The Problem of Cross-World Predication*

Alex Kocurek *Journal of Philosophical Logic*, Vol. 45, No. 6, pp. 697–742.

Abstract. While standard first-order modal logic is quite powerful, it cannot express even very simple sentences like "I could have been taller than I actually am" or "Everyone could have been smarter than they actually are". These are examples of *cross-world predication*, whereby objects in one world are related to (sometimes the same) objects in another world. Extending first-order modal logic to allow for cross-world predication in a motivated way has proven to be notoriously difficult. In this paper, I argue that the standard accounts of cross-world predication all leave something to be desired. I then propose an account of cross-world predication based on *quantified hybrid logic* and show how it overcomes the limitations of these previous accounts. I will conclude by discussing various philosophical consequences and applications of such an account.

§1 Introduction

Consider the following sentence:

(Tall) I could have been taller than I actually am.

Such a sentence cannot be expressed in standard first-order modal logic. To see why, let Taller be a taller-than predicate, where "$\mathsf{Taller}(x, y)$" is read as "x is taller than y", and let me be a constant denoting me. Now, ask yourself: how would one formalize (Tall)? It doesn't take much to see that

$$\Diamond \mathsf{Taller}(\mathsf{me}, \mathsf{me}) \tag{1}$$

is no good. For this says that there's a possible world where I (in that world) am taller than myself (in that world). But this is simply nonsense: nothing can be taller than itself.

Does adding an actuality operator @ help? No. For

$$@\Diamond \mathsf{Taller}(\mathsf{me}, \mathsf{me}) \tag{2}$$

says that actually, there's a possible world where I (in that world) am taller than myself (in that world), which is just as nonsensical as before. And

$$\Diamond @\mathsf{Taller}(\mathsf{me}, \mathsf{me}) \tag{3}$$

is equivalent to $@\mathsf{Taller}(\mathsf{me}, \mathsf{me})$, which says (nonsensically) that I am actually taller than myself. It doesn't seem like any of the straightforward attempts to formalize (Tall) in first-order modal logic work.

Consider another example:

*Many thanks to Johan van Benthem, Russell Buehler, Balder ten Cate, Sophie Dandelet, Melissa Fusco, Wesley Holliday, Grace Paterson, Justin Vlasits, Kai Wehmeier, and Seth Yalcin for all their helpful comments and suggestions. A version of this paper was presented at the Berkeley-Stanford Circle in Logic and Philosophy in October 2014. A version of this paper was also presented at a C-ALPHA-sponsored talk at UC Irvine in March 2015 and at the Logic Seminar at Stanford in May 2015. I am very grateful for all the helpful comments and discussion from these talks.

(Smart) Everyone could have been smarter than they actually are.

Let Smarter be a smarter-than predicate, where "$\mathsf{Smarter}(x, y)$" is read as "x is smarter than y". How would one formalize (Smart)? Obviously,

$$\Diamond \forall x \, \mathsf{Smarter}(x, x) \tag{4}$$

won't do. This says that in some possible world, everyone (in that world) is smarter than themselves (in that world)—nonsense. Similarly,

$$\forall x \, \Diamond \mathsf{Smarter}(x, x) \tag{5}$$

doesn't work, since this says that everyone has a possible world where they (in that world) are smarter than themselves (in that world)—again, nonsense. And one can check that adding an actuality operator doesn't help; no matter where you put @ in either (4) and (5), the reading one obtains is nonsense. Again, no straightforward attempt at formalization is successful.

Finally, consider a slightly more complicated example:

(Polar) A polar bear could be bigger than a grizzly bear could be.[1]

Semi-formally, this says that there's a possible world w and a polar bear x in w such that, for any possible world v, and for any grizzly bear y in v, x in w is bigger than y is in v.

After adding the appropriate predicates once more, how would one formalize (Polar)? Unfortunately, the most natural attempt yields

$$\Diamond \exists x \; (\mathsf{Polar}(x) \wedge \Box \forall y \; (\mathsf{Grizzly}(y) \rightarrow \mathsf{Bigger}(x, y))) \,, \tag{6}$$

which, while not nonsense, isn't a correct formalization of (Polar). For (6) says that there's a possible world w and a polar bear x in w such that, for any possible world v, and for any grizzly bear y in v, x <u>*in v*</u> is bigger than y in v. But what we want is for x <u>*in w*</u> to be bigger than y in v—it's x's size <u>*in w*</u> that places an upper bound on how big a grizzly bear can get. And as one can check, no permutation of the modals and quantifiers, nor any addition of an actuality operator, yields a correct formalization of (Polar). Once again, we're stuck.

Now, to be fair, such arguments aren't proofs. Just because the most straightforward attempt to formalize an ordinary sentence into a formal language fails, it doesn't follow that there aren't more subtle or indirect ways to formalize that sentence. However, using standard techniques from modal model theory, these informal arguments can be backed up by rigorous proof: one can prove these sentences really aren't expressible in any standard first-order modal logic (see §A.3).

Sentences like (Tall), (Smart), and (Polar) are examples of ***cross-world predications***,[2] which involve relating objects in one world to (perhaps the same) objects in another world. That is, cross-world predications are predications between objects *across worlds* rather than *within worlds*. First-order modal logic, as we've just seen, is unequipped to formalize even very simple examples of cross-world predication, and it's not clear what kind extension is needed to accommodate them. Call this ***the problem of cross-world predication***.

The problem of cross-world predication is widespread, and a solution to it would have broad applications. I will illustrate with three such applications.

[1] Originally from von Stechow [1984, p. 35].
[2] I borrow the terminology from Wehmeier [2012].

Application 1: Fiction. Consider a famous puzzle in the philosophy of fiction. Both of these sentences seem true:

(Detective) Sherlock Holmes is a detective.

