How things are elsewhere
Adventures in counterpart semantics

Wolfgang Schwarz
02 May 2012

1 Introduction

When quantifiers and modal operators mingle, all sorts of troubles arise. Legend has it that after some initial confusion about how to make sense of formulas like $\exists x \Diamond Fx$, the issue was finally settled by Saul Kripke, who put forward what is now known as *Kripke semantics* for quantified modal logic. Formulas like $\exists x \Diamond Fx$ are interpreted by models consisting of some "possible worlds", each equipped with a quantifier domain, and an interpretation function that specifies which individuals satisfy which predicates relative to which worlds. Modal operators function as quantifiers over the worlds, restricted by an "accessibility" relation. $\exists x \Diamond Fx$ is true at a world w iff there is an individual in the domain of w that satisfies F relative to some world accessible from w.

Kripke semantics works well for many applications, but it also has its limitations. Both philosophers and logicians have therefore been toying with alternative, more powerful interpretations. One such alternative is *counterpart semantics*, originally developed by David Lewis in [Lewis 1968]. Models of counterpart semantics are similar to models of Kripke semantics. However, $\exists x \Diamond Fx$ now counts as true at world w iff some individual in the domain of w has a *counterpart* at some accessible world that satisfies F relative to that world.

While the philosophical and model-theoretic virtues of counterpart semantics have seemed appealing in certain quarters, there remains a widespread suspicion that the framework is counter-intuitive, involves implausible metaphysical commitments, and gives rise to an unmanageable and non-standard modal logic. I will try to defend counterpart semantics against these charges. However, what I will defend is rather different from Lewis's original proposal. A general theme of this paper will be that we should dissociate counterpart semantics from various Lewisian doctrines that are commonly lumped together under the heading of "counterpart theory".

Lewis himself regarded his proposal less as an interpretation of quantified modal logic than as a means to dispense with it. He offered his semantics in the form of translation rules from modal logic into the (in his view) superior language of first-order logic with explicit quantifiers over worlds and their inhabitants. Fortunately, our understanding of modal logic has improved a lot since the 1960s, and the contrast between modal logic and first-order logic now looks much more subtle than it must have appeared at the time. In fact, contemporary logicians often treat first-order logic as a branch of propositional modal logic (see e.g. [Blackburn et al. 2001]).

I want to begin by saying a bit more about this contemporary perspective on modal logic. In section 3, I then present a simple counterpart semantics that overcomes some shortcomings of both Kripke semantics and the semantics in [Lewis 1968]. In section 4, I briefly compare my proposal to another venerable alternative to Kripke semantics, *individual concept semantics*. I also explain why counterpart semantics has no untoward implications for the nature of ordinary objects. In section 5, I consider the extension of counterpart semantics to quantified hybrid logic. I end with a brief comment on recent criticisms of counterpart semantics.

2 Relational structures and counterpart models

Let me begin with standard models for *propositional* modal logic. Here, a model here consists of a set W of objects, a relation R on W, and an interpretation function V that assigns to each sentence letter a subset of W. On the traditional alethic interpretation, the objects in W are possible worlds, R is the relation of relative possibility, and the subsets of W are propositions. On a temporal interpretation, W is a set of times, R their temporal order, and the subsets of W are tensed propositions – things that can be true at one time and false at another.

But the applicability of modal logic extends far beyond these traditional examples, especially if we move to multi-modal logics where the single relation R is replaced by a set $\mathbb{R}$ of relations (with arbitrary arity). For example, the objects in W may just as well be people, the subsets of W properties, and the members of $\mathbb{R}$ relations between people. Ultimately, a model for propositional modal logic is just an ordinary first-order model in disguise: W is the domain of quantification, and the subsets and relations on W are the values of predicates and relation symbols.

Unlike first-order logic, modal logic talks about such structures from the perspective of a particular object, which I will call the *centre* of the model.[1] To express that the centre has property p in modal logic, one can simply say 'p'. Every sentence letter on its own is understood as attributing the relevant property to the centre. To talk about other objects, modal operators have to be used. For example, '$\Diamond_R p$' says that some object that is R-related to the centre has property p.

Standard modal logic does not have the resources to explicitly talk about particular objects other than the centre, or to quantify over all objects, irrespective of how they are related to the centre. While this expressive weakness keeps the logic safely decidable, it means that many perhaps important facts remain inexpressible. One way to overcome these limitations, first introduced by Arthur Prior, is to enrich the modal language by special sentence letters $a, b, c, \ldots$, called *nominals*, that express properties true of a single object in W. With the help of a *description operator*, which I will write as a colon, one can then say things like '$a : p$', meaning that the a object has property p. (Another common notation for '$a : p$' is '$@_a p$'.) Nominals function much like names in first-order logic, where one would write 'Pa' instead of '$a : p$'. I will reserve the special nominal '$\mathtt{c}$' for the (singleton of the) centre; '$\mathtt{c}$:' is also known as the *actually operator*.

Following Prior, we might go further and introduce quantifiers into nominal position, allowing for statements like $\forall x \exists y(x : \Box_R y)$. Unsurprisingly, the resulting language has the same expressive power as first-order logic. Indeed, there is a simple translation from this modal language into first-order logic and back that preserves truth in any model (see [Blackburn 2006]).

Given that *propositional* modal logic is an "internalised" version of first-order logic (or a fragment of first-order logic), what is *first-order* modal logic? This depends on the underlying semantics, but it typically emerges as (a fragment of) a two-sorted first-order logic, with an internal perspective on one sort of objects,

1 Adding an explicit centre to models is not strictly necessary, and has fallen a bit out of fashion, but it will simplify some of the discussions in this paper.

and an external perspective on the other. In classical alethic logic, the two sorts of objects are the possible worlds and the individuals inhabiting those worlds. Sentences of quantified modal logic describe the space of possibilities from the internal perspective of a particular world, while they look at individuals from the familiar external perspective of first-order logic.

More generally, in models of quantified modal logic, each *primary object* $w \in W$ is associated with a domain D_w of *secondary objects* that in some sense exist from the perspective of w. Predicate letters stand for properties which secondary objects may have relative to a given primary object. Thus $\Diamond_R \exists x F x$ says that relative to some primary object R-related to the centre, there exists a secondary object which is F. In what follows, I will call primary objects 'worlds' and secondary objects 'individuals'; I trust that these mnemonic labels won't mislead the reader into thinking that primary objects must always be world-like; a "world" may well be a time or a person or a state in a computer program.

Now return to the formula $\exists x \Diamond F x$. As I mentioned above, Kripke semantics counts this as true iff there is an individual in the domain of the centre world such that *this very individual* lies in the extension of F at some world accessible from the centre. By contrast, counterpart semantics treats $\exists x \Diamond F x$ as true iff there is an individual at the centre such that *some counterpart of it* lies in the extension of F at some accessible world. Counterparts thereby serve as representatives of individuals at other worlds.

