Representing Counterparts

Andrew Bacon*

Department of Philosophy
University of Southern California
Los Angeles
CA 90089
abacon@usc.edu

Abstract

This paper presents a counterpart theoretic semantics for quantified modal logic based on a fleshed out account of Lewis's notion of a 'possibility'. According to the account a possibility consists of a world and some haecceitistic information about how each possible individual gets represented *de re*. Following Hazen [4], a semantics for quantified modal logic based on evaluating formulae at possibilities is developed. It is shown that this framework naturally accommodates an actuality operator, addressing recent objections to counterpart theory (see [2], [1]), and is equivalent to the more familiar Kripke semantics for quantified modal logic with an actuality operator.

One of the most important insights of possible world semantics is captured by the Leibnizian biconditionals which relate the notions of necessity and possibility to the notion of a possible world:

LB$\Box$ p is necessary if and only if p is true at every possible world.

LB$\Diamond$ p is possible if and only if p is true at some possible world.

*I would like to thank the audience of the Oxford-Paris Workshop on Language and Ontology in June 2008 for many helpful comments. Special thanks are due to Jeff Russell and Timothy Williamson for providing me with substantial comments on an earlier version of this paper and an anonymous reviewer for this journal who suggested some valuable improvements for this version. Since the acceptance of this paper by this journal (in 2011) several other solutions to the problem addressed in this paper have appeared in print which are for this reason regrettably absent from my discussion (see Russell [13], Meyer [10], Rigoni and Thomason [12].)

For possible world semantics to be useful to the study of modality it suffices only that these biconditionals be necessary.[1] It is not important in this regard that the biconditionals serve as an analysis of possibility or necessity.

Indeed, it might initially seem unlikely that they could serve as an analysis of possibility and necessity since modal notions appear in the right hand side of the biconditional. The concept of a possible world – roughly, a way things could have been, as opposed to a way things couldn't have been – appears to presuppose the very modal concepts we are trying to analyse.

Nonetheless philosophers have attempted such reductions. Here is a putative reductive analysis of 'possible world', inspired by Lewis [9], that defines a world in terms of the parthood relation and the property of being an open region of spacetime.[2] Since neither of these concepts require an antecedent understanding of possibility an analysis appears to be in the offing. With the notion of a fusion introduced in the ordinary way, a possible world is a region which is not the fusion of two disjoint open regions and which is furthermore identical to any open region which it is a part of and which is itself not the fusion of two disjoint open regions.[3]

To have a reductive analysis of possibility and necessity via the Leibnizian biconditionals one also needs an explication of what it means for a proposition to be *true at* one of these spacetime regions. Luckily we have a decent grasp of what kind of things are going on in a region of spacetime, and it plausibly doesn't require an antecedent understanding of modality. We know, for example, how to determine what was going on in the region of spacetime occupied by Italy throughout 49BC – in that region Julius Caesar was crossing the Rubicon. On the other hand, Neville Chamberlain was not declaring war on Germany at that region – Chamberlain is not even present in that region.

Note, however, that it is exactly this naïve conception of *truth at* that forces one to relinquish the Leibnizian biconditionals. Let us suppose that Humphrey is not a figure skater, but that he could have been. Then $\mathsf{LB}\Diamond$ states that:

[1] Or, without assuming the $\mathsf{S4}$ principle, that they are necessarily necessary, necessarily necessarily necessarily and so on.

[2] Given these concepts, an arbitrary region of spacetime may be defined as any part of the fusion of all open regions.

[3] Thus, according to my definition, a world need not be path connected or locally connected. Lewis talks of two objects being spatio-temporally interelated, but does not distinguish between the possible things this could mean; the above is merely one way making that precise. (Sometimes he glosses this as there being a 'distance between them – be it great or small, spatial or temporal' (p70 [9]), however this gloss seems to be inadequate for spacetimes which cannot be consistently assigned a global metric.)

> The proposition that Humphrey is a figure skater is possible iff the proposition that Humphrey is a figure skater is true at some possible world.

Yet this instance of LB◇ is false: the proposition that Humphrey is a figure skater is clearly possible, yet there is no world at which it is true. By hypothesis Humphrey is not a figure skater at the region of spacetime spatiotemporally connected to us. Furthermore, he isn't even present at regions disconnected from us so, going by the analogy with truth in Italy 49BC, he isn't a figure skater at these regions either.

Lewis, however, rejects the orthodox possible world semantics founded on the Leibnizian biconditionals, and proposes that we adopt instead his own counterpart theory. The resulting theory is slightly awkward, and it is often noted that it delivers some unwanted results (see for example Hazen [4], Lewis [8] and Kripke [6] p45, fn13). The most recent of these have been pointed out by Fara and Williamson [2] and arise when one wants to make sense of claims formulated using an actuality operator.

In this paper I suggest a couple of alternative ways to develop Lewis's framework. Unlike Lewis's theory, however, both accounts retain the Leibnizian biconditionals that are characteristic of orthodox possible worlds semantics. According to the first account, instead of Lewisian worlds, one evaluates propositions relative to slightly more fine grained entities: *possibilities* – a notion introduced by Lewis himself, and developed by Hazen [4]. The second theory, discussed in §2, works with a slightly more sophisticated understanding of *truth at*.

The paper is organised as follows. In §1 I outline some of the technical difficulties involved in giving quantified modal logic without identity[4] a counterpart theoretic semantics, especially with respect to the problem of actuality operators. In §3 I expand on the second counterpart semantics and show that the logic coincides with that validated by the most general class of Kripke models (in fact, Kripke's original semantics comes out as a special case of counterpart semantics in which the counterpart relation is the identity relation.) In the appendix I address some miscellaneous issues such as how identity should be treated in the framework, and extensional formulations of the theory.

1 Counterpart semantics

To correctly evaluate the proposition that Humphrey could have been a figure skater, according to Lewis, we must put Humphrey to one side and instead

[4]The treatment of identity is slightly more subtle; I deal with these in the appendix.

look at the men sufficiently similar to Humphrey, his counterparts, and see whether there are figure skaters among them. This is, at a first parse, the central idea behind counterpart theory.

However the devil is in the details. Lewis's original 1968 formulation of counterpart theory has seen it's fair share of problems and has undergone various transformations by Lewis and others.[5] The most recent round of problems, affecting all the aforementioned versions, are those raised by Fara and Williamson [2] associated with translating sentences containing the actuality operator into counterpart theory.

Fara and Williamson show that various natural ways of translating sentences of quantified modal logic with an actuality operator fail and thus, insofar as that part of English is correctly intertranslatable with quantified modal logic, counterpart theory fails to faithfully represent English. To be precise, these translations fail because they translate inconsistent formulae of quantified modal logic to consistent formulae of counterpart theory.[6]

One might quibble that this is too stringent a criteria of failure. For example, one of the heralded features of counterpart theory is that it has a non-standard logic of identity. In what follows I will restrict attention to the identity free fragment of the language of quantified modal logic with an actuality operator (for short: $\mathsf{QML@}$); the worries we shall be considering will be completely independent of the counterpart theoretic treatment of identity.

Let us begin by outlining Lewis's translation of (identity free) quantified modal logic without the actuality operator (QML for short) into counterpart theory and see how talk about actuality might be accommodated. I shall formulate Lewis's theory with two primitive predicates: Ixw and Cxy to be read as 'x is in w' and 'y is a counterpart of x' respectively, a name, w^*, for the actual world, and two sorts of variables: w, w', that vary over worlds and x, y, z which range over individuals.