(Non-exist) Sherlock Holmes does not exist.

However, they also seem incompatible. In order to be a detective, one must exist. If Holmes doesn't exist, then he can't be a real detective. What's going on?

One view, which I'll call the ***operator view of fiction***, is that when we assert a sentence like (Detective), we really mean to be asserting it as if it were in the scope of some kind of operator, like "in the Holmes stories", which is amenable to a modal treatment. So on this view, what (Detective) typically means is something like:

(Dectective*) In the Holmes stories, Sherlock Holmes is a detective.

which, semi-formally, says that in all fictional worlds w compatible with the Homles stories, Sherlock Holmes in w is a detective.

But here's a problem for the operator view. Consider the following sentence:[3]

(Bilbo) Bilbo in *The Lord of the Rings* is taller than Thumbelina in *Thumbelina*.

Suppose we have operators like $[\mathsf{LotR}]$ for "in all fictional worlds compatible with *The Lord of the Rings*" and $[\mathsf{Thumb}]$ for "in all fictional worlds compatible with *Thumbelina*". How do we express (Bilbo) with such operators? Clearly,

$$[\mathsf{LotR}]\,\mathsf{Taller}(\mathsf{bilbo},\mathsf{thumb}) \tag{7}$$

doesn't work since it says that, in *The Lord of the Rings*, Bilbo is taller than Thumbelina. But Thumbelina doesn't even exist in *The Lord of the Rings*. Similarly,

$$[\mathsf{LotR}]\,[\mathsf{Thumb}]\,\mathsf{Taller}(\mathsf{bilbo},\mathsf{thumb}) \tag{8}$$

doesn't work since Bilbo doesn't exist in *Thumbelina*. Once again, we can't seem to find an adequate formalization of this rather natural thought.

It's not hard to see that (Bilbo) is just a special case of cross-world predication—or rather, cross-*fictional* predication. As we've just seen, sentences like (Bilbo) could be used as a preliminary criticism of the operator view of fiction. But if we had a way of solving the problem of cross-world predication, we might also have an analogous way of defending the operator view of fiction.

Application 2: Counteridenticals. Counterfactuals provide very natural examples of cross-world predication. For example, we might say "If I were taller, I would be a basketball player" to mean that if I were taller *than I actually am*, I would be a basketball player. But consider the following sentence:

(Horse) If I were you, I wouldn't bet on that horse.

[3] From Button [2012, p. 245].

It's difficult to say exactly how to understand sentences like (Horse), even semi-formally. Such sentences are known as ***counteridenticals***, which, very roughly, involve counterfactuals whose antecedent is some kind of identity claim.[4]

So how do we semi-formalize (Horse)? Notice the naïve thought doesn't work: we can't read (Horse) as saying that in the nearest possible world where I am identical to you, I (in that world) don't bet on that horse. For one thing, it's simply unclear how I could possibly be identical to you. But moreover, the relation in the antecedent of counteridenticals isn't generally reflexive, symmetric, or transitive. For instance, against reflexivity, the sentence

(Bold) I would be bolder if I weren't me.

is non-trivial. But it would involve an impossible antecedent if the antecedent was interpreted literally. Against symmetry, (Horse) isn't equivalent to:

(Horse*) If you were me, I wouldn't bet on that horse.

And against transitivity, this sentence doesn't seem necessarily true, but it would be if we interpret these as identity statements:

(Sally) If I were you and you were Sally, I would be Sally.

Furthermore, such an analysis of counteridenticals would render the following (obviously sensical) sentence nonsensical:

(Star) If I were you and you were me, I would be a rock star and you wouldn't.

In the nearest possible world where we're both identical, we can't have conflicting properties. Hence, so-called "counteridenticals" can't be *literally* interpreted as counterfactuals with identity statements in the antecedent. Something else must be going on.

Here's a rather plausible reading of (Horse): in the nearest possible world where "I am in your shoes", I wouldn't bet on that horse (in that world). Now, how we interpret "in your shoes" will most likely vary with the context. Let's set aside this context-sensitivity for now and just stipulate a "in your shoes" predicate, $\mathsf{Shoes}(x, y)$, where "$\mathsf{Shoes}(x, y)$" is read as "x is in the same position as y", or more loosely "x is in y's shoes". Such an approach seems quite natural, and has been suggested elsewhere.[5]

But this approach suffers a problem. For clearly, we can't formalize (Horse) as:

$$\mathsf{Shoes}(\mathsf{me}, \mathsf{you}) \mathrel{\Box\!\!\rightarrow} \neg\, \mathsf{Bet}(\mathsf{me}, \mathsf{horse}). \tag{9}$$

This says that in the nearest possible world where I (in that world) am in the same position that you *in that world* are in, I don't bet on that horse. But we want to consider the nearest possible world where I (in that world) am in the same position that you are *actually* in.

Once again, it's clear that this problem is just a special case of the problem of cross-world predication—it's just one where the cross-world predication is in the antecedent of a counterfactual. So, assuming a solution to cross-world predication is compatible with the standard analysis of counterfactuals, a solution to the problem of cross-world predication will generate a solution to this problem for this approach to counteridenticals.

[4] I'm grateful to Mike Deigan for drawing my attention to these kinds of sentences.

[5] For instance, Pollock [1976, pp. 6–7], though Pollock thinks it has several problems (pp. 114–115).

Application 3: Supervenience. Finally, consider one version of a famous thesis in the philosophy of mind:

(Super) The mental supervenes on the physical—there can't be any difference in the mental without a difference in the physical.