What does it take for an individual to represent an individual from another world? The simplest answer reads representation as identity: every individual represents itself and nothing else. This version of counterpart semantics coincides with Kripke semantics. Things get interesting if representation is some relation other than identity. For the alethic case, where the members of W are possible worlds, Lewis [1986] argues that representation is a matter of qualitative similarity: your counterparts at other possible worlds are individuals that are sufficiently like you, as you are at the actual world. Stalnaker [1987] agrees that the representation relation is not identity, but argues that it is not a matter of qualitative similarity either. In the temporal case, an individual's representatives at other times might be things that stand in a suitable relation of causal continuity and connectedness to the original object (see [Sider 2001]), but again one can disagree about the details of this relation.

From the perspective of formal model theory, it will be best not to settle on a specific choice of the counterpart relation. Just as any relation on W may be chosen as the accessibility relation, any relation between individuals at worlds may be chosen as the counterpart relation. If no individual exists at more than one world, counterpart relations can be understood simply as binary relations between individuals. Since I don't want to impose the restriction of disjoint domains, I will instead define counterpart relations as binary relations between pairs $\langle w, d \rangle$ of a world w and an individual d. This allows us to say that an individual d at world w is represented by a different individual d' at w', even if d itself happens to exist at w'.

I have mentioned that some applications of modal logic require more than one accessibility relations. Similarly, there are reasons to introduce more than one counterpart relation (see e.g. [Lewis 1986: 252–258], [Ghilardi 2001]). If we have multiple counterpart relations, an individual d' at w' may represent d *in one way*, or *along one path*, while a different individual d'' at the same world represents d in another way or along another path. For simplicity, I will here stick to models with a single counterpart relation and a single accessibility relation.

By fixing representation to be identity, Kripke semantics is slightly simpler than counterpart semantics. Why bother with the extra complication? Different authors have been motivated by different reasons.

For logicians, one attractive feature of counterpart semantics is that it provides a model theory for systems of quantified modal logic for which Kripke semantics breaks down. In the next section, we will see how this comes about. Lewis himself was to a large extent motivated by metaphysical views on the nature of modality: he believed that statements about possibility and necessity should be interpreted on a particular structure of real but spatio-temporally disjoint universes with ordinary individuals as parts. Since an individual that is part of one universe is never part of another, the parts of other universes can at most be counterparts of individuals at our own universe.[2] Other philosophers are often attracted by the deflationary attitude towards

2 Disjointness of domains allows Lewis to define counterparthood directly on individuals, and to drop the world-relativity from the interpretation of (simple) predicates: if an individual at world w is bent, then it is bent at every world where it exists, so we don't have to say that it is bent *relative to w*. This was of some importance to Lewis; see e.g. [Lewis 1986: 199–202, 228].

de re modality allowed by counterpart semantics, or by its track record at solving metaphysical puzzles (see e.g. [Sider 2001]). This is not the place to review these matters. Nevertheless, it will be useful later on to have one or two examples in mind where it seems that the representation relation cannot be identity.

Consider doxastic possibilities. Puzzling Pierre believes that the city he has come to know in England under the name 'London' is a different city to the one known as 'Londres' in France ([Kripke 1979]). The worlds doxastically accessible to Pierre – the worlds compatible with everything he believes – presumably contain two relevant cities, playing different parts of the role that London plays in the actual world. We may want to say that both of them represent London, although they can hardly both be *identical* to London, seeing as they are not identical to one another. Similar problems arise in the temporal setting when individuals fission, or when a time traveler visits her younger self: we then have two individuals at the relevant time with equal claim to being the continuation of a single individual from an earlier time; so "continuation" can't be identity.

There is another, more general motivation for using counterpart semantics, especially with multiple counterpart relations. A counterpart model can represent "trans-world" relations between individuals for which there is no place in a Kripke model. An individual d' at w' might be a cause, or a successor, or a daughter of an individual d at w (think of temporal models, where W is a set of times); such relational facts are not represented anywhere in a Kripke model. In counterpart semantics, they can be taken into account – although we will see in section 5 that one can still not say as much about them as one might perhaps like.

3 Basic counterpart semantics

In this section, I will look at counterpart semantics for basic quantified modal logic, without nominals or nominal quantifiers. Thus our *language* $\mathcal{L}$ is the language of first-order logic with identity, plus a monadic sentence operator $\Box$. I use free variables as individual constants; if you want proper constants, simply declare some of the variables unbindable.[3]

3 The distinction between constants and free variables is unimportant in the semantics below because quantifiers are interpreted by considering variations of the original interpretation

A *counterpart model* for $\mathcal{L}$ consists of a counterpart structure together with an initial interpretation. A *counterpart structure* is a quintuple $\langle W, @, R, D, C\rangle$ such that

1. W is a non-empty set (of "primary objects" or "worlds"),
2. $@$ is a member of W (the "centre"),
3. R is a binary relation on W (the "accessibility" relation),
4. D is a family of sets indexed by worlds, i.e. a function that assigns to each $w \in W$ a set D_w (of "secondary objects" or "individuals"),
5. C is a binary relation on $\{\langle w, d\rangle : w \in W, d \in D_w\}$ (the "counterpart" relation).

An *initial interpretation* on such a structure is a function V such that

1. for each $w \in W$ and non-logical predicate P^n, $V_w(P^n) \subseteq D_w^n$,
2. for each $w \in W$, $V_w(=) = \{\langle d, d\rangle : d \in D_w\}$, and
3. for each individual variable x, $V(x)$ is either a member of $D_@$ or undefined.

To complete the semantics, we have to specify under what conditions a formula of $\mathcal{L}$ is true in a counterpart model. The following definition corresponds to the translation rules in [Lewis 1968].

Let $\mathcal{M}$ be a counterpart model consisting of a structure $\mathcal{S} = \langle W, @, R, D, C\rangle$ and an interpretation V. A formula A is *true* in $\mathcal{M}$ iff $\mathcal{S}, V, @ \Vdash A$, where $\Vdash$ is defined as follows.

$\mathcal{S}, V, w \Vdash P^n x_1 \dots x_n$	iff $\langle V(x_1), \dots, V(x_n)\rangle \in V_w(P^n)$.
$\mathcal{S}, V, w \Vdash \neg A$	iff $\mathcal{S}, V, w \nVdash A$.
$\mathcal{S}, V, w \Vdash A \wedge B$	iff $\mathcal{S}, V, w \Vdash A$ and $\mathcal{S}, V, w \Vdash B$.
$\mathcal{S}, V, w \Vdash \forall x A$	iff $\mathcal{S}, V', w \Vdash A$ for all x-variants V' of V with $V'(x) \in D_w$.
$\mathcal{S}, V, w \Vdash \Box A(x_1, \dots, x_n)$	iff $\mathcal{S}, V', w' \Vdash A(x_1 \dots, x_n)$ for all w', V' such that wRw' and V' is an $x_1, \dots, x_n$-variant of V with $\langle w, V(x_i)\rangle C \langle w', V'(x_i)\rangle$.