A way to translate a formula of QML into Lewis's counterpart theory is described below. Translation is always with respect to a world variable.

$$(Px_1, \ldots, x_n)^w \mapsto Px_1, \ldots, x_n$$
$$(\neg\phi)^w \mapsto \neg(\phi^w)$$
$$(\phi \vee \psi)^w \mapsto (\phi^w \vee \psi^w)$$
$$(\exists x\phi)^w \mapsto \exists x(Iwx \wedge \phi^w)$$
$$(\Diamond\phi)^w \mapsto \exists w'\exists x_1, \ldots, x_n(Iw'x_1 \wedge \ldots \wedge Iw'x_n \wedge Ct_1x_1 \wedge \ldots \wedge Ct_nx_n \wedge \phi^{w'})$$

In the last clause $t_1, \ldots, t_n$ are the free terms in ϕ.

[5] Lewis [8], Ramachandran [11], Forbes [3].

[6] 'Inconsistent' here is presumably supposed to mean more than merely 'unsatisfiable according to the Kripke semantics'.

Fara and Williamson note that extending this translation to the richer language containing an actuality operator is not as simple as it seems. To demonstrate we'll consider the two most obvious translations. According to the first, to be actually F some of your actual counterparts must be F, and according to the second all of them must be:

$$(@\phi)^w \mapsto \exists x_1, \ldots, x_n(Iw^*x_1 \wedge \ldots \wedge Iw^*x_n \wedge Ct_1x_1 \wedge \ldots \wedge Ct_nx_n \wedge \phi^{w^*})$$
$$(@\phi)^w \mapsto \forall x_1, \ldots, x_n((Iw^*x_1 \wedge \ldots \wedge Iw^*x_n \wedge Ct_1x_1 \wedge \ldots \wedge Ct_nx_n) \rightarrow \phi^{w^*})$$

where $t_1, \ldots, t_n$ are the free terms in ϕ. To illustrate: to say that Humphrey is actually a figure skater is to say that he has an actual figure skating counterpart, or to say that all his actual counterparts are figure skaters, depending on which translation we choose.

The problem with both of these suggestions is that they both translate the patently inconsistent (1) into a consistent formula of counterpart theory.

$$\Diamond\exists x(@Fx \leftrightarrow @\neg Fx) \tag{1}$$

If we adopt the first translation we get $\exists w\exists x(Ixw \wedge (\exists y(Iyw^* \wedge Cxy \wedge Fy) \leftrightarrow \exists y(Iyw^* \wedge Cxy \wedge \neg Fy)))$ and $\exists w\exists x(Ixw \wedge (\forall y(Iyw^* \wedge Cxy \rightarrow Fy) \leftrightarrow \forall y(Iyw^* \wedge Cxy \rightarrow \neg Fy)))$ if we adopt the second. But both these come out true if there is some possible object that has an actual F counterpart and an actual $\neg F$ counterpart or if there is a possible object with no actual counterparts. If Humphrey has no actual counterparts, he has no actual F, nor $\neg F$ counterparts so according to the first suggestion he's actually F if and only if he's actually $\neg F$. Similarly all his actual counterparts are F and are $\neg F$ so according to the second schema he's also actually F if and only if he's actually $\neg F$.

Fara and Williamson generalise these problems in various ways to allow for alternatives to the two natural clauses, and to apply to other versions of counterpart theory. To see the source of the problem, however, this simple example will do for now.

The heart of the problem seems to be that if we think of worlds as regions of spacetime, and we think of a region as representing something a certain way if it contains a counterpart that is that way, then worlds will sometimes represent an object in many different ways (if it contains multiple counterparts of that object), or not represent the object at all (if it contains no counterparts.) Thinking in terms of the Leibnizian biconditionals again, it seems as if the kinds of entities that play the possible world role in those biconditionals must be more fine-grained than Lewisian worlds: a single Lewisian world containing multiple counterparts of Humphrey can serve to represent several possibilities for him. Thus worlds are insufficient if we want to talk about *specific* ways things could be (such as the actual way.)

This issue is irrelevant in Lewis's '68 account because he is only concerned about possibility and necessity: there is a way things could have been for Humphrey in which he is a figure skater precisely if there is a world with a figure skating counterpart of Humphrey. When we are talking about specific ways Humphrey could be, say, the *actual way*, then the issue becomes important. When we are talking about the way things actually are for Humphrey we cannot just be referring to a region of spacetime, which could represent Humphrey in multiple ways. Moreover, no condition of the form 'something (everything) of a certain kind in the actual world is a figure skater' will do either, as Fara and Williamson convincingly show.

It seems that to characterise the *de re* possibilities for Humphrey, worlds must be supplemented with a choice of representative for him. This idea is actually due to Lewis. It doesn't find its way into his formal counterpart theory, but he makes it very clear that in admitting multiple counterparts in a single world we are cutting possibilities finer than worlds:

> "To illustrate, consider these two possibilities for me. I might have been one of a pair of twins. I might have been the first-born one, or the second-born one. These two possibilities involve no qualitative difference in the way the world is. [...] The haecceitist says: two possibilities, two worlds. They *seem* just alike, but they must differ somehow. They represent, *de re*, concerning someone. Hence they must differ with respect to the determinants of the representation *de re*; and these must be non-qualitative, since there are no qualitative differences to be had. I say: two possibilities, sure enough. And they do indeed differ in representation *de re*: according to one I am the first-born twin, according to the other I am the second-born. But they are not two worlds. They are two possibilities within a single world. The world in question contains twin counterparts of me, under a counterpart relation determined by intrinsic and extrinsic qualitative similarities (especially, match of origins.) Each twin is a possible way for a person to be, and in fact is a possible way for me to be. I might have been one, or I might have been the other. There are two distinct possibilities for me. But they involve only one such possibility for the world: it might have been the world inhabited by two such twins." – David Lewis, 'On the Plurality of Worlds', p231

In this passage Lewis wants to reconcile two things. He wants to deny a version of haecceitism that states that there can be qualitatively identical possible worlds which differ only with respect to what *de re* possibilities they represent for some individual, while making sense of the intuitive claim that the first born twin might have switched places with the second born twin while keeping everything else fixed. To do this he introduces possibilities.

While we diverge from Lewis in the details, this is the basic motivation behind our solution. One and the same world can represent an object *de re* in multiple ways. Each of these ways is a possibility. If possibilities are worlds 'plus some extra information', it is this extra information that allows for us to account for the haecceitistic differences between the possibilities, but it is also what accounts for the differences in representation *de re*. The extra information determines how each individual gets represented in the possibility. A natural way to think of a possibility is an ordered pair of a world, and a function taking possibilia to objects existing in the world - each individual is taken to their representative in that world.

To fix ideas, consider Lewis's twins. There are two possibilities within one world; one in which Lewis is the first born, and one on which he is the second born.

1. $\langle w, \sigma \rangle$ where σ is a function taking Lewis to the first born twin. That is, in this possibility Lewis is represented by the first born twin.

2. $\langle w, \sigma' \rangle$ where σ' is a function taking Lewis to the second born twin, but otherwise takes the same values as σ. In this possibility Lewis is represented by the second born twin.

While the difference in σ and σ' comprises a haecceitistic difference in the possibilities, it involves no difference in the world coordinate.

Once we have chosen the actual possibility, it is quite simple to give the truth clause for $@\phi$ with respect to a sequence of individuals and a world. We simply look at the representatives of each object in the sequence at the actual possibility, and ask whether this new sequence of representatives satisfies ϕ at the world coordinate.