Lewis [1986, pp. 14–17] argued that such a thesis couldn't be expressed in first-order modal logic. Let $\sim_{\mathsf{M}}$ and $\sim_{\mathsf{P}}$ be predicates where "$x \sim_{\mathsf{M}} y$" and "$x \sim_{\mathsf{P}} y$" are read respectively as "There's no mental difference between x and y" and "There's no physical difference between x and y". The natural attempt to formalize (Super) would be:

$$\Box \forall x \, \forall y \; (x \sim_{\mathsf{P}} y \rightarrow x \sim_{\mathsf{M}} y)\,. \tag{10}$$

But (10) is too weak. It says that for all worlds w, and for any x and y, if there's no physical difference between x and y <u>*in w*</u>, then there's no mental difference between x and y <u>*in w*</u>. But as Lewis [1986, p. 16] points out, (Super) says something stronger. It says that if x in some world w is not physically different from y in some world v, then x in w is not mentally different from y in v. Even if the supervenience thesis was upheld *within* every world, it wouldn't follow that the supervenience thesis was upheld *across* every world.

You can probably guess how the story goes—(Super) is yet another instance of cross-world predication that resists formalization into first-order modal logic. Thus, if we can solve the problem of cross-world predication, it might in doing so lend us useful tools for framing important philosophical debates.

The problem of cross-world predication is not new. A number of philosphers and logicians have theorized about different ways to extend or revise first-order modal logic to accommodate cross-world predication. However, as we'll see below, most of the approaches that have been proposed are unsatisfactory. Many approaches contain some gap in expressive power, rendering them too weak to capture cross-world predication in full generality. Others, while powerful enough to express cross-world predication in full, have a number of odd, undesirable features. None of these approaches strikes the right balance between expressive power and parsimony.

The goal of this paper is to present a parsimonious solution to the problem of cross-world predication and to provide a principled motivation for its adoption. As I'll show, the approach defended here is, on the one hand, powerful enough to capture cross-world predication while, on the other hand, more conservative and more unified than previous approaches.

An outline of the paper is as follows. In §2, we'll briefly examine a phenomenon which is closely related to cross-world predication, and motivate seeking a unified solution to both problems. In §3, we'll briefly recount the standard syntax and semantics for (two-dimensional) first-order modal logic. In §4–6, I present and criticize some accounts of cross-world predication that have already been proposed in the literature. In §7, I present another account, based on *quantified hybrid logic*, and show how it overcomes the difficulties of the previous views. I will conclude with a philosophical discussion in §8.

§2 Cross-world Quantification

Before diving in, let's examine another phenomenon, similar to cross-world predication, which also evades formalization in first-order modal logic. Consider the sentence:

(Rich) The rich could have all been poor.[6]

Semi-formally, this says (at least on one reading) that there's a possible world v such that everyone in the actual world w that's rich in w is poor in v.

Now, whether or not (Rich) is expressible in first-order modal logic depends on whether the quantifiers we use are actualist (that is, ranging over the local domain of the world of evaluation) or possibilist (that is, ranging over the global domain of the whole model).

If we only use actualist quantifiers, then (Rich) isn't expressible in first-order modal logic (even with the actually operator @). Consider a first attempt at formalization:

$$\Diamond\forall x\ (@\mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x))\,. \tag{11}$$

Semi-formally, this says that there's a possible world v such that anyone *in v* that's rich in the actual world w is poor in v. But this is too weak; we want to quantify over every actual rich person *in w*, not just those in v. In particular, (11) would be true if v was the only accessible world from w, and if no one in w existed in v; but (Rich) wouldn't be true.

We could try to resolve this issue by pulling the quantifier out, yielding

$$\forall x\,\Diamond\,(@\mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x))\,, \tag{12}$$

which, assuming there are accessible worlds, will be equivalent to

$$\forall x\ (@\mathsf{Rich}(x) \rightarrow \Diamond\mathsf{Poor}(x))\,. \tag{13}$$

But (13), unlike (Rich), could be satisfied in a situation where every actual rich person has a different possible world in which they are poor, even when there's no single possible world where they are all poor *together*. Thus, at least when we restrict ourselves to actualist quantifiers, first-order modal logic cannot express (Rich), even with @.[7]

If we use the possibilist quantifier Π instead of $\forall$, then

$$\Diamond\Pi x\ (@\mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x)) \tag{14}$$

is a fine formalization of (Rich). For this says that there's a possible world v such that anyone *in the entire model* that's rich in the actual world w is poor in v. But, presumably, if Rich is to really mean "is rich" in our models, we'll need to impose a semantic constraint that nothing could be rich at a world without existing at that world: being rich is an existence-entailing property. So adding this contraint to our models, (14) equivalently says that there's a possible world v such that anyone *in w* that's rich in w is poor in v—precisely what (Rich) says.

[6] From Cresswell [1990, p. 34].

[7] There are several proofs that first-order modal logic *without* @ can't express (Rich), regardless of whether we use possibilist or actualist quantifiers [Hodes, 1984; Wehmeier, 2001]. A proof that adding @ doesn't help (at least for the language with just actualist quantifiers) can be found in Kocurek [2015].

Some might use this as an argument in favor of adopting a possibilist reading of the quantifiers over an actualist reading. However, we're not *that* much better off with possibilist quantification. For consider the necessitation of (Rich):

(Rich*) Necessarily, the rich could have all been poor.

Semi-formally, at least on one reading, this says that in all possible worlds v, there's a possible world u such that everyone in v that's rich in v is poor in u. This simple modification of (Rich) is inexpressible in first-order modal logic, even with possibilist quantification. To illustrate, consider the necessitation of (14):

$$\Box\Diamond\Pi x\ (@\mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x))\,. \tag{15}$$

This says that in every possible world v, there's a possible world u where everyone that's rich *in the actual world w* is poor in u. But we just want to consider everyone that's rich *in v*, not in w—the @ takes us back too far. Although possibilist quantification offers a solution to the formalization of (Rich), it can't handle even slightly more complicated examples.[8]

Both (Rich) and (Rich*) are examples of what I will call ***cross-world quantification***, which involves quantifying outside the scope of a modal operator. As we've just seen, first-order modal logic is unequipped to formalize even very simple examples of cross-world quantification. Call this ***the problem of cross-world quantification***.