All of this is just as in Kripke semantics, except for the last clause, concerning the box. Here $A(x_1, \dots, x_n)$ is a formula whose free variables are $x_1, \dots, x_n$.

function, which makes an extra assignment function redundant; see e.g. [Bostock 1997: 81–90] for discussion.

So $\Box A$ is true relative to a world w and interpretation V iff A is true relative to all w-accessible worlds w' and interpretations V' that assign to each free variable x in A a counterpart (at w') of the individual (at w) assigned to x by V.

For example, $\Box Fx$ says that all counterparts of x at all accessible worlds are F. It is not required that x has a counterpart at all accessible worlds. In general, worlds where the relevant individuals lack counterparts are ignored in the truth-conditions for the box. Hence $\Box(Fx \wedge Gy)$ does not entail $\Box Fx$: if all y-counterparts are G and the only x-counterparts that are not F inhabit worlds where y has no counterpart, then $\Box(Fx \wedge Gy)$ is true and $\Box Fx$ false. Lewis's semantics therefore fails to validate basic principles of modal logic like[4]

(K) $\quad \Box(A \supset B) \supset (\Box A \supset \Box B)$.

On the other hand, as Lewis notes, the rather controversial "necessity of existence" postulate

(NE) $\quad \Box\exists x(x = y)$

comes out valid, because the box only quantifies over worlds where y has a counterpart.

There is a system behind these oddities. Lewis's logic takes a different internal perspective on relational structures than standard modal logic. It looks at a structure not from the perspective of a particular world, but from the perspective of a particular *individual* at a world. The modality in $\Box Fx$ ranges not over worlds accessible from the present world, but over individuals accessible from x – in the alethic case, over alternative ways x might have been (see [Lewis 1986: 230-235]).

Consider a possible world where history keeps repeating itself. This might represent one alternative for the world, but many alternatives for *me*: I might live in the first epoch, or in the second, or third. In the other direction, a possible world in which there is nothing but empty spacetime does not represent any alternative for me: no way I might have been is compossible with the world

4 This is pointed out in [Hazen 1979] and [Woollaston 1994]. The S1 axiom $\Box(A \supset B) \supset (\Box(B \supset C) \supset \Box(A \supset C))$ also fails. Strictly speaking, all closed formulas of these forms are valid in the semantics of [Lewis 1968], because Lewis's modal language does not have individual constants.

being like that. Such worlds will therefore be ignored when we look at ways I might have been.

In a logic of individual possibility, truth is relative not to worlds, but to individuals, or sequences of individuals; the counterpart relation replaces the accessibility relation. This “individualistic” interpretation of quantified modal logic has been rigorously developed by Silvio Ghilardi, Giancarlo Meloni and Giovanna Corsi (see especially [Ghilardi and Meloni 1991], [Corsi 2002], [Braüner and Ghilardi 2006: 591–616]). The result is a beautiful and powerful framework, but it is not a framework for classical, world-centred modal logic.

Note that counterpart semantics and individualism are independent. Centring on individuals does not require tracking individuals by a counterpart relation. Individualistic interpretations for quantified modal logic have been proposed in otherwise standard Kripke semantics (e.g. in [van Benthem 1983]). Conversely, the idea that individuals should be tracked by a counterpart relation does not entail that modality should be centred on individuals rather than worlds. So here is another opportunity for separating inessential ingredients from Lewis's original proposal. Since my focus in this paper is on counterpart semantics and not on individualism, I will henceforth develop the more standard “worldly” approach.

The clause for the box should then be revised. When evaluating $\Box Fx$, we have to take into account all accessible worlds, including worlds where x has no counterpart. Evaluated at those worlds, ‘x’ behaves like an empty term. We have essentially the same options that arise in the semantics of free logic. We could say that if ‘x’ is empty at a world, then every atomic predication Fx is false at that world, and $\neg Fx$ true. This would be a *negative interpretation*. Alternatively, we could say that non-existence is no bar to satisfying atomic predicates, so that Fx may still be trued at worlds where x has no counterpart. This would be a *positive* interpretation. Finally, we could choose a *nonvalent* interpretation on which neither Fx nor $\neg Fx$ is true if ‘x’ is empty.

Kripke semantics faces the same choice if individuals may fail to exist at some world. All three options can be found in the literature, with most authors opting for either positive or negative interpretations (see [Garson 1984]). The semantic definitions above most naturally extend to a negative interpretation: since $V_w(F)$ is a subset of D_w, Fx is false at w whenever $V(x) \notin D_w$, and then $\neg Fx$ is true. To reach a positive semantics, the domain of each world would have to be split into an *inner* domain that provides the world-relative domain of quantification, and an *outer* domain of things that don't exist at the world, but can nevertheless satisfy predicates. One should then stipulate that if something has no counterpart in the inner domain of a world, it must nevertheless have a counterpart in the outer domain, so that terms can never go genuinely empty. Positive interpretations roughly along these lines are presented in [Forbes 1982], [Kracht and Kutz 2002] and [Schwarz 2011]. A nonvalent semantics will be sketched in section 5 below.

We still need a new clause for the box. The guiding thought is the same as before. When we evaluate $\Box Fx$ relative to a world w and interpretation V, we consider the truth-value of Fx at all w-accessible worlds w' relative to variations V' of V that assign to 'x' a counterpart at w' of the original value of 'x'. If there is no such counterpart, then $V'(x)$ is undefined. To fix some useful terminology, call an interpretation function V' *a w'-image of V at w* iff V' and V agree on all predicates, and for every variable x, if there are individuals d with $\langle w, V(x)\rangle C \langle w', d\rangle$, then $V'(x)$ is one of these individuals, otherwise $V'(x)$ is undefined. Then

$\mathcal{S}, V, w \Vdash \Box A$ iff $\mathcal{S}, V', w' \Vdash A$ for all w', V' such that wRw' and V' is a w'-image of V at w.

The rest stays as before.

It is easy to check that on this interpretation, $\Box(Fx \wedge Gy)$ entails $\Box Fx$. If $\Box(Fx \wedge Gy)$ is true relative to some world w and interpretation V, then $Fx \wedge Gy$ is true relative to all w-accessible worlds w' and w'-images V' of V at w. By the clause for $\wedge$, it follows that both Fx and Gy are true relative to all such w', V'. So $\Box Fx$ is true relative to w and V. We no longer have a counterexample to (K). Moreover, (NE) has become invalid: if y has no counterpart at some accessible world, then $\Box \exists x(x = y)$ is false.

