mathematical Logic* **8**: 11–29.

Davies, M. (1981), *Meaning, Quantification, Necessity*, Routledge & Kegan Paul.

Forbes, G. (1982), 'Canonical counterpart theory', *Analysis* **42**: 33–37.

Forbes, G. (1985), *The Metaphysics of Modality*, Oxford University Press, Oxford.

Forbes, G. (1989), *Languages of Possibility*, Blackwell, Oxford.

Forbes, G. (1990), 'Counterparts, logic and metaphysics: Reply to Ramachandran', *Analysis* **50**: 167–173.

Gibbard, A. (1975), 'Contingent identity', *Journal of Philosophical Logic* **4**: 187–221.

Hazen, A. (1976), 'Expressive completeness in modal languages', *Journal of Philosophical Logic* **5**: 25–46.

Hazen, A. (1979), 'Counterpart-theoretic semantics for modal logic', *Journal of Philosophy* **76**: 319–338.

Hazen, A. (1990), 'Actuality and quantification', *Notre Dame Journal of Formal Logic* **31**: 498–508.

Hodes, H. (1984*a*), 'Axioms for actuality', *Journal of Philosophical Logic* **13**: 27–34.

Hodes, H. (1984*b*), 'Some theorems on the expressive limitations of modal languages', *Journal of Philosophical Logic* **13**: 13–26.

Kripke, S. (1963), 'Semantical considerations on modal logic', *Acta Philosophica Fennica* **16**: 83–94.

Kripke, S. (1980), *Naming and Necessity*, Harvard University Press.

Lewis, D. (1968), 'Counterpart theory and quantified modal logic', *Journal of Philosophy* **65**: 113–126. Reprinted in (Lewis 1983, pp. 26–39).

Lewis, D. (1983), *Philosophical Papers*, Vol. I, Oxford University Press, Oxford.

Lewis, D. (1986), *On the Plurality of Worlds*, Blackwell, Oxford.

Linsky, B. & Zalta, E. (1994), 'In defense of the simplest quantified modal logic', *Philosophical Perspectives* **8**: 431–58.

Prior, A. N. (1955), *Formal Logic*, Clarendon Press, Oxford.

Prior, A. N. (1956), 'Modality and quantification in S5', *Journal of Symbolic Logic* **21**: 60–62.

Ramachandran, M. (1989), 'An alternative translation scheme for counterpart theory', *Analysis* **49**: 131–141.

Ramachandran, M. (1990*a*), 'Contingent identity in counterpart theory', *Analysis* **50**: 163–166.

Ramachandran, M. (1990*b*), 'Unsuccessful revisions of cct', *Analysis* **50**: 173–177.

Westerståhl, D. (1989), 'Quantifiers in formal and natural languages', in D. Gabbay, ed., *Handbook of Philosophical Logic*, Vol. IV, D. Reidel Publishing Company, pp. 1–131.

Williamson, T. (1996), 'The necessity and determinateness of distinctness', in S. Lovibond & S. G. Williams, eds., *Essays for David Wiggins: Identity, Truth and Value*, Blackwell, Oxford.

Williamson, T. (1998), ‘Bare possibilia’, *Erkenntnis* **48**: 257–273.

Williamson, T. (2002), ‘Necessary existents’, in A. O’Hear, ed., *Logic, Thought and Language*, Cambridge University Press, Cambridge.