While many philosophers have discussed (Rich) and have sought a solution to this problem, not many have offered a *combined* solution to the problems of cross-world predication and cross-world quantification. But in retrospect, this is quite odd, since both phenomena seem deeply connected. It would be nice if a solution to the problem of cross-world predication managed to also account for cross-world quantification and perhaps even explain what the connection between the two is. As we'll see, the account proposed in this paper (presented in §7) can do this.

§3 First-Order Modal Logic

Before we can examine the various extensions of first-order modal logic to accommodate cross-world predication, we'll need to first get clear about what exactly we're taking first-order modal logic to be. In setting out the details of first-order modal logic, we are forced to make a number of hard choices regarding its syntax, models, and semantics. Some of these choices are methodological, others philosophical. We review the most important choices below. For many of the others, we will silently make a decision and carry on without comment. In all cases, our silent decisions are motivated towards making first-order modal logic as expressive as possible.

§3.1 Syntax

We start by laying out the syntax for our first-order modal language, $\mathcal{L}^{\mathsf{1M}}$. The signature we adopt is the same throughout. Let:

[8] The proof of this claim can also be found in Kocurek [2015].

- $\mathsf{CON} = \{c_1, c_2, c_3, \ldots\}$ (the set of ***constants***);
- $\mathsf{VAR} = \{x_1, x_2, x_3, \ldots\}$ (the set of ***(object) variables***);
- $\mathsf{TERM}^{\mathsf{1M}} := \mathsf{CON} \cup \mathsf{VAR}$ (the set of ***(object) terms*** or ***$\mathcal{L}^{\mathsf{1M}}$-terms***);
- $\mathsf{PRED}^n = \{P^n_1, P^n_2, P^n_3, \ldots\}$ for each $n \geqslant 1$ (the set of ***n-place predicates***);
- $\mathsf{PRED} := \bigcup_{n \geqslant 1} \mathsf{PRED}^n$ (the set of ***predicates***).

The set of ***formulas in $\mathcal{L}^{\mathsf{1M}}$*** or ***$\mathcal{L}^{\mathsf{1M}}$-formulas***, $\mathsf{FORM}^{\mathsf{1M}}$, is defined recursively:

$$\varphi ::= P^n(\tau_1, \ldots, \tau_n) \mid \tau_1 \approx \tau_2 \mid \mathsf{E}(\tau_1) \mid \neg\,\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid @\varphi \mid \forall x\, \varphi$$

where $P^n \in \mathsf{PRED}^n$, $\tau_1, \ldots, \tau_n \in \mathsf{TERM}^{\mathsf{1M}}$, and $x \in \mathsf{VAR}$. The usual abbreviations for $\vee$, $\rightarrow$, $\exists$, and $\Diamond$ all apply. To be clear, $\approx$ is the identity symbol, E is an existence predicate, and $@$ an actuality operator.[9] We'll let $\mathcal{L}^{\mathsf{1M}}_{\Pi}$ be the language obtained by adding to $\mathcal{L}^{\mathsf{1M}}$ the universal possibilist quantifier Π (the existential counterpart is Σ). When no ambiguity arises, we'll drop parentheses for readability.

§3.2 Models

Next, we look at how first-order modal models are defined.

Definition 3.1 *(First-Order Modal Models)*. An ***$\mathcal{L}^{\mathsf{1M}}$-model***, or ***(modal) model***, is an ordered tuple $\mathcal{M} = \langle W, R, D, \delta, I \rangle$ where:

- W is a nonempty set (the ***state space***);
- $R \subseteq W \times W$ (the ***accessibility relation***), where $R[w] := \{v \in W \mid R(w, v)\}$;
- D is a nonempty set (the ***(global) domain***);
- $\delta \colon W \rightarrow \wp(D)$ is a function (the ***local domain assignment***), where for each $w \in W$, $\delta(w)$ is the ***local domain of*** w;
- I is a function (the ***interpretation function***) such that:
  - for each $c \in \mathsf{CON}$, $I(c, w) \in D$;
  - for each $P^n \in \mathsf{PRED}^n$, $I(P^n, w) \subseteq D^n$.

Let me emphasize three features of our models. First, we don't require constants to denote rigidly—a constant may denote different objects in different worlds. Second, our models are variable domain models—some objects may fail to exist in some worlds. Third, we don't require that objects exist in order to instantiate predicates—non-existents can still have properties, stand in relations, etc. These assumptions are made to construe "standard" first-order modal logic in its most general form and may be dropped if desired. They won't affect the limitations considered in what follows.

[9] One could consider adding λ-abstraction, in the spirit of Fitting and Mendelsohn [1998, Chps. 9-10], but doing so doesn't increase the expressive power enough to solve the problems of cross-world predication and quantification [Kocurek, 2015]. All λ-abstraction does is essentially allow for "rigidification" of non-rigid terms, which doesn't help much (though see Fitting [2013] for an approach that combines λ-abstraction with the approach in §5 for sentences like (Tall) without quantifiers).

§3.3 Semantics

Finally, we examine the semantics for first-order modal logic.

Definition 3.2 *(Variable Assignment)*. Let $\mathcal{M}$ be an $\mathcal{L}^{1\mathsf{M}}$-model. A ***variable assignment for*** $\mathcal{M}$ is a function assigning members of its global domain to variables. The set of variable assignments for $\mathcal{M}$ will be $\mathsf{VA}(\mathcal{M})$.