More generally, it can be shown that all principles of the minimal normal modal logic K are valid in this semantics, as are all principles of negative free logic N (with one minor caveat). To get a complete logic, the following principles must be added to the standard axioms and rules of K and N:

(1) $x \neq x \supset \Box x \neq x$.

(2) $x = y \supset \Box(x = x \supset y = y)$.

In negative semantics, $x = x$ is equivalent to $\exists y(x = y)$. So (2) says that if x is identical to y, and x has a counterpart at some world, then y also has a counterpart at that world. This would be invalid if different names for the same individual could be associated with different counterpart relations, but our current semantics does not allow for that. (1) says that if x doesn't exist, then x doesn't have any counterparts. This should not be confused with the claim that no individual exists at any world that isn't a counterpart of something at the centre. The point of (1) is merely that no such alien individual is denoted by a name x. To rule out aliens altogether would require something like the Barcan Formula, $\forall x \Box A \supset \Box \forall x A$. This isn't valid. For example, if $W = \{w, w'\}$, $@ = w$, wRw', $D_w = \emptyset$ and $D_{w'} = \{0\}$, then $\forall x \Box x \neq x$ is true and $\Box \forall x(x \neq x)$ false.

Now for the minor caveat. Some care is required when applying principles from first-order logic that involve substitution. For example, not all instances of

(LL) $\quad x = y \supset (A \supset A[y/x])$

are valid. This much is true already in standard first-order logic, where e.g.

(3) $\quad x = y \supset (\exists y(x \neq y) \supset \exists y(y \neq y))$

does not count as a legitimate instance of Leibniz' Law. Intuitively, the problem is that (LL) only applies if $A[y/x]$ says about y what A says about x. And while $\exists y(x \neq y)$ says that there are individuals other than x, $\exists y(y \neq y)$ does not say that there are individuals other than y. We can make this more precise: $A[y/x]$ says about y what A says about x iff for any interpretation V and world w in any structure $\mathcal{S}$: $\mathcal{S}, V, w \Vdash A[y/x]$ iff $\mathcal{S}, V^{[y/x]}, w \Vdash A$, where $V^{[y/x]}$ is like V except that $V^{[y/x]}(x) = V(y)$.

In first-order logic, $A[y/x]$ says about y what A says about x whenever y is *free for* x, i.e. whenever x does not occur in the scope of an operator that binds y. In counterpart logics, we need an analogous restriction for modal operators, because these effectively re-bind all variables in their scope (as pointed out in [Lewis 1983]). Thus

(4) $\quad x = y \supset (\Diamond x \neq y \supset \Diamond y \neq y)$

is not a legitimate instance of Leibniz' Law on the present semantics: if $x = y$, then $\Diamond x \neq y$ is true whenever the individual picked out by 'x' (and 'y') has multiple counterparts at some accessible world. However, $\Diamond y \neq y$ is only true if that individual has no counterpart at some accessible world. (In section 5, I will consider an alternative reading of $\Diamond y \neq y$ on which (4) is valid.)

The appropriate syntactical restriction on substitution principles like (LL) is not entirely obvious, but the following turns out to work. Call a variable y *modally free for x in A* iff

(i) no modal operator in A has free occurrences of both x and y in its scope, *or*

(ii) A has the form $\Box B$ and y is modally free for x in B.

The definition is recursive. Clause (ii) reflects the fact that (4) is valid if the diamond is replaced with a box or a string of boxes. A proof that the resulting logic is indeed sound and complete on the class of all counterpart models can be found in [Schwarz 2011]. I also show that the complete logic for positive models is simply the combination of positive free logic and K, again with substitution principles restricted by the condition of modal freedom.

These are very weak logics. As a result, counterpart semantics can serve as a model theory for a large class of modal logics for which Kripke semantics becomes unusable. For example, consider the perhaps simplest quantified modal logic QK: the combination of K with standard (non-free) first-order logic Q. This logic is not characterised by any class of Kripke models, because every Kripke model validates

(5) $\quad x \neq y \supset \Box x \neq y,$

which is not provable in QK. In counterpart semantics, QK is characterised by the class of *total functional* structures, i.e. the structures in which every individual has exactly one counterpart at every other world.

4 Counterparts and intensional objects

Before I turn to counterpart semantics for enriched modal languages, I want to pause for a moment to clarify what counterpart semantics entails about the nature of ordinary objects when used in alethic or temporal logic. Recall that for Lewis, ordinary objects like you and I exist only at a single possible world. Similarly, Ted Sider holds as part of his temporal counterpart theory that you and I exist only at a single time (see [Sider 2001]). However, in counterpart semantics, if an individual x at the centre has a counterpart at some other world where p is true, then $\Diamond(p \wedge \exists y(y = x))$ is true. In words: there is an accessible world where p holds and x exists. This seems to entail that x exists at other worlds. So it is unfortunate if one has to add that in fact, ordinary objects exist at no world other than the centre. Philosophers do sometimes distinguish the strict and literal truth from what one can get away with when speaking with the vulgar. Still, the mismatch between "object language" and "meta-language" in the Lewis-Sider position is unattractive, and has caused many philosophers to shy away from counterpart semantics.

Here is a remedy. Start with a model of the Lewis-Sider type. For each individual d that exists at a world w, define the *graph* $\Gamma(w, d)$ *induced at* w *by* d as the pair $\langle w, d\rangle$ together with all its counterparts from other worlds. That is,

$$\Gamma(w, d) =_{def} \{\langle w, d\rangle\} \cup \{\langle w', d'\rangle : \langle w, d\rangle C \langle w', d'\rangle \text{ and } w' \neq w\}.$$

Thus whenever some counterpart of Hubert Humphrey wins an election at a world w', then the graph of his present stage contains a pair $\langle w', d'\rangle$ such that d' wins an election at w'.[5]

Next, introduce a new class O of *intensional objects* (for short: *objects*) into counterpart models, together with new variables $\xi_1, \xi_2, \ldots$ that range over objects. Each object is uniquely associated with a graph. To keep things simple, let's assume that objects *are* graphs.

We need a clause for interpreting predications with object terms. To this end, define the *trace* $\tau(o, w)$ of object o at world w as the unique individual $d \in D_w$ such that $\langle w, d\rangle \in o$. If a graph has no member for world w, or if it has more than one, then the trace at w is undefined. For individuals d, let $\tau(d, w)$ simply be d. Note that if $o = \Gamma(w, d)$, then $\tau(o, w) = d$.