So this leaves the crucial question: which possibility is the actual possibility? The non-haecceitistic facts must certainly match the way things actually are, and the actual individuals must certainly be represented by themselves. Thus the actual possibility consists of a pair $\langle w^*, \sigma^* \rangle$ where w^* is the actual world and $\sigma^*(x) = x$ for every actual individual x. This just leaves it open how the non-actual objects get represented. Perhaps there is a canonical way of assigning non actual objects representatives: each possible individual is assigned its best actual representative. I would not hold out hope for such a miraculous coincidence. But even if it were true, it is surely not even epistemically necessary that the the non-actual objects have unique best counterparts to act as their representatives. If b and c are two actual counterparts of a which are on a par I say it is indeterminate whether b or c actually represents a. If it is also the case that b is F and c isn't, I say it is indeterminate whether a is actually F. More generally, if a formula of $\mathsf{QML@}$ is true no matter which possibility we choose as actual, so long as it matches actuality in the non-haecceitistic facts, and represents actual objects as themselves, say the formula is determinately true. If it is false no matter what, call it determinately false. Indeterminacy in the intended model will be rare and harmless - it will only occur when we are able to refer to non actual objects which have several actual counterparts with different properties.

Note that for the standard QML semanticist this distinction does not even arise. For her non-actual objects don't get represented at all at the actual world (after all, you can only represent where you exist, and such individuals don't actually exist.) It is only when you have a counterpart in the actual world distinct from yourself that it makes a difference, and even then, it only makes a difference when there is more than one such counterpart.

2 Refining the notion of 'truth at a world'

The theory sketched above, and given more fully in §3, is not the only way a counterpart theorist might go. There is a quick fix that does not deviate too much from the original counterpart theory of Lewis. The fix relies on a slightly more subtle way of understanding the *truth at* relation that appears in the Leibniz biconditionals. According to the naïve understanding of *truth at*, an object is F at a region of spacetime, for an atomic predicate F, if it is both present there (i.e. it is located in that region) and it is F simpliciter. According to the proposed modification we instead say that something is F at a region if it has an F counterpart there. On this basis we can extend the notion of *truth at a region* compositionally to arbitrary formulae.

In particular, given a predicate F and a Lewis style interpretation of F, X (intuitively the set of possibilia that are F), we may give a more traditional interpretation to F that varies its extension from world to world: The extension of the predicate, F at a world, w, call it $[\![F]\!]_w$, is the set of objects that have a counterpart in X that is also in w. $[\![F]\!]_.$ is exactly the sort of intension you would assign to a monadic predicate in a Kripke model. Starting with a Lewisian metaphysical picture we can in this way construct a perfectly classical Kripke model that delivers intuitively correct truth values to most English sentences (more details can be found in the appendix.) Since we are interpreting QML@ standardly, the logic of @ will be perfectly classical. Indeed it can be shown that validity with respect to the class of models defined in this way is sandwiched between validity with respect to two well studied classes of Kripke models.[7]

However, although the *logic* might be acceptable – after all it is sandwiched between two acceptable notions of validity – there is no guarantee it will assign the right truth conditions. Let me outline an argument that the correct truth conditions will be indeterminate. We will see that the current proposal does not postulate any indeterminacy in the case described.[8] Suppose we have non-actual Ned for whom, in the actual world, there are two perfect candidates to be his counterpart: Ted and Fred. We may assume for the sake of argument that the actual world is a perfect mirror world containing only two objects, Ted and Fred, who are perfect duplicates of one another. Let T and F be the predicates of being Ted and being Fred, and let a be a name for Ned.[9] Given the description of the situation it is natural to think that the following points hold.

1. The semantics should not commit us to the truth of one of $@Ta$ or $@Fa$ without committing us to the other.

2. $@(Ta \vee Fa)$ should come out true on the semantics.

3. $\neg@(Ta \wedge Fa)$ should come out true on the semantics.

Point 1 is true because, by stipulation, both Fred and Ted are equally good candidates. How could Ned possibly be represented by one and not the other; they are perfect and symmetrical duplicates. It would be intolerably arbitrary to assert or deny one of $@Ta$ and $@Fa$ but not the other. Points 2 and 3 are simple. There are only two candidate counterparts for Ned in the actual world so Ned is one or the other, but he cannot be both.

In fact our proposal violates 3: since Ned has Ted and Fred as a counterpart at the actual world he is contained in the extension of both T and F relative to the actual world, so $@(Ta \wedge Fa)$ comes out true. But there is a general problem for any form of counterpart theory that fails to accommodate indeterminacy here. For suppose that the semantics commits us to the determinate truth of $@Ta$ or to the determinate truth of $\neg@Ta$. Either way, assuming 1, the semantics must commit us to $(@Ta \leftrightarrow @Fa)$. The fact that @ commutes with the truth functional connectives, in conjunction with

[7]See definition 3.0.8.

[8]This is a straightforward adaptation of an argument presented by Fara and Williamson.

[9]The use of predicates is just so we do not need to talk about identity.

2. and 3. gives us $(@Ta \leftrightarrow @Fa) \wedge (@Ta \vee @Fa) \wedge \neg(@Ta \wedge @Fa)$, which is an inconsistency in propositional logic.

Someone postulating indeterminacy, on the other hand, may accept 1, 2 and 3. Suppose that there are at least two admissible actualities, one that represents Ned as Ted, and one that represents Ned as Fred. $@Ta$ is true on the first actuality, and $@Fa$ is true on the second actuality. Neither, however, are true on both, so neither is determinately true – they are both borderline. Nonetheless $(@Ta \leftrightarrow @Fa)$ is determinately false – every actuality that makes $@Ta$ true makes $@Fa$ false, and vice versa. The semantics doesn't commit us to $@Ta$ or $@Fa$, as required by 1, but it doesn't commit us to their negations either. Fara and Williamson took this to be a reductio of counterpart theory. However I think a better conclusion to draw is that if counterpart theoretic semantics is to succeed in accommodating an actuality operator it ought to be a semantics which permits indeterminacy about what gets represented by what. It is to such a semantics which we turn to now.[10]

3 A counterpart theoretic semantics for quantified modal logic

In what follows we shall provide a counterpart theoretic semantics for an identity free quantified modal logic with an actuality operator QML@. We then show that validity on this counterpart theoretic semantics is equivalent to validity with respect to the ordinary variable domain Kripke semantics [5].

Our object language, $\mathcal{L}$, consists of the following symbols:

- A denumerable set of variables, $x_1, x_2, \ldots \in Var$
- Predicate symbols of various arity: $P_1^n, P_2^n, \ldots \in Pred^n, n \in \omega$
- Logical connectives $\vee, \neg$

[10] An anonymous referee has pointed out to me that there are ways to tweak the proposal in this section to allow for indeterminacy. For example, by allowing the intended model to consist of several Kripke models based on different counterpart relations as described above. This would do the job within the constraints I have set, however one might wonder if such a proposal could really be called a counterpart theoretic semantics. For it to work the 'counterpart' relations must be sensitive to non-qualitative haecceitistic differences – for example there must be a counterpart relation which relates Ned to Fred but not Ted, and one which relates Ned to Ted but not Fred. The approach I espouse has a better claim to being called a 'counterpart theoretic semantics' as it is a qualitative counterpart relation which determines which possibilities there are. However, I don't mean to suggest that ensuing proposal is forced on us by these considerations.