If a variable assignment g on $\mathcal{M}$ agrees with a variable assignment g' on $\mathcal{M}$ on every variable except possibly x, then g and g' are ***x-variants***, $g \sim_x g'$. The variable assignment $g[x \mapsto a]$ or (more compactly) g^x_a is the x-variant of g that sends x to a.

With @ in our language, our semantics will need to be two-dimensional (à la Davies and Humberstone [1980, pp. 4-5]). That is, indices will have two worlds. The first world is to be interpreted as the world "considered as actual", and the second as the world of evaluation.

Definition 3.3 *(Denotation)*. Let τ be a term, let $\mathcal{M}$ be an $\mathcal{L}^{1\mathsf{M}}$-model, let $w, v \in W$, and let $g \in \mathsf{VA}(\mathcal{M})$. The ***denotation of*** τ ***at*** $\langle \mathcal{M}, w, v, g\rangle$, $[\![\tau]\!]^{\mathcal{M},w,v,g}$, is defined as follows:

$$[\![\tau]\!]^{\mathcal{M},w,v,g} = \begin{cases} I(c, v) & \text{if } \tau = c \text{ where } c \in \mathsf{CON} \\ g(x) & \text{if } \tau = x \text{ where } x \in \mathsf{VAR}. \end{cases}$$

Note that w (the world considered as actual) makes no difference to denotations. Nevertheless, we include it to ease the transition into §6, where it *will* make a difference.

Definition 3.4 *(Satisfaction)*. The ***satisfaction relation***, $\Vdash$, is defined recursively over $\mathcal{L}^{1\mathsf{M}}_{\Pi}$-formulas, for all $\mathcal{L}^{1\mathsf{M}}$-models $\mathcal{M} = \langle W, R, D, \delta, I\rangle$, all $w, v \in W$, and all $g \in \mathsf{VA}(\mathcal{M})$:

$$\begin{array}{lcl}
\mathcal{M}, w, v, g \Vdash P^n(\tau_1, \ldots, \tau_n) & \Leftrightarrow & \left\langle [\![\tau_1]\!]^{\mathcal{M},w,v,g}, \ldots, [\![\tau_n]\!]^{\mathcal{M},w,v,g} \right\rangle \in I(P^n, v) \\
\mathcal{M}, w, v, g \Vdash \tau \approx \sigma & \Leftrightarrow & [\![\tau]\!]^{\mathcal{M},w,v,g} = [\![\sigma]\!]^{\mathcal{M},w,v,g} \\
\mathcal{M}, w, v, g \Vdash \mathsf{E}(\tau) & \Leftrightarrow & [\![\tau]\!]^{\mathcal{M},w,v,g} \in \delta(v) \\
\mathcal{M}, w, v, g \Vdash \neg\varphi & \Leftrightarrow & \mathcal{M}, w, v, g \nVdash \varphi \\
\mathcal{M}, w, v, g \Vdash \varphi \wedge \psi & \Leftrightarrow & \mathcal{M}, w, v, g \Vdash \varphi \text{ and } \mathcal{M}, w, v, g \Vdash \psi \\
\mathcal{M}, w, v, g \Vdash \Box\varphi & \Leftrightarrow & \forall v' \in R[v]\colon\ \mathcal{M}, w, v', g \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash @\varphi & \Leftrightarrow & \mathcal{M}, w, w, g \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash \forall x\, \varphi & \Leftrightarrow & \forall a \in \delta(v)\colon\ \mathcal{M}, w, v, g^x_a \Vdash \varphi \\
\mathcal{M}, w, v, g \Vdash \Pi x\, \varphi & \Leftrightarrow & \forall a \in D\colon\ \mathcal{M}, w, v, g^x_a \Vdash \varphi.
\end{array}$$

If Γ is a set of $\mathcal{L}^{1\mathsf{M}}_{\Pi}$-formulas, then $\mathcal{M}, w, v, g \Vdash \Gamma$ if for all $\varphi \in \Gamma$, $\mathcal{M}, w, v, g \Vdash \varphi$.

§4 The Two-Sorted Language

We now turn to the various extensions of first-order modal logic that have been proposed to handle cross-world predication. The first proposal is quite simple: just move to a two-sorted language—one sort for objects, the other for worlds.

§4.1 Formalism

Here's what such a proposal would look like. The signature of this two-sorted language, $\mathcal{L}^{2S}$, will include CON and VAR as before, but will also include the following set:

- SVAR = $\{s_1, s_2, s_3, \dots\}$ (the set of ***state variables***).

The set of ***(object) terms in $\mathcal{L}^{2S}$*** or ***$\mathcal{L}^{2S}$-terms***, TERM^{2S}, is defined as follows:

$$\tau ::= x \mid c(s)$$

where $x \in$ VAR, $c \in$ CON, and $s \in$ SVAR. In other words, in $\mathcal{L}^{2S}$, c is treated as a *unary function symbol*. The reason for doing this is that we allowed for constants in $\mathcal{L}^{1M}$ to be non-rigid. To capture this in $\mathcal{L}^{2S}$, we need some way of keeping track of what world we're evaluating the denotation of c relative to. If we had required constants in $\mathcal{L}^{1M}$ to be rigid, then we could have just treated constants in $\mathcal{L}^{1M}$ as constants in $\mathcal{L}^{2S}$ also. In what follows, if c is meant to be interpreted as a rigid constant in $\mathcal{L}^{1M}$, we'll write "c" in $\mathcal{L}^{2S}$-formulas in place of "$c(s)$" for some arbitrary s.

In place of PRED^n, we have following set in our signature:

- PRED$^{n/m}$ = $\left\{P_1^{n/m}, P_2^{n/m}, P_3^{n/m}, \dots\right\}$ for each $n, m \geqslant 1$ (the set of ***n/m-place predicates***).