The new clause now says that $F\xi$ is true relative to a world w and an interpretation V iff the trace of the object denoted by ξ at w is in the extension of F at w. In general, for arbitrary terms $x_1, \ldots, x_n$,

$$\mathcal{S}, V, w \Vdash P^n x_1 \ldots x_n \text{ iff } \langle \tau(V(x_1), w), \ldots, \tau(V(x_n), w)\rangle \in V_w(P^n).$$

5 If an individual d at w has another individual d' at w as counterpart, then $\langle w, d'\rangle$ is not part of $\Gamma(w, d)$. However, there is then a different graph $\Gamma(w, d')$ which, as we will see, can be regarded as a counterpart of $\Gamma(w, d)$.

Finally, we adjust what it takes for one interpretation function to be an image of another. Recall that so far, V' is a w'-image of V at w iff (i) V' and V agree on all predicates, and (ii) for any individual variable x, if there is an individual d such that $\langle w, V(x)\rangle C \langle w', d\rangle$ then $V'(x)$ is some such d, otherwise it is undefined. This says nothing about object variables. So we add: (iii) for any object variable ξ, if there is an individual d such that $\langle w, \tau(V(\xi), w)\rangle C \langle w', d\rangle$ then $V'(\xi)$ is $\Gamma(w', d)$ for some such d, otherwise it is undefined.

The idea is as follows. Suppose ξ denotes an intensional object o, whose trace at the present world w is d. When we move the point of evaluation to another world w', we first look for a counterpart d' of d at w' and then let ξ denote some object whose graph is induced by d' at w'. The counterpart relation between individuals thereby determines a counterpart relation between objects: object o' at w' is a counterpart of object o at w iff the trace of o' at w' is a counterpart of the trace of o at w.

Call the result of all these changes *intensional counterpart semantics*. It is a semantics for a two-sorted language, with terms for both individuals and intensional objects. We can go one step further by assuming that all terms denote intensional objects. After all, reference to an individual d can always be mimicked by referring instead to an object whose graph is the set of $\langle w, d\rangle$ pairs (for all worlds w where d exists). And then we might as well use latin variables $x_1, x_2, \ldots$ instead of the greek $\xi_1, \xi_2, \ldots$. The resulting pure intensional counterpart semantics turns out be nothing other than simple, non-intensional counterpart semantics.

To see why, take any model $\mathcal{M}^i$ of intensional counterpart semantics, and define a corresponding counterpart model $\mathcal{M}^c$ as follows. $W^c, @^c, R^c$ are as in $\mathcal{M}^i$; for each world w, let D^c_w be the set of objects in $\mathcal{M}^i$ induced at w by an individual in D^i_w; C^c holds between $\langle w, o\rangle$ and $\langle w', o'\rangle$ iff $\langle w, \tau(o, w)\rangle C^i \langle w', \tau(o', w')\rangle$; V^c is like V^i except that $V^c_w(P^n)$ is the set of object tuples $\langle o_1, \ldots, o_n\rangle$ for which the corresponding tuple of traces at w is in $V^i_w(P^n)$. A simple induction shows that $\mathcal{M}^i$ and $\mathcal{M}^c$ verify exactly the same sentences.[6]

The upshot is that counterpart semantics doesn't require individuals to be world-bound slices or stages. On the account just outlined, 'Hubert Humphrey'

6 My intensional counterpart semantics is inspired by [Schurz 1997: 217–222], [Fitting 2004] and [Kracht and Kutz 2005]. The collapse into non-intensional semantics resembles the main storyline in [Hughes and Cresswell 1996: ch.18].

denotes an individual which (strictly and literally) exists at many worlds – at all worlds about which one can truly say that Humphrey exists there. Nevertheless, what matters for the truth of ‘possibly (or: at some time), Humphrey won the election’ are the properties of his *counterparts*. In easy cases, those counterparts will simply be Humphrey himself. But in strange cases, when it comes to fission or fusion or time travel, the counterparts will not be Humphrey, but other trans-world objects that merely share various traces with Humphrey.

Intensional counterpart semantics resembles Carnap’s semantics of *individual concepts* (see [Carnap 1947]).[7] The main difference between the two is that in concept semantics, the denotation of object variables remains constant under imaging. That is, if ξ denotes an object o, and we want to evaluate $F\xi$ at another world w', then we simply check whether the w'-trace of o is in the extension of F at w'. This presupposes that the graph of each object is a function, which is indeed stipulated in concept semantics: an individual concept is simply an object whose graph is a function.

Models of individual concept semantics can be understood as a restricted class of intensional counterpart models (with multiple counterpart relations). The two accounts diverge when an individual has several counterparts at some accessible world. More importantly, the conceptual account breaks down when the counterpart relation is asymmetrical or when individuals can have other individuals at their own world as counterparts, as it can happen in cases of time travel.

5 Quantified hybrid logic

I mentioned in section 2 that the expressive power of modal logic can be strengthened by adding nominals for talking about particular worlds. In the resulting language of *quantified hybrid logic*, one can for example say that something at the centre world is F at world a – $\exists x(a\colon Fx)$, or that something at world a is F at the centre – $a\colon \exists x(\mathsf{c}\colon Fx)$. Extending counterpart semantics to formulas like these will bring to light a problem that I have swept under the carpet when the only modal operators were boxes and diamonds.

7 Like in concept semantics, we might spice up the logic by re-interpreting identity so that $x = y$ says that x and y have the same trace at the relevant world.

The language $\mathcal{L}_h$ of quantified hybrid logic is just like $\mathcal{L}$ with the addition of some special zero-ary predicates $a, b, c, \ldots$ (the nominals) and the colon operator : that forms a sentence by taking a nominal to its left and an arbitrary sentence to its right. Under any interpretation function V, the nominals must be interpreted to be true at exactly one world; for convenience, I will write $V(a)$ for the world at which a is true. In standard hybrid logic, the rule for the colon is that $a\colon A$ is true at any world iff A is true at the a world. That is,

$\mathcal{S}, V, w \Vdash a\colon A$ iff $\mathcal{S}, V, V(a) \Vdash A$.

In counterpart semantics, this interpretation is not very useful, as it doesn't properly track individuals by the counterpart relation. Roughly, we want to say that $\mathcal{S}, V, w \Vdash a\colon A$ iff $\mathcal{S}, V', V(a) \Vdash A$, where V' is the $V(a)$-image of V at w. But what if there is more than one such image?

This is the problem from under the carpet: what should we say about an individual if it has multiple counterparts at the world under consideration? Suppose, for concreteness, that at some future time t, Alice the time traveler is about to wake up her younger self. Should we say that at this point, Alice is asleep? That she is awake? That she is both asleep and awake? Neither asleep nor awake?