- Quantifier $\exists$
- Modal operators $\Diamond$, @

The well formed formulae of $\mathcal{L}$ are given as follows:

- If $x_{i_1} \ldots x_{i_n} \in Var$ and $P_i^n \in Pred^n$, then $P_i^n x_{i_1} \ldots x_{i_n} \in Form(\mathcal{L})$
- If $\phi, \psi \in Form(\mathcal{L})$ then $(\phi \vee \psi), \neg\phi, \exists x_i \phi, \Diamond\phi, @\phi \in Form(\mathcal{L})$
- If S satisfies the above conditions, then $Form(\mathcal{L}) \subseteq S$.

Definition 3.0.1. *A* **counterpart structure** *is a quintuple* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^* \rangle$ *satisfying the following conditions:*

1. $\mathcal{W}$ *and* $\mathcal{D}$ *are non-empty.*
2. $\mathcal{C} \subseteq \mathcal{D} \times \mathcal{D}$ *is a reflexive relation.*
3. $Ind : \mathcal{W} \to \mathcal{P}(\mathcal{D})$
4. $w^* \in \mathcal{W}$ *and* $Ind(w^*) \neq \emptyset$

Informally we refer to $\mathcal{W}$ *as the worlds,* $\mathcal{D}$ *the individuals,* $\mathcal{C}$ *the counterpart relation and* w^* *the actual world of the counterpart structure.*

A counterpart structure is essentially a variable domain Kripke structure with the addition of a counterpart relation. Since counterpart structures are simple generalisations of Kripke structures, they are compatible with the same metaphysical hypotheses the Kripke structures can be used to model. In particular, counterpart structures and Kripke structures are both compatible with transworld identity. Unlike Lewis, we allow the worlds domains to overlap.

Notice that allowing the domains to overlap, and treating the counterpart relation as a relation over individuals allows the counterpart relation to be stronger than a qualitative similarity relation. This raises the following worry. Suppose you are an anti-essentialist, and do not want, say, John in this world to be a counterpart of John in a poached egg world, a world where John is a poached egg. Since the the counterpart relation cannot see which world the individual it is considering is at, and it is reflexive, John is always a counterpart of himself. One fix would be to treat the counterpart relation as a four place relation, relating an individual at a world, to an individual at a possibly different world. This maybe conceptually more satisfying, but is a deviation from currently entrenched literature. It is simple to accommodate the anti-essentialist within the above framework by forcing the domains to be disjoint by taking ersatz individuals: individual world pairs. It should not be assumed that the intended interpretation treats $\mathcal{D}$ as a set of ordinary objects.

Definition 3.0.2. *A* **counterpart model** *is a sextuple* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^*, [\![\cdot]\!]^{\cdot} \rangle$ *where* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^* \rangle$ *is a counterpart structure, and*

- $[\![\cdot]\!]^{\cdot} : [Pred^n \rightarrow [\mathcal{W} \rightarrow \mathcal{P}(\mathcal{D}^n)]]$

Intuitively, we may think of $[\![P_i^n]\!] : \mathcal{W} \rightarrow \mathcal{P}(\mathcal{D}^n)$ *as the intension of* P_i^n*, assigning* P_i^n *an extension at each world. If the extension of every predicate at a world is constructed from the objects in the domain of that world, say that the counterpart model is* **serious**. *Stating this constraint more precisely:*

- *For* $w \in \mathcal{W}, P_i^n \in Pred^n$, $[\![P_i^n]\!]^w \subseteq Ind(w)^n$

In what follows we restrict our attention to non-serious counterpart models and non-serious Kripke models. The seriousness constraint on models forces predicates to take their extension at a world from the domain at that world. Intuitively this corresponds to objects only being allowed to instantiate atomic properties at worlds where they exist. Since serious counterpart models and serious Kripke models are special cases of the general models defined here, we ignore the seriousness constraints. Similarly we concentrate on variable domain models, since the fixed domain models are special cases of these.

Definition 3.0.3. *Given a counterpart structure,* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^* \rangle$*, and a world,* $w \in \mathcal{W}$*, we say that* σ *is a* **counterpart function** *for* w *iff* σ *is a function from* $\mathcal{D}$ *into* $\mathcal{D}$ *which is a subset of* $\mathcal{C}$*. We write it as follows:*

- $CF(\sigma, w) \Leftrightarrow \sigma : \mathcal{D} \rightarrow D, \sigma \subseteq C$

Definition 3.0.4. *Given a counterpart structure,* $\mathfrak{A} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^* \rangle$*, we may define the set of* **possibilities***, S, with respect to the structure:*

- $S(\mathfrak{A}) := \{\langle w, \sigma \rangle \mid w \in W, CF(\sigma, w)\}$

Each world is paired with a counterpart function for that world, which provides the extra information concerning how things get represented de re.

Counterpart functions encode ways possibilia might be represented at a world: haecceitistic information, or information about how the possible individuals get represented *de re*. Possibilities are worlds plus haecceitistic information. On some metaphysical views, worlds alone do not contain haecceitistic information, but even if they do, it may not be this information that makes our ordinary modal talk true.

As it stands the world coordinate, w, serves only to restrict the range of the quantifiers at a possibility to objects in that world's domain, and to determine the extension of the atomic predicates.[11] It is consistent with our constraints that there is a possibility, $\langle w, \sigma \rangle$, such that no member of the domain w represents any possibilia at all according to σ. Although I think the current framework has a perfectly legitimate interpretation, one might want to work with a restricted set of possibilities. In particular one might want to stipulate that any possibility based on a world w, must be one in which every member of w's domain represents something. Let us call this condition the 'exhaustion condition', which a pair $\langle w, \sigma \rangle$ satisfies iff

$$\forall y \in Ind(w), \exists x \in \mathcal{D} \text{ such that } \sigma(x) = y.$$

The exhaustion condition ensures that every object in the domain of a possibility represents something according to that possibility. One might want to restrict attention to possibilities which satisfy the exhaustion condition. A stronger condition is the 'identity condition', according to which every individual in the domain of the world coordinate of a possibility represents itself.

$$\forall x \in Ind(w), \sigma(x) = x.$$

As it happens the proofs of theorem 3.1 and 3.2 below will go through if one instead defined $S(\mathfrak{A})$ to be the set of counterpart functions for $\mathfrak{A}$ which satisfy the exhaustion condition, or the identity condition. For the purposes of framework building I shall not take sides on which class of possibilities is philosophically correct.

Presumably among these possibilities there are ones that represent the way things actually are. Although it seems natural for actual objects to be represented by themselves in actual possibilities, typically non-actual objects can be represented in multiple ways. There is nothing particularly special about any one way of representing the individuals, and indeed we should expect there to be multiple ways to represent the non-actual objects all equally compatible with the way we use modal idioms in natural languages. When a sentence is sensitive to the multiple ways of assigning counterparts compatible with facts about English, we should not expect the truth of that sentence to be settled by use facts and the state of the modal universe. This motivates the following definition.

Definition 3.0.5. *An admissible actuality, for a counterpart structure* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^* \rangle$, *and a world,* $w \in \mathcal{W}$ *is a possibility of the form* $\langle w^*, \tau \rangle$ *such that* $\forall x \in Ind(w^*)(\tau(x) = x)$

[11] See the definition of satisfaction at a possibility below.