For a n/m-place predicate $P^{n/m}$, n is the object-arity, while m is the state-arity. Thus, $P^{n/m}$ takes exactly n object terms and m state variables as arguments to be well-formed.[10]

The set of ***formulas in $\mathcal{L}^{2S}$*** or ***$\mathcal{L}^{2S}$-formulas***, FORM^{2S}, is defined recursively:

$$\varphi ::= P^{n/m}(\tau_1, \dots, \tau_n; t_1, \dots, t_m) \mid \tau_1 \approx \tau_2 \mid t_1 \approx t_2 \mid \mathsf{E}(\tau_1; t_1) \mid \mathsf{R}(t_1, t_2)$$
$$\mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \forall x\, \varphi \mid \forall t\, \varphi$$

where $P^{n/m} \in$ PRED$^{n/m}$, $\tau_1, \dots, \tau_n \in$ TERM^{2S}, $t, t_1, \dots, t_m \in$ SVAR, and $x \in$ VAR.[11] The models and semantics for $\mathcal{L}^{2S}$ are just the standard models and semantics of fist-order logic with two sorts. The details are left to §A.

Using this two-sorted language, we can formalize (Tall), (Smart), (Polar), (Rich), and (Rich*) (where by convention s picks out the starting world of evaluation) as follows:[12]

[10] We'll use ";" to separate object terms and state variables. Also, we'll use "P^n" in place of "$P^{n/1}$".

[11] We won't really need formulas of the form $t_1 \approx t_2$ in this section, but we include them for the sake of generality. Such $\mathcal{L}^{2S}$-formulas will be discussed in §8.

[12] The formalizations assume possibilist object quantifiers, but the actualist readings could be obtained by relativizing the object quantifiers to E appropriately. We also assume (just for simplicity) that we always start truth evaluation at ***diagonal*** points of evaluation, whereby the world considered as actual is the world of evaluation. That way, we don't need to treat separate readings of the sentences, where the free state variables either pick out the starting world of evaluation or the world considered as actual.

(Tall) I could have been taller than I actually am.

$$\exists t\ (\mathsf{R}(s,t) \wedge \mathsf{Taller}\,(\mathsf{me},\mathsf{me};t,s)) \tag{16}$$

(Smart) Everyone could have been smarter than they actually are.

$$\forall x\,\exists t\ (\mathsf{R}(s,t) \wedge \mathsf{Smarter}(x,x;t,s)) \tag{17}$$

(Polar) A polar bear could be bigger than a grizzly bear could be.

$$\exists t\ [\mathsf{R}(s,t) \wedge \exists x\ [\mathsf{Polar}(x;t) \wedge \forall t'\ (\mathsf{R}(t,t') \to \forall y\ (\mathsf{Grizzly}(y;t') \to \mathsf{Bigger}(x,y;t,t')))]] \tag{18}$$

(Rich) The rich could have all been poor.

$$\exists t\ (\mathsf{R}(s,t) \wedge \forall x\ (\mathsf{Rich}(x;s) \to \mathsf{Poor}(x;t))) \tag{19}$$

(Rich*) Necessarily, the rich could have all been poor.

$$\forall s\,\exists t\ (\mathsf{R}(s,t) \wedge \forall x\ (\mathsf{Rich}(x;s) \to \mathsf{Poor}(x;t)))\,. \tag{20}$$

As one can check, all of the other example sentences mentioned so far can be formalized in $\mathcal{L}^{2S}$ in a similar manner.

§4.2 Expressive Power

Two facts about $\mathcal{L}^{2S}$ are important to note. First, every $\mathcal{L}^{1M}$-formula is equivalent to some $\mathcal{L}^{2S}$-formula. Second, not every $\mathcal{L}^{2S}$-formula is equivalent to some $\mathcal{L}^{1M}$-formula.

To see why the first fact holds, one can just translate every $\mathcal{L}^{1M}$-formula into an equivalent $\mathcal{L}^{2S}$-formula, just as one would translate every propositional modal formula into a first-order correspondence language.[13] The proof that this translation is accurate is sketched in §A.

Definition 4.1 *(Standard Translation).* Let τ be a $\mathcal{L}^{1M}$-term, and let $s,t \in \mathsf{SVAR}$. The ***standard translation of*** τ ***in*** $\langle s,t\rangle$, $\mathsf{st}_{s,t}(\tau)$, is a $\mathcal{L}^{2S}$-term defined as follows:

$$\mathsf{st}_{s,t}(\tau) = \begin{cases} c(t) & \text{if } \tau = c \text{ where } c \in \mathsf{CON} \\ x & \text{if } \tau = x \text{ where } x \in \mathsf{VAR}. \end{cases}$$

[13] See e.g., Blackburn et al. [2001].

Let φ be a $\mathcal{L}^{1M}_{\Pi}$-formula, and let $s, t \in$ SVAR. The ***standard translation of*** φ ***in*** $\langle s, t\rangle$, $\mathsf{ST}_{s,t}(\varphi)$, is a $\mathcal{L}^{2S}$-formula defined recursively:

$$
\begin{aligned}
\mathsf{ST}_{s,t}(P^n(\tau_1,\ldots,\tau_n)) &= P^n(\mathsf{st}_{s,t}(\tau_1),\ldots,\mathsf{st}_{s,t}(\tau_n);t)\\
\mathsf{ST}_{s,t}(\tau_1 \approx \tau_2) &= \mathsf{st}_{s,t}(\tau_1) \approx \mathsf{st}_{s,t}(\tau_2)\\
\mathsf{ST}_{s,t}(\mathsf{E}(\tau)) &= \mathsf{E}(\mathsf{st}_{s,t}(\tau);t)\\
\mathsf{ST}_{s,t}(\neg\varphi) &= \neg\,\mathsf{ST}_{s,t}(\varphi)\\
\mathsf{ST}_{s,t}(\varphi \wedge \psi) &= \mathsf{ST}_{s,t}(\varphi) \wedge \mathsf{ST}_{s,t}(\psi)\\
\mathsf{ST}_{s,t}(\Box\varphi) &= \forall t' \left(\mathsf{R}(t,t') \to \mathsf{ST}_{s,t'}(\varphi)\right)\\
\mathsf{ST}_{s,t}(@\varphi) &= \mathsf{ST}_{s,s}(\varphi)\\
\mathsf{ST}_{s,t}(\forall x\,\varphi) &= \forall x\,(\mathsf{E}(x;t) \to \mathsf{ST}_{s,t}(\varphi))\\
\mathsf{ST}_{s,t}(\Pi x\,\varphi) &= \forall x\,\mathsf{ST}_{s,t}(\varphi)\,.
\end{aligned}
$$

where t' is the next state variable not occurring anywhere in $\mathsf{ST}_{s,t}(\varphi)$.

To see why the second fact holds, look more closely at how we formalize examples of cross-world predication. Each of (16)–(18) requires the presence of some 2/2-place predicate. But a cursory inspection of **Definition 4.1** reveals that no $\mathcal{L}^{1M}$-formula ever translates into a 2/2-place predicate, and indeed no translated formula would be equivalent to a simple, atomic 2/2-place formula. This argument is made more precise in §A.

The situation is a bit trickier with cases of cross-world quantification. In (19)–(20), there is no n/m-predicate where $m > 1$. Thus, the argument above does not apply. One can supply a different argument by appealing to the notion of a *bisimulation*, familiar from standard modal model theory. The details are in Kocurek [2015].

§4.3 Are We Done?

Is the two-sorted language $\mathcal{L}^{2S}$ the best solution to the problem of cross-world predication? A number of philosophers think so. For instance, Lewis [1986, pp. 13–14], Melia [2003, p. 32], and Mackay [2013] argue that the only viable solution to the problem of cross-world predication is to adopt $\mathcal{L}^{2S}$. And philosophers such as Cresswell [1990] and Melia [2003, pp. 31–32] think that we will inevitably need to use $\mathcal{L}^{2S}$ to solve the problem of cross-world quantification. If these philosophers are correct, then the ascent to $\mathcal{L}^{2S}$ is inevitable. So why delay the inevitable?

Three reasons. First, a number of philosophers are mistaken. As we'll see below, there are strictly less expressive formal languages that are capable of dealing with a number of the problems many have claimed we need the two-sorted language to solve—the ascent is not inevitable. And since we're seeking a parsimonious solution to the problems of cross-world predication and quantification, the two-sorted language is best seen as a baseline, rather than our final answer.

Second, many will feel that the two-sorted approach is unsatisfactory. As we'll discuss in §8, the two-sorted language can make distinctions that one might think aren't real distinctions. Moreover, the fact that such a language like the two-sorted language is powerful enough to express cross-world predication is hardly an insightful observation. It essentially amounts to the claim that cross-world predications express well-defined thoughts: for what can be said clearly can be said in a first-order language. It should be *obvious* that a language as powerful as the two-sorted language has the capacity to express cross-world predication.

But this observation, in itself, doesn't give us any insight into cross-world predication—it doesn't get at the *heart* of the matter. Compare this with ordinary necessity: we could have just used the two-sorted language to analyze intra-world modal claims like "It could have rained" from the start. But then we would have deprived ourselves of much of the insight that derives from the modal framework. So while we could surrender to the two-sorted language, let us seek a more minimal solution first.

§5 The Degree Approach

Here's a very simple proposal to solve to the problem of cross-world predication. Consider the following sentence:

(Height) My height could have been greater than my actual height.

Intuitively, (Tall) and (Height) say the same thing: (Height) is just a paraphrase of (Tall). So suppose we took this paraphrase at face value by including *in our object language* the ability to talk directly about heights. Then a natural formalization of (Tall) is obtained as follows:[14]

$$\exists h\ (\mathsf{Height}(\mathsf{me}, h) \wedge \Diamond \exists h'\ (\mathsf{Height}(\mathsf{me}, h') \wedge h < h')) \quad (21)$$

where Height and $<$ are predicates in the object language, and where h and h' are intended to be variables over "heights", which are explicitly included in the domain of a model.

All of the examples of cross-world predication that we've encountered thus far seem formalizable in a similar manner. For instance, (Smart) can be formalized as:

(IQ) Everyone could have had an IQ greater than their actual IQ.

$$\forall x\, \exists i\ (\mathsf{Intelligence}(x, i) \wedge \Diamond \exists i'\ (\mathsf{Intelligence}(x, i') \wedge i < i')) \quad (22)$$

where i and i' are intended to be variables over "IQs". As another example, (Polar) can be formalized as:

(Size) A polar bear could have a size greater than the size any grizzly bear could have.

$$\Diamond \exists x\ (\mathsf{Polar}(x) \wedge \exists s\ (\mathsf{Size}(x, s) \wedge \Box \forall y\, \forall s'\ (\mathsf{Grizzly}(y) \wedge \mathsf{Size}(y, s') \rightarrow s' < s))) \quad (23)$$

[14] Normally, we would need (or want) to add @ in front of all our formalizations to ensure that the shifted worlds are accessible from the actual world, so that the formalization can more accurately capture the meaning of the sentence in context. But as noted in footnote **12**, since we're assuming our starting point of evaluation is a diagonal one, we may drop @ without loss.

where s and s' are intended to be variables over "sizes". It seems as though the problem of cross-world predication is not so hard to solve after all. Call this ***the degree approach***.[15]

Many will have the feeling that this solution is at best a hack. The reason—or so I will argue—is that while this solution does seem powerful enough to deal with the problem of cross-world predication, the way in which it deals with it is unnatural.