In section 3, the question could be avoided because it didn't really matter what is true of an individual at any particular world other than the centre. By the semantics presented there, $\Box Fx$ requires that *all* counterparts at *all* accessible worlds be F, and so $\Diamond Fx$ (i.e. $\neg\Box\neg Fx$) requires that *some* counterpart at *some* world be F. This reading of the box and diamond is natural in an individualistic setting, where the modal operator in $\Box Fx$ quantifies not over alternatives for the world, but over alternatives for x. Things are less clear-cut if we assume the perspective of a world: if $\Diamond Fx$ says that some accessible world is such that some x counterpart is F there, shouldn't $\Box Fx$ say that all accessible worlds are like this? With the addition of nominals, the question becomes more pressing: if $a\colon Fx$ says that the a world is so-and-so (whatever that is), shouldn't $\Box Fx$ say that all accessible worlds are so-and-so?

At any rate, even if we stick to the old clause for the box, we have a problem with the colon operator. If x has a unique counterpart at the a world, then $a\colon Fx$ should be true iff that counterpart is F. If x has no counterpart at the a world, we have the three options for empty terms discussed in section 3. Things become complicated if x has several counterparts at the a world. When we shift the point of evaluation to the a world, the term 'x' becomes ambiguous in a sense: it denotes several individuals at once.

One option then is to follow the old clause for the box and require all counterparts at the a world to be F. That is,

$$\mathcal{S}, V, w \Vdash a\colon A \text{ iff } \mathcal{S}, V', V(a) \Vdash A \text{ for all } V(a)\text{-images } V' \text{ of } V \text{ at } w.$$

This might be called a *supervaluationist* interpretation, mirroring supervaluationism in the theory of ambiguity and vagueness. Alternatively, one might give a *subvaluationist* account, requiring only that A be true on *some* image V'.[8]

A characteristic feature of both 'valuationisms is that connectives and quantifiers no longer behave classically. For instance, a disjunction $A \vee B$ can be true on all disambiguations, while neither A nor B is true on all disambiguations (try $B = \neg A$). Likewise, A and B may both be true on some disambiguation, while $A \wedge B$ is not true on any disambiguation. In the present context, these facts show up as the invalidity of either $a : (A \vee B) \supset (a : A \;\vee\; a : B)$ or $a\colon A \wedge a\colon B \supset a\colon (A \wedge B)$. Either account therefore invalidates the distribution principle

(HK) $\quad a\colon (A \supset B) \supset (a\colon A \supset a\colon B)$,

which is one of the standard axioms in hybrid logic (see [Blackburn et al. 2001: ch.7]). If $\Box A$ is understood as true iff $a : A$ is true on any assignment of accessible worlds to 'a', the corresponding distribution principle (K) for the box fails as well.

Moreover, if x has multiple counterparts at a only one of which is F, then $a\colon Fx$ and $\neg a\colon \neg Fx$ are both true on the subvaluationist account, and both false on the supervaluationist version. So the standard principle of self-duality also has to go:

(HD) $\quad a\colon A \leftrightarrow \neg a\colon \neg A$.

Aside from such logical peculiarities, the 'valuationist accounts have the drawback that they ignore mixed disambiguations on which different occurrences

8 [Lewis 1973: 37–43] suggests a supervaluationist interpretation for the counterfactual operator $\Box\!\!\rightarrow$, which may be understood as a generalisation of the colon that takes arbitrary propositions on either side (see also [Lewis 1973: 111–117]).

of a term may denote different individuals. For example, 'at t, Alice wakes up Alice' would be true on a mixed interpretation that maps the first occurrence of 'Alice' to the older counterpart, and the second to the younger.

To allow for such mixed interpretations, it proves useful to redefine $\Vdash$ so that it links a formula to a world w and a *set* $\mathbb{V}$ of interpretation functions. $\mathbb{V}$ contains all "disambiguations" of the language relative to w. The clause for the colon becomes

$\mathcal{S}, \mathbb{V}, w \Vdash a\colon A$ iff $\mathcal{S}, \mathbb{V}', V(a) \Vdash A$, where V is an arbitrary member of $\mathbb{V}$ and $\mathbb{V}'$ is the set of $V(a)$-images of members of $\mathbb{V}$ at w.

A sentence A is true in a model $\langle \mathcal{S}, V \rangle$ iff $\mathcal{S}, \{V\}, @ \Vdash A$, where @ is the centre of $\mathcal{S}$. So $a\colon A$ is true in a model iff A is true relative to the a world and the set of interpretation functions that are a-images of the original interpretation function V. To recover the 'valuationist accounts from above, we would add that for any sentence A,

$\mathcal{S}, \mathbb{V}, w \Vdash A$ iff $\mathcal{S}, V, w \Vdash A$ for all (some) $V \in \mathbb{V}$.

But we can now also state separately for each semantic operation what it should do when its arguments are ambiguous. For example, one might confine the 'valuationist treatment to atomic formulas, as follows.

$\mathcal{S}, \mathbb{V}, w \Vdash P^n x_1 \ldots x_n$	iff $\langle V(x_1), \ldots, V(x_n) \rangle \in V_w(P^n)$ for all (some) $V \in \mathbb{V}$.
$\mathcal{S}, \mathbb{V}, w \Vdash \neg A$	iff $\mathcal{S}, \mathbb{V}, w \nVdash A$.
$\mathcal{S}, \mathbb{V}, w \Vdash A \land B$	iff $\mathcal{S}, \mathbb{V}, w \Vdash A$ and $\mathcal{S}, \mathbb{V}, w \Vdash B$.
$\mathcal{S}, \mathbb{V}, w \Vdash \forall x A$	iff $\mathcal{S}, \mathbb{V}', w \Vdash A$ for all $\mathbb{V}'$ such that for some $d \in D_w$, every $V' \in \mathbb{V}'$ is an x-variant of some $V \in \mathbb{V}$ with $V'(x) = d$.
$\mathcal{S}, \mathbb{V}, w \Vdash a\colon A$	iff $\mathcal{S}, \mathbb{V}', V(a) \Vdash A$, where V is any member of $\mathbb{V}$ and $\mathbb{V}'$ is the set of $V(a)$-images of members of $\mathbb{V}$ at w.
$\mathcal{S}, \mathbb{V}, w \Vdash \Box A$	iff $\mathcal{S}, \mathbb{V}', w' \Vdash A$ for all $w', \mathbb{V}'$ where wRw' and $\mathbb{V}'$ is the set of w'-images of members of $\mathbb{V}$ at w.

The set $\mathbb{V}$ of disambiguations is unfolded only in the first clause. Connectives and quantifiers behave classically. (HK) and (HD) are valid in this semantics.

However, uniformity is still enforced within each atomic sentence. To get a reading on which 'at t, Alice wakes up Alice' can be true, the clause for atomic formulas has to be replaced so as to confine the resolution of ambiguity even further:[9]

$\mathcal{S}, \mathbb{V}, w \Vdash P^n x_1 \ldots x_n$ iff $\langle V_1(x_1), \ldots, V_n(x_n) \rangle \in V_w(P^n)$ for all (some) $V, V_1, \ldots, V_n \in \mathbb{V}$.