As usual, we define a valuation to be a function $v : \mathbb{N} \rightarrow \mathcal{D}$. Note that valuations can take their values from individuals which needn't exist in the same world. For valuations v and u, and $n \in \mathbb{N}$ write $v[n]u$ to mean $v(x_m) = u(x_m), \forall m \neq n$. Given a counterpart model $\mathcal{M} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^*, [\![\cdot]\!] \rangle$, and an admissible actuality $s = \langle w^*, \sigma^* \rangle$ for that structure, we define the satisfaction relation with respect to a possibility and a valuation, $\langle \mathcal{M}, s \rangle, \langle w, \sigma, v \rangle \models \phi$, as follows (we omit the model and actuality when there is no possibility of confusion)

$$\langle w, \sigma, v \rangle \models P_i^n x_1, \ldots, x_n \Leftrightarrow \langle \sigma(v(x_1)), \ldots, \sigma(v(x_n)) \rangle \in [\![P_i^n]\!]^w$$
$$\langle w, \sigma, v \rangle \models \neg\phi \Leftrightarrow \langle w, \sigma, v \rangle \not\models \phi$$
$$\langle w, \sigma, v \rangle \models (\phi \wedge \psi) \Leftrightarrow \langle w, \sigma, v \rangle \models \phi \text{ and } \langle w, \sigma, v \rangle \models \psi$$
$$\langle w, \sigma, v \rangle \models @\phi \Leftrightarrow \langle w^*, \sigma^*, v \rangle \models \phi$$
$$\langle w, \sigma, v \rangle \models \Diamond\phi \Leftrightarrow \langle w', \sigma', v \rangle \models \phi \text{ for some } \langle w', \sigma' \rangle \in S$$
$$\langle w, \sigma, v \rangle \models \exists x_i \phi \Leftrightarrow \langle w, \sigma, v' \rangle \models \phi \text{ for some } v'[i]v \text{ such that } v'(x) \in Ind(w)$$

We can then introduce the standard definitions of truth, validity and consequence as follows:

Definition 3.0.6. *Given a counterpart structure $\mathfrak{A} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^* \rangle$, and a model $\mathcal{M} = \langle \mathfrak{A}, [\![\cdot]\!] \rangle$ based on the structure, say that a formula, ϕ is*

- **true** in the model $\mathcal{M} = \langle \mathfrak{A}, [\![\cdot]\!] \rangle$ with respect to an admissible actuality $s = \langle w^*, \sigma^* \rangle$ and a valuation v iff $\langle \mathcal{M}, s \rangle, \langle w^*, \sigma^*, v \rangle \models \phi$.

  valid in $\mathfrak{A}$ iff $\langle \mathcal{M}', s \rangle, \langle w^*, \sigma^*, u \rangle \models \phi$ for every model, $\mathcal{M}' = \langle \mathfrak{A}, [\![\cdot]\!]' \rangle$, based on $\mathfrak{A}$, every valuation u and every admissible actuality $s = \langle w^*, \sigma^* \rangle$.

- **a consequence** of Γ in $\mathfrak{A}$ iff for any model, $\mathcal{M}' = \langle \mathfrak{A}, [\![\cdot]\!]' \rangle$, based on $\mathfrak{A}$, any admissible actuality $s = \langle w^*, \sigma^* \rangle$ and any valuation v, if $\langle \mathcal{M}', s \rangle, \langle w^*, \sigma^*, v \rangle \models \psi, \forall \psi \in \Gamma$, then $\langle \mathcal{M}', s \rangle, \langle w^*, \sigma^*, v \rangle \models \phi$.

We now have a notion of validity for formulae of QML@ based on a counterpart theoretic semantics. How do we know that the semantics we have introduced does not validate the "wrong" formulae? Luckily, it is straightforward to check our semantics gives the correct results with respect to the problematic formulae Fara and Williamson identify. But we can be a bit more general: we can show that, if a standard Kripke semantics for QML@ gets the right results, then so does the counterpart theoretic semantics. Let's start by outlining the Kripke semantics for QML@. I also consider a variation, the 'serious Kripke semantics', for contrast, however our target is the broader notion of validity for Kripke models.

Definition 3.0.7. *A* **Kripke structure** *is a quadruple* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), w^* \rangle$ *satisfying the following conditions:*

1. $\mathcal{W}$ *and* $\mathcal{D}$ *are non-empty.*

2. $Ind : \mathcal{W} \to \mathcal{P}(\mathcal{D})$

3. $w^* \in \mathcal{W}$ *and* $Ind(w^*) \neq \emptyset$

Definition 3.0.8. *A* **Kripke model** *is a quintuple* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), w^*, [\![\cdot]\!]^{\cdot} \rangle$ *where* $\langle \mathcal{W}, \mathcal{D}, Ind(\cdot), w^* \rangle$ *is a Kripke structure, and*

- $[\![\cdot]\!]^{\cdot} : [Pred^n \to [\mathcal{W} \to \mathcal{P}(\mathcal{D}^n)]]$

Intuitively, we may think of $[\![P_i^n]\!] : \mathcal{W} \to \mathcal{P}(\mathcal{D}^n)$ *as the intension of* P_i^n*. Once again we say that a Kripke model is* **serious** *if it satisfies the following condition for every predicate:*

- *For* $w \in \mathcal{W}, P_i^n \in Pred^n$*,* $[\![P_i^n]\!]^w \subseteq Ind(w)^n$

Truth in a Kripke model with respect to a valuation and world, $\mathcal{M}, \langle w, v \rangle \models \phi$ is given as usual (again omit the model when there is no ambiguity):

$$\langle w, v \rangle \models P_i^n x_1, \ldots, x_n \Leftrightarrow \langle v(x_1)), \ldots, v(x_n)) \rangle \in [\![P_i^n]\!]^w$$
$$\langle w, v \rangle \models \neg\phi \Leftrightarrow \langle w, v \rangle \not\models \phi$$
$$\langle w, v \rangle \models (\phi \land \psi) \Leftrightarrow \langle w, v \rangle \models \phi \text{ and } \langle w, v \rangle \models \psi$$
$$\langle w, v \rangle \models @\phi \Leftrightarrow \langle w^*, v \rangle \models \phi$$
$$\langle w, v \rangle \models \Diamond\phi \Leftrightarrow \langle w', v \rangle \models \phi \text{ for some } w' \in \mathcal{W}$$
$$\langle w, v \rangle \models \exists x_i \phi \Leftrightarrow \langle w, v' \rangle \models \phi \text{ for some } v'[i]v \text{ such that } v'(x) \in Ind(w)$$

Definition 3.0.9. *Given a Kripke structure* $\mathfrak{A} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), w^* \rangle$*, and a model* $\langle \mathfrak{A}, [\![\cdot]\!] \rangle$ *based on the structure, say that a formula,* ϕ *is*

- **true** in the model $\mathcal{M} = \langle \mathfrak{A}, [\![\cdot]\!] \rangle$ with respect to a valuation v iff $\mathcal{M}, \langle w^*, v \rangle \models \phi$.

- **valid** in $\mathfrak{A}$ iff $\mathcal{M}', \langle w^*, u \rangle \models \phi$ for every model, $\mathcal{M}' = \langle \mathfrak{A}, [\![\cdot]\!]' \rangle$, based on $\mathfrak{A}$ and every valuation u.

- **a consequence** of Γ in $\mathfrak{A}$ iff for any model, $\mathcal{M}' = \langle \mathfrak{A}, [\![\cdot]\!]' \rangle$, based on $\mathfrak{A}$, and any valuation v, if $\mathcal{M}', \langle w^*, v \rangle \models \psi, \forall\psi \in \Gamma$, then $\mathcal{M}', \langle w^*, v \rangle \models \phi$.