§5.1 Formalism

Let's see how the degree approach would work. For simplicitly, let's suppose we're just dealing with one kind of degree. Generalization is straightforward. The syntax for the degree language $\mathcal{L}^{\mathsf{D}}$ is constructed from a two-sorted first-order signature, similar to the one for $\mathcal{L}^{1\mathsf{M}}$ in §3.1, except we add:

- $\mathsf{DEG} = \{h_1, h_2, h_3, \ldots\}$ (the set of ***degree variables***);

and replace PRED^n with:

- $\mathsf{PRED}^{n/m} = \left\{P_1^{n/m}, P_2^{n/m}, P_3^{n/m}, \ldots\right\}$ where $n+m \neq 0$ (the set of n/m***-place predicates***);
- $\mathsf{PRED} := \bigcup_{n,m\geqslant 0} \mathsf{PRED}^{n/m}$ (the set of ***predicates***).

In a predicate $P^{n/m}$, n is the arity of the object-sort, while m is the arity of the degree-sort. Thus, $P^{n/m}$ takes n object terms and m degree variables as arguments to count as well-formed. So, for instance, Height is a 1/1-ary predicate, while $<$ is a 0/2-ary predicate.

The set of ***formulas in*** $\mathcal{L}^{\mathsf{D}}$ or $\mathcal{L}^{\mathsf{D}}$***-formulas***, $\mathsf{FORM}^{\mathsf{D}}$, is given by:

$$\begin{aligned}\varphi ::= {} & P^{n/m}(\tau_1, \ldots, \tau_n; h_1, \ldots, h_m) \mid \tau_1 \approx \tau_2 \mid h_1 \approx h_2 \mid \mathsf{E}(\tau) \\ & \mid \neg\,\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid @\varphi \mid \forall x\,\varphi \mid \forall h\,\varphi\end{aligned}$$

where $P^{n/m} \in \mathsf{PRED}^{n/m}$, $\tau_1, \ldots, \tau_n, \tau \in \mathsf{CON} \cup \mathsf{VAR}$, $h_1, \ldots, h_m, h \in \mathsf{DEG}$, and $x \in \mathsf{VAR}$. If we want, we could add possibilist quantifiers to $\mathcal{L}^{\mathsf{D}}$ to get $\mathcal{L}^{\mathsf{D}}_{\mathsf{p}}$.

Modal models are as in §3.2, except now we add a set of *degrees* to the model.

Definition 5.1 *(Degree Models)*. An $\mathcal{L}^{\mathsf{D}}$***-model*** or ***degree model*** is an ordered tuple $\mathcal{M} = \langle W, R, D, \delta, \mathsf{Deg}, I\rangle$ where W, R, D, and δ are as before, where Deg is a set (the set of ***degrees***), and where I is as before except:

- $I(P^{n/m}, w) \subseteq D^n \times \mathsf{Deg}^m$.

Degree variable assignments are like regular variable assignments, except they also assign members of Deg to degree variables. The semantic clauses are given below:

[15] The most well-known defendents of this approach are von Stechow [1984] and Cresswell [1990, Chp. 5]. Fitting [2013] also defends a version of this approach in a language λ-abstraction instead of quantifiers.

Definition 5.2 *(Degree Satisfaction)*. The ***degree satisfaction relation*** or ***D-satisfaction relation***, $\Vdash_{\mathsf{D}}$, is defined as in **Definition 3.4** for D-models, except:

$$\begin{array}{lcl} \mathcal{M}, w, v, g \Vdash_{\mathsf{D}} P^{n/m}(\tau_1, \ldots, \tau_n; h_1, \ldots, h_m) & \Leftrightarrow & (*) \\ \mathcal{M}, w, v, g \Vdash_{\mathsf{D}} h \approx h' & \Leftrightarrow & g(h) = g(h') \\ \mathcal{M}, w, v, g \Vdash_{\mathsf{D}} \forall h\, \varphi & \Leftrightarrow & \forall d \in \mathsf{Deg}\colon \mathcal{M}, w, v, g^h_d \Vdash_{\mathsf{D}} \varphi \end{array}$$

where $(*)$ is:

$(*)$ $\quad \left\langle [\![\tau_1]\!]^{\mathcal{M},w,v,g}, \ldots, [\![\tau_n]\!]^{\mathcal{M},w,v,g}; g(h_1), \ldots, g(h_m)\right\rangle \in I(P^{n/m}, v).$

§5.2 Ontological Commitment

One problem philosophers have with the degree approach is its "ontological commitment" to degrees like heights, sizes, and so forth. Perhaps this is fine for quantitative measures, but many find it harder to accept the existence of degrees in examples like:

(Happy) Everyone could have been as happy as they could possibly be.

(Difficult) Our homework was as difficult as possible.

(Way) Any way you could be would be better than the way I am.

Respectively, we'd need to accept the existence of degrees of happiness (or a quantity of "hedons"), levels of difficulty, and degrees of "goodness amongst ways of being". The degree approach puts us in an awkward position insofar as it forces us to take a stand with respect to various ontological questions that it seems logic should remain neutral on.

Now, it's unclear that this is a real problem for two reasons. First, it's not obvious that such a commitment is very costly. Since ordinary discourse makes use of locutions such as "my happiness" or "the level of difficulty", perhaps such ontological commitment cannot be avoided anyway.[16] If this is right, then this ontological cost is just a sunk cost.

Second, even if we grant that this ontological commitment would be costly, it's not obvious that adopting the degree approach really requires such commitments. Consider an analogous debate. Many linguists and philosophers now argue that English has at least the full expressive power of a language with explicit quantification over worlds and times.[17] Let's suppose this is correct for a moment. Even so, it's not clear that English speakers are therefore committed