Now that mixed disambiguations are in play, the hybrid "necessitation" rule

(HN) $\vdash A \quad \Rightarrow \quad \vdash a\colon A$

becomes invalid, as does the standard necessitation rule for the box. For example, $\forall y Gyy \wedge x{=}x \supset Gxx$ and $\forall y \neg Gyy \wedge x{=}x \supset \neg Gxx$ are valid, but either $a\colon (\forall y Gyy \wedge x{=}x \supset Gxx)$ or $a\colon (\forall y \neg Gyy \wedge x{=}x \supset \neg Gxx)$ is invalid on the present semantics: nothing can be two meters away from itself, and yet at t, Alice is two meters away from Alice!

Finally, rather than confining ambiguity, we might let it spread: if 'x' is ambiguous, then 'Fx' is also ambiguous, and so are '$\neg Fx$' and '$Fx \wedge Gy$'. Any sentence may then be true on some disambiguation and false on another. Unlike in 'valuationist treatments, the relevant disambiguations need not be uniform. The basic semantic values on this account are not *true* and *false*, but *true-on-some-disambiguation* (for short: 1) and *false-on-some-disambiguation* (for short: 0). Since a sentence can have both value 0 and 1, it is not enough for a semantics to state under which conditions a sentence has value 1. We must also state under which conditions it has 0.

This approach goes nicely with a nonvalent treatment of empty terms: if x has a no counterpart at world a, then $a\colon Fx$ is true on no (mixed) choice of counterparts, and neither is $a\colon \neg Fx$. So there are four possibilities: a sentence can have just value 1 at a world, just value 0, both (if there are multiple counterparts), or neither (if there are no counterparts). The result resembles a well-known system from [Dunn 1976]. For atomic formulas, we have

$\mathcal{S}, \mathbb{V}, w \Vdash_1 P^n x_1 \ldots x_n$ iff there are $V, V_1 \ldots V_n \in \mathbb{V}$ such that $\langle V_1(x_1), \ldots, V_n(x_n) \rangle \in V_w(P^n)$.

9 The interpretations of 'actually' in [Ramachandran 1989], [Forbes 1982], and follow-up work by these authors resemble this approach.

$\mathcal{S}, \mathbb{V}, w \Vdash_0 P^n x_1 \ldots x_n$ iff there are $V, V_1 \ldots V_n \in \mathbb{V}$ such that $\langle V_1(x_1), \ldots, V_n(x_n) \rangle \notin V_w(P^n)$.

For complex formulas, it should suffice to give the clauses for negation and the colon; the others are similar.

$\mathcal{S}, \mathbb{V}, w \Vdash_1 \neg A$ iff $\mathcal{S}, \mathbb{V}, w \Vdash_0 A$.

$\mathcal{S}, \mathbb{V}, w \Vdash_0 \neg A$ iff $\mathcal{S}, \mathbb{V}, w \Vdash_1 A$.

$\mathcal{S}, \mathbb{V}, w \Vdash_1 a\colon A$ iff $\mathcal{S}, \mathbb{V}', V(a) \Vdash_1 A$, where V is any member of $\mathbb{V}$ and $\mathbb{V}'$ is the set of $V(a)$-images of members of $\mathbb{V}$ at w.

$\mathcal{S}, \mathbb{V}, w \Vdash_0 a\colon A$ iff $\mathcal{S}, \mathbb{V}', V(a) \Vdash_0 A$, where V is any member of $\mathbb{V}$ and $\mathbb{V}'$ is the set of $V(a)$-images of members of $\mathbb{V}$ at w.

We also need to specify under what conditions a sentence shall count as true in a model. As before, we might require truth on all choices of counterparts, so that A is true in $\langle \mathcal{S}, V \rangle$ iff $\mathcal{S}, \{V\}, @ \Vdash_1 A$ and $\mathcal{S}, \{V\}, @ \nVdash_0 A$. Or we might only require truth on some choice of counterparts, so that A is true in $\langle \mathcal{S}, V \rangle$ iff $\mathcal{S}, \{V\}, @ \Vdash_1 A$. The first yields a paracomplete logic: if x has two counterparts at the a world only one of which is F, then neither $a\colon Fx$ nor $\neg a\colon Fx$ is true in the model. On the second, weaker reading, the logic is both paracomplete and paraconsistent (indeed, dialethic): in the situation just described, $a\colon Fx$ and $\neg a\colon Fx$ are both true; in a situation where x has no counterpart at the a world, they are both untrue. Nonetheless, these logics are in many respects quite conservative. Call a sentence *weakly true* in a model if it is not determinately false, i.e. if the value of the sentence at $@, \{V\}$ is not just 0. (HK) and (HD) are weakly true in every model, as are all principles of nonvalent free logic and K, even without limiting classical substitution principles.

We have seen many options for extending counterpart semantics to hybrid languages. Which is the right one? None of them strike me as seriously flawed or unusable, so the choice should be based on whatever proves most convenient for a given application. Which is best for the logic of possibility, or the logic of time? I would recommend something like the weak reading of the last option I discussed, but I won't argue the point here. Admittedly, none of the proposals can be called entirely obvious or unproblematic. But this was to be expected. Remember the kind of problem they are meant to address: when a time traveler visits her younger self, what should we say about sentences like 'at t, Alice wakes up Alice', or 'at t, Alice is asleep and Alice is awake'? It would be surprising if these questions had an obvious and unproblematic answer.

One answer that I haven't mentioned, and that I suspect most people would be tempted to give, is to reject the questions: if ambiguity gets confusing, the best response is to disambiguate. This would mean using different expressions to pick out the two Alice counterparts at t. Of course it is not enough to simply use new names, say, 'Jennifer' and 'Natalie'. We also have to say that Jennifer and Natalie are, in a sense, just Alice. That is, they are both counterparts of Alice. The problem is that in the language of quantified hybrid logic one can't (in any ordinary way) express things like 'at t, there are two counterparts of Alice', or 'Natalie at t is a counterpart of Alice at t''. Without drastically revising the syntax of quantified modal logic, resolving the ambiguity is impossible, so we have to settle for some way of coping with it.

6 Contradictions

Let me return to the differences between Kripke semantics and counterpart semantics. I have not raised any objections to Kripke semantics. After all, Kripke semantics *is* counterpart semantics with identity as the counterpart relation. For certain applications, this special case may be all we need, just as there are applications where we can fix accessibility to be the universal relation between all worlds. However, there are good reasons – both in logic and in philosophy – to look at the more general case. I have argued that this step is not as daring as it is sometimes presented. In its most general form, counterpart semantics does not require disjoint domains. It does not imply that ordinary individuals exist only at a single world or time. Nor does it require centring on individuals rather than worlds. You can have non-trivial counterpart models in which (K) is valid and (NE) invalid. You do not have to buy the whole package of "counterpart theory" to use counterpart semantics.