It should be noted that in concentrating on Kripke models, and not serious Kripke models, I have taken sides on a substantial issue. According to a non-serious Kripke model, an individual may satisfy an atomic predicate even if it doesn't exist at that world. If I have the instructions and pieces to make a toy plane, and I am imagining the plane that would have been built if I had followed the instructions, one might want to say that there could have been something I'm actually imagining, namely the toy that would have been built, but which doesn't actually exist. This sentence, formalised as $\Diamond\exists x(@Iax \wedge @\neg\exists yx = y)$, is not satisfiable in any serious Kripke model, but is satisfiable over the wider class of Kripke models.

Our conception of a possibility reflects this choice. A possibilia's representative at a possibility needn't be in the domain at that possibility. Thus, for example, the object representing the possible toy plane at a given possibility may not belong to that possibilities domain.[12]

We are now in a position to compare the notions of Kripke validity and counterpart validity.

Theorem 3.1. *Let* $\mathcal{M} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^*, [\![\cdot]\!]\rangle = \langle \mathfrak{A}, [\![\cdot]\!]\rangle$ *be a counterpart model, and* $s = \langle w^*, \sigma^*\rangle$ *be an admissible actuality for* $\mathfrak{A}$*. Then there is a Kripke model* $\mathcal{M}' = \langle \mathcal{W}', \mathcal{D}', Ind'(\cdot), w^{*\prime}, [\![\cdot]\!]'\rangle$ *such that the following are equivalent for any formula* ϕ*:*

- $\langle \mathcal{M}, s\rangle, \langle w, \sigma, v\rangle \models \phi$ *for every* $\langle w, \sigma\rangle \in S(\mathfrak{A})$ *and every valuation* v
- $\mathcal{M}', \langle w, v\rangle \models \phi$ *for every* $w \in \mathcal{W}'$ *and valuations* v*.*

Proof. Given a counterpart model $\mathcal{M} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), \mathcal{C}, w^*, [\![\cdot]\!]\rangle$ we define a Kripke model $\mathcal{M}' = \langle \mathcal{W}', \mathcal{D}', Ind'(\cdot), w^{*\prime}, [\![\cdot]\!]'\rangle$ as follows:

- $\mathcal{W}' := S(\mathfrak{A})$ the set of possibilities for $\mathfrak{A}$
- $D' := D$
- $Ind'(\langle w, \sigma\rangle) := Ind(w)$
- $w^{*\prime} := \langle w^*, \sigma^*\rangle$
- $[\![P^n_i]\!]'^{\langle w,\sigma\rangle} := \{\langle a_1, \ldots, a_n\rangle \mid \langle \sigma(a_1), \ldots, \sigma(a_n)\rangle \in [\![P^n_i]\!]^w\}$

[12]A corresponding notion of a serious possiblity could be defined as pair $\langle w, \sigma\rangle$ where σ is a *partial* function $\sigma : D \rightharpoonup D$, which is a subset of the counterpart relation, and is such that the range of σ, $\sigma(D)$ is a subset of $Ind(w)$. According to this conception the representative of a possible individual at a possibility always exists at at the possibility in question.

The proof is an easy induction on the complexity of ϕ. Our inductive hypothesis is that for every $\langle w, \sigma \rangle \in S(\mathfrak{A}) = \mathcal{W}'$ and every valuation v: $\langle \mathcal{M}, s \rangle, \langle w, \sigma, v \rangle \models \phi \Leftrightarrow \mathcal{M}', \langle \langle w, \sigma \rangle, v \rangle \models \phi$.

Base case: by the truth clause for counterpart models $\langle \mathcal{M}, s \rangle, \langle w, \sigma, v \rangle \models P_i^n x_1 \ldots x_n$ iff $\langle \sigma(v(x_1)), \ldots, \sigma(v(x_n)) \rangle \in [\![P_i^n]\!]^w$. But by definition of our Kripke model, this holds just in case $\langle v(x_1), \ldots, v(x_n) \rangle \in [\![P_i^n]\!]'^{\langle w, \sigma \rangle}$ which by the atomic truth clause for Kripke models means that $\mathcal{M}', \langle \langle w, \sigma \rangle, v \rangle \models P_i^n x_1 \ldots x_n$.

Inductive step: the cases for $\neg\phi$, $(\phi \vee \psi)$, $\Diamond\phi$ and $@\phi$ are straightforward. The $\exists x_i \phi$ case is worth noting: $\langle \mathcal{M}, s \rangle, \langle w, \sigma, v \rangle \models \exists x_i \phi$ iff $\langle \mathcal{M}, s \rangle, \langle w, \sigma, v' \rangle \models \phi$ for some $v'[i]v$ such that $v'(i) \in Ind(w)$. By inductive hypothesis $\langle \mathcal{M}, s \rangle, \langle w, \sigma, v' \rangle \models \phi$ holds iff $\mathcal{M}', \langle \langle w, \sigma \rangle, v' \rangle \models \phi$. Also $v'(i) \in Ind'(\langle w, \sigma \rangle)$ since $Ind'(\langle w, \sigma \rangle) = Ind(w)$, so by the truth clause for $\exists$, $\mathcal{M}', \langle \langle w, \sigma \rangle, v \rangle \models \exists x_i \phi$. The converse is similar.

Note that in particular $\langle \mathcal{M}, s \rangle, \langle w^*, \sigma^*, v \rangle \models \phi \Leftrightarrow \mathcal{M}', \langle w^{*\prime}, v \rangle \models \phi$. □

Theorem 3.2. *Let $\mathcal{M} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), w^*, [\![\cdot]\!] \rangle = \langle \mathfrak{A}, [\![\cdot]\!] \rangle$ be a Kripke model. Then there is a counterpart model $\mathcal{M}' = \langle \mathcal{W}', \mathcal{D}', Ind'(\cdot), \mathcal{C}, w^{*\prime}, [\![\cdot]\!]' \rangle = \langle \mathfrak{A}', [\![\cdot]\!]' \rangle$ and an admissible actuality, s for that model such that the following are equivalent for any formula ϕ:*

- *$\langle \mathcal{M}', s \rangle, \langle w, \sigma, v \rangle \models \phi$ for every $\langle w, \sigma \rangle \in S(\mathfrak{A}')$ and every valuation v*
- *$\mathcal{M}, \langle w, v \rangle \models \phi$ for every $w \in \mathcal{W}$ and valuations v.*

Proof. Note that the counterpart semantics is just a generalization of the Kripke semantics. Given $\mathcal{M} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), w^*, [\![\cdot]\!] \rangle$ we take our counterpart model simply to be $\mathcal{M} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), =, w^*, [\![\cdot]\!] \rangle$ where the counterpart relation is just the identity relation. In these kinds of models the distinction between worlds and possibilities collapse, and the truth clauses match those for the Kripke semantics.

□

Corollary 3.3. *A formula, ϕ, in the language of* QML@ *is valid in every counterpart structure if and only if it is valid in every Kripke structure.*

Proof. This is a direct consequence of theorems (3.1) and (3.2). □

3.1 Concluding remarks

Fara and Williamson's strategy in [2] is to show that informally inconsistent formulae of QML@ have a consistent interpretation in variants of Lewis's counterpart theory. It is a straightforward consequence of corollary 3.3 that there can be no such argument against the counterpart theoretic semantics for QML@ we have given here. Assuming that no informally inconsistent formula of QML@ is satisfied in a Kripke model, it follows that no inconsistent formula is satisfiable in a counterpart theoretic model.

As an example of this general fact, let us reconsider the formula (1) introduced in section 1:

$$\Diamond\exists x(@Fx \leftrightarrow @\neg Fx) \tag{2}$$

This formula is not satisfied in any counterpart model, relative to any admissible actuality $\langle w^*, \sigma^* \rangle$. For otherwise there would be some possibility, $\langle w, \sigma \rangle$, and some object $a \in Ind(w)$ such that $w, \sigma, v \models @Fx \leftrightarrow @\neg Fx$ where v is any valuation with $v(x) = a$. Following through the satisfaction clauses this would imply that $w^*, \sigma^*, v \models Fx$ if and only if $w^*, \sigma^*, v \not\models Fx$ which is a contradiction.

Indeed, it seems as if there is no logical difference between the counterpart theoretic and Kripke semantics for quantified modal logic. The dispute, at it's heart, is about the correct analysis of modal predication. What I hope to have shown is that the basic tenets of Lewis's analysis of modal predication can be reconciled with the logic, syntax and, perhaps, the ontological innocence of traditional uses of quantified modal logic.

4 Appendix

Here we address some miscellaneous issues raised in the paper.

4.1 Extending the language

So far we have considered only languages without identity. In such languages the Kripke and the counterpart theoretic model theory coincide. We shall see that this result extends to languages with identity, provided we interpret the identity relation in a certain way. However it can be seen that one can also introduce a relation of 'loose identity' which captures the notion of two objects being 'represented as the same'. The inclusion of this relation allows the counterpart theorist to say, for example, that the statue and the lump are only contingently identical. One cannot do this using the ordinary identity relation.

Let us augment QML@ by two binary relations, $=$ and $\simeq$. Thus our language is defined as in section 3, except that we add the extra clauses

- If $x_i, x_j \in Var$ then $x_i = x_j \in Form(L)$ and $x_i \simeq x_j \in Form(L)$.

We adopt the same satisfaction clauses as before with the addition of:

$$\langle w, \sigma, v\rangle \models x_i = x_j \Leftrightarrow v(x_i) = v(x_j)$$
$$\langle w, \sigma, v\rangle \models x_i \simeq x_j \Leftrightarrow \sigma(v(x_i)) = \sigma(v(x_j))$$

a and b are strictly identical at a world iff they are the same object, whereas a and b are loosely identical at a world iff they are represented as the same object there. If a and b are strictly identical they are loosely identical, however the converse may fail.

Logically speaking, the logic of counterpart structures in a language with $=$ and $\simeq$ conservatively extends the logic of Kripke structures in a language with only $=$. Thus $=$ satisfies a completely standard logic of identity. In particular, the formulae containing only strict identity which are valid in every counterpart model are the same as those valid in every Kripke model.[13]

Loose identity, $\simeq$, does not behave like strict identity. For example, loose identity between two objects may be contingent. Similarly, loose identity does not obey Leibniz's law. The inclusion of loose identity is an important feature, since this is the first point at which the Kripke semantics and the counterpart semantics differ in the object language. I leave it to a future project, however, to determine what the logic of $\simeq$ is with respect to these models.

4.2 A non-supervaluationist semantics

According to the variable domain Kripke semantics, the extension of a predicate at a world need not necessarily consist of objects existing at that world. Some modal logicians, the 'serious actualists', prefer to consider a smaller class of models than Kripke did and stipulate that an individual can only have an atomic property at a world at which it exists, restricting, for example, the extension of a monadic predicate at a world to being a subset of the domain of that world. The more general semantics outlined by Kripke, however is compatible with individuals being world bound, while varying their properties from world to world.

An alternative way to give a counterpart theoretic interpretation QML@ involves simply using the Kripke semantics directly. The basic idea is that the extension of a predicate, P, at a world, w, is just the set of objects that have an F counterpart in w (here F is the Lewis style property, a set of possibilia, that interprets P.) Being F-at-a-world is simply having an F counterpart there. This gives us a general method for converting a

[13]This can be seen by a trivial modification of the proofs of theorem 3.1 and 3.2.

Lewis model for first order counterpart theory, into a Kripke model. Suppose we have a first order Lewis style model for counterpart theory, $\langle D, [\![\cdot]\!]^L \rangle$. Let D be the domain, let $[\![\cdot]\!]^L$ be the interpretation function for our model, and let $\mathcal{C} := [\![C]\!]^L$ be the relation that interprets the counterpart relation. The primitive predicates of Lewis's counterpart theory are: W, I, C and @ - the predicates for worldhood, world parthood, counterparthood and the name of the actual world. We assume also that the language contains non counterpart theoretic vocabulary: P_i^n for $n, i \in \omega$. To get a Kripke model, $\langle \mathcal{D}, \mathcal{W}, Ind(\cdot), w^*, [\![\cdot]\!]^K \rangle$, we need a domain, a set of worlds, a domain for each world, an actual world and an interpretation for the non-logical predicates. We obtain them from the Lewis interpretation as follows:

- $\mathcal{D} := D$
- $\mathcal{W} := [\![W]\!]^L$
- $Ind(w) := \{x \in D \mid \langle x, w \rangle \in [\![I]\!]^L\}$
- $w^* := [\![@]\!]^L$
- $[\![P_i^n]\!]^K(w) := \{\langle x_1 \dots x_n \rangle \in D^n \mid \exists y_1 \dots y_n \in Ind(w)(\forall i \leq n, \mathcal{C}x_i y_i \wedge \langle y_1 \dots y_n \rangle \in [\![P_i^n]\!]^L)\}$

What should we expect the logic to look like? Does every Kripke model represent a logically possible way of interpreting the language? Does the counterpart theoretic constraint on the interpretation of the atomic predicates only enter the picture at the intended interpretation? If so, we should expect the logic to be equivalent to the Kripke semantics. However, this is not a particularly interesting response. One might want to know if the logic is still acceptable if we keep the constraint in place across models. The theorem below demonstrates that it is still acceptable. Call the set formulae valid with respect to Kripke models obtained by first order counterpart models **CT**, and call the formulae valid with respect to all Kripke models, and all serious Kripke models **K** and **SK** respectively. We show that $\mathbf{K} \subseteq \mathbf{CT} \subseteq \mathbf{SK}$. Every valid formula according to the Kripke semantics is in **CT**, so **CT** satisfies the minimum requirement of containing formulae that are valid in the widest sense. **CT** does not deem outright inconsistent formulae consistent. However, **CT** does not make too many formulae valid either. For example **CT** does not say that every object exists necessarily, or that every property is had necessarily, if had possibly. For any formula in **CT** is also valid according to the class of serious Kripke models. **SK** provides an upperbound.

Theorem 4.1. $\boldsymbol{K} \subseteq \boldsymbol{CT} \subseteq \boldsymbol{SK}$

Proof. Clearly $\mathbf{K} \subseteq \mathbf{CT}$ since the counterpart models are Kripke models by construction. To show $\mathbf{CT} \subseteq \mathbf{SK}$ we find a counterpart model for each serious Kripke model which makes the same formulae true.

Let $\mathcal{M} = \langle \mathcal{W}, \mathcal{D}, Ind(\cdot), w^*, [\![\cdot]\!] \rangle$ be a serious Kripke model and v an assignment of variables to elements of $\mathcal{D}$. We then construct a first order counterpart model, $\langle D, [\![]\!]^L \rangle$ as follows.