One of the strengths of counterpart semantics, I have argued, is that it determines a very weak modal logic. Some philosophers have tried to turn this strength into a weakness, by applying the following recipe (see e.g. [Fara and Williamson 2005]). Start with some formula like

(6) $\quad \Diamond \exists x(\mathsf{c}\colon Fx \land \mathsf{c}\colon \neg Fx)$,

that is satisfiable in counterpart semantics but not in Kripke semantics. Next, intuit that the formula is a "logical contradiction", and fault counterpart semantics for rendering it satisfiable. Let me conclude with two brief comments on this type of argument.

First, a formula like (6) can obviously be interpreted in ways on which it is not only not contradictory, but true. So if you have intuitions about whether or not (6) is contradictory, you have probably assigned some meaning to the logical terms, perhaps by informally reading the diamond as 'possibly' (or 'at some time') and '$\mathtt{c}$:' as 'actually' (or 'now'). But then you are already thinking about a specific *model*, or class of models. One would not criticise Kripke semantics for rendering

(7) $\Box Fx \wedge \neg Fx$

satisfiable, although this is intuitively a contradiction if the box is read as 'necessarily' or 'always'. If we want to use Kripke semantics for the logics of possibility and time, we simply have to limit the relevant models (or rather frames) so that (7) can never be true. The same should hold for counterpart semantics. If (6) is contradictory in the logic of possibility or time, then we can simply limit the relevant counterpart structures to functional structures.

Second, the intuition that (6) is contradictory in the logics of possibility and time is at best questionable. Remember Frege's Basic Law V: some general principles seem utterly compelling until you think of the peculiar instances in which they fail. If it seems to you that (6) must always be false, think about the kind of scenario in which a counterpart theorist might say that it is true. One such scenario is that of Alice the time traveler: according to some versions of counterpart semantics, (6) is true if '$\mathtt{c}$' picks out the time at which Alice visits her younger self and 'F' stands for something like being awake. (6) may then be expressed in English as the claim that at some time there was a person such that now this person is awake and now this person is not awake. Under the peculiar circumstances of the scenario, this does not strike me as obviously false. It may not be determinately true either, and I've outlined alternative versions of counterpart semantics on which it comes out false. But other, equally strange things will then come out true. This is as it should be: under strange conditions, strange things may be true.

References

Patrick Blackburn [2006]: “Arthur Prior and Hybrid Logic”. *Synthese*, 150: 329–372

Patrick Blackburn, Maarten de Rijke and Yde Venema [2001]: *Modal Logic*. Cambridge: Cambridge University Press

David Bostock [1997]: *Intermediate Logic*. Oxford: Clarendon Press

Torben Braüner and Silvio Ghilardi [2006]: “First-order Modal Logic”. In Patrick Blackburn, Johan van Benthem and Frank Wolter (Eds.) *Handbook of Modal Logic,* Dordrecht: Elsevier, 549–620

Rudolf Carnap [1947]: *Meaning and Necessity*. Chicago: University of Chicago Press

Giovanna Corsi [2002]: “Counterpart Semantics. A Foundational Study on Quantified Modal Logics”. Manuscript

J. M. Dunn [1976]: “Intuitive Semantics for First Degree Entailment and ‘Coupled Trees”’. *Philosophical Studies*, 29: 149–168

Michael Fara and Timothy Williamson [2005]: “Counterparts and Actuality”. *Mind*, 114: 1–30

Melvin C. Fitting [2004]: “First-order intensional logic”. *Annals of Pure and Applied Logic*, 127: 171–193

Graeme Forbes [1982]: “Canonical Counterpart Theory”. *Analysis*, 42: 33–37

James W. Garson [1984]: “Quantification in Modal Logic”. In D. Gabbay and F. Guenthner (Eds.) *Handbook of Philosophical Logic,* vol 2. Dordrecht: Reidel, 249–307

Silvio Ghilardi [2001]: “Substitution, quantifiers and identity in modal logic”. In E. Morscher and A. Hieke (Eds.) *New essays in free logic,* Dordrecht: Kluwer, 87–115

Silvio Ghilardi and Giancarlo Meloni [1991]: “Philosophical and mathematical investigations in first-order modal logic”. In *Problemi fondazionali in teoria del significato. Atti del convegno di Pontignano,* Firenze: Olsckhi, 77–107

Allen P. Hazen [1979]: “Counterpart-Theoretic Semantics for Modal Logic”. *Journal of Philosophy*, 76: 319–338

G. E. Hughes and Max J. Cresswell [1996]: *A New Introduction to Modal Logic*. London and New York: Routledge

Markus Kracht and Oliver Kutz [2002]: “The semantics of modal predicate logic I. Counterpart frames”. In Frank Wolter, Heinrich Wansing, Maarten de Rijke and Michael Zakharayaschev (Eds.) *Advances in Modal Logic,* vol 3. World Scientific Publishing Company

— [2005]: “The semantics of modal predicate logic II. Modal individuals revisited”. In Reinhard Kahle (Ed.) *Intensionality,* Los Angeles: A. K. Peters

Saul Kripke [1979]: “A Puzzle About Belief”. In Avishai Margalit (ed.), *Meaning and Use*, Dordrecht: Reidel

David Lewis [1968]: “Counterpart Theory and Quantified Modal Logic”. *Journal of Philosophy*, 65: 113–126

— [1973]: *Counterfactuals*. Oxford: Blackwell

— [1983]: “Postscript to [Lewis 1968]”. In *Philosophical Papers I,* Oxford: Oxford University Press, 39–46

— [1986]: *On the Plurality of Worlds*. Malden (Mass.): Blackwell

Murali Ramachandran [1989]: “An Alternative Translation Scheme for Counterpart Theory”. *Analysis*, 49: 131–141

Gerhard Schurz [1997]: *The is-ought problem. A study in philosophical logic*, vol 1 of *Trends in Logic*. Dordrecht: Kluwer

Wolfgang Schwarz [2011]: “Generalising Kripke semantics for quantified modal logics”. Manuscript

Theodore Sider [2001]: *Four-Dimensionalism*. Oxford: Clarendon Press

Robert Stalnaker [1987]: "Counterparts and Identity". *Midwest Studies in Philosophy*, 11: 121–140

Johan van Benthem [1983]: *Modal Logic and Classical Logic*. Naples: Bibliopolis

Lin Woollaston [1994]: "Counterpart Theory as a Semantics for Modal Logic". *Logique et Analyse*, 147–148: 255–263