- $D := \{\langle x, w\rangle \in \mathcal{D} \times \mathcal{W} \mid x \in Ind(w)\}$
- $[\![W]\!]^L := \mathcal{W}$
- $[\![I]\!]^L := \{\langle\langle x, w\rangle, w'\rangle \mid w = w'\}$
- $[\![C]\!]^L := \{\langle\langle x, w\rangle, \langle x', w'\rangle\rangle \mid x = x'\}$
- $[\![@]\!]^L := w^*$
- $[\![P_i^n]\!]^L := \{\langle\langle x_1, w\rangle, \ldots, \langle x_n, w\rangle\rangle \mid \langle x_1, \ldots, x_n\rangle \in [\![P_i^n]\!](w)\}$

It is easy to check that this model satisfies Lewis's original axioms of counterpart theory. Now consider the Kripke model induced by this counterpart model: $\mathcal{M}^K = \langle \mathcal{W}^K, \mathcal{D}^K, Ind^K(\cdot), w^{*K}, [\![\cdot]\!]^K \rangle$. Using choice, we may match each individual in the original $\mathcal{D}$ with an individual in $\mathcal{D}^K$; for each $x \in \mathcal{D}$ we pick a world such that x exists in w, and match x with $\langle x, w\rangle \in \mathcal{D}^K$. We convert the variable assignment, v, for $\mathcal{M}$ to a variable assignment, v^K, for $\mathcal{D}^K$ by letting $v^K(x_i)$ take the value corresponding to $v(x_i)$, for each i. Now we must check for each formula, world and valuation, ϕ, w and v, that $\mathcal{M}, w, v \models \phi$ if and only if $\mathcal{M}^K, w, v^K \models \phi$. The proof is a straightforward induction - we begin by checking the atomic formulae. To save time we consider only monadic predicates. Note that $\mathcal{M}^K, w, v^K \models Fx_i$ if and only if $v^K(i) := \langle x, w'\rangle \in [\![F]\!]^K(w)$ which happens if and only if there is a $\langle y, w''\rangle$ such that $I\langle y, w''\rangle w$ holds, $C\langle x, w'\rangle\langle y, w''\rangle$ holds, and $\langle y, w''\rangle \in [\![F]\!]^L$. $I\langle y, w''\rangle w$ holds iff $w'' = w$, $C\langle x, w'\rangle\langle y, w''\rangle$ holds iff $x = y$, so the condition simplifies to $\langle x, w\rangle \in [\![F]\!]^L$. By the construction of the counterpart model, this happens iff $x \in [\![F]\!](w)$, which is the condition for $\mathcal{M}, w, v \models Fx_i$. The truth functional and modal clauses are standard. For the $\exists$ clause note that $\mathcal{M}^K, w, v^K \models \exists v_i \phi$ iff for some $u^K[i]v^K$ such that $u^K(i) \in Ind^K(w)$, $\mathcal{M}^K, w, u^K \models \phi$. Now, there is a variable assignment for $\mathcal{M}$, u, that is matched to u^K in the way described earlier. Thus by inductive hypothesis we get $\mathcal{M}^K, w, u^K \models \phi$ iff $\mathcal{M}, w, u \models \phi$. The way u is matched with u^K ensures that $u(i) \in Ind(w)$ iff $u^K(i) \in Ind^K(w)$ and since $u[i]v$, we get $\mathcal{M}, w, v \models \exists v_i \phi$. This completes the proof.

□

4.3 Extensional counterpart theory.

Possibilists traditionally treat counterpart theory as an extensional first order theory in which one can quantify over the non-actual counterparts of actual objects (see, e.g. [7].) Here we describe a first order language and a translation schema from QML@ to our language which accords with the counterpart theoretic semantics we gave for QML@.

Standardly counterpart theorists will need the two primitive symbols: Iwx and Cxy. I is the relation of being a part of a world, C is the counterpart relation. We shall use the primitive, $Rsxy$, interpreted as x is represented by y in the possibility s (in our previous notation: $x = \sigma(y)$ where $s = \langle w, \sigma \rangle$), and Isx interpreted as x is part of the world w. The predicate As is interpreted as saying that s is an admissible actuality. Variables x, y, z ...range over possibilia, variables s, t, u, v range over possibilities. We reserve one designated variable, s^* which ranges only over admissible actualities. Finally for each non counterpart theoretic n-ary predicate, P, we assign an $n+1$-ary predicate P'. We can give a translation schema of for QML@ as follows:

$$
\begin{aligned}
&(Px_1, \ldots, x_n)^s \mapsto \exists y_1, \ldots, y_n(Rsy_1x_1 \wedge \ldots \wedge Rsy_nx_n \wedge P'y_1 \ldots y_ns)\\
&(\neg\phi)^s \mapsto \neg(\phi^s)\\
&(\phi \wedge \psi)^s \mapsto (\phi^s \wedge \psi^s)\\
&(\exists x\phi)^s \mapsto \exists x(Isx \wedge \phi^s)\\
&(\Diamond\phi)^s \mapsto \exists s'\phi^{s'})\\
&(@\phi)^s \mapsto \phi^{s^*}
\end{aligned}
$$

The translation of a formula, ϕ, of QML@ is given by ϕ^{s^*}. When doing the standard Tarskian model theory for first order languages, we sometimes need a notion of truth in a model for formulae containing free variables. It is traditional to supervaluate: a formula is true if it is true with respect to every assignment to the variables, and false if it is false with respect to every assignment to the variables. Understanding free variables in this way delivers the same results as the supervaluationist semantics presented in §3. The axioms below provide an extensional framework for formulating counterpart theory based on the notion of a possibility:

1. $\forall x \exists s Rsxx$

   Everything is represented by itself in some possibility.

2. $\forall x \exists s Ixs$

   Everything exists in some world.

3. $\forall s \forall x \forall y \forall z ((Rsxy \wedge Rsxz) \rightarrow y = z)$

   An object never has more than one representative in a possibility.

4. $\forall x \forall s (Ixs \rightarrow \exists y Rsxy)$

   If an object is part of a possibility, it represents some object.

5. $\forall x (Ixs^* \rightarrow Rs^*xx)$

   Actual objects represent themselves in the actual possibility.

6. $\exists x Ixs^*$

   There is at least one actual thing.

References

[1] Delia Graff Fara. Dear haecceitism. *Erkenntnis*, 70(3):285–97, 2009.

[2] Michael Fara and Timothy Williamson. Counterparts and actuality. *Mind*, 114:1–30, 2005.

[3] Graeme Forbes. *The Metaphysics of Modality*. Oxford University Press, Oxford, 1985.

[4] Allen Hazen. Counterpart-theoretic semantics for modal logic. *The Journal of Philosophy*, 76:319–38, 1979.

[5] Saul Kripke. Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16:83–94, 1963.

[6] Saul Kripke. *Naming and Necessity*. Harvard University Press, Harvard, 1980.

[7] David Lewis. Counterpart theory and quantified modal logic. In *Philosophical Papers Vol. 1*, pages 26–39. Oxford University Press, 1983.

[8] David Lewis. Postscripts to "counterpart theory and quantified modal logic". In *Philosophical Papers Vol. 1*, pages 39–46. Oxford University Press, 1983.

[9] David K Lewis. *On the plurality of worlds*. Cambridge University Press, 1986.

[10] Ulrich Meyer. Counterpart theory and the actuality operator. *Mind*, 2013.

[11] Murali Ramachandran. An alternative translation scheme for counterpart theory. *Analysis*, 49:131141, 1989.

[12] Adam Rigoni and Richmond H Thomason. The logic of counterpart theory with actuality. *Journal of Philosophical Logic*, Forthcoming.

[13] Jeffrey Sanford Russell. Actuality for counterpart theorists. *Mind*, 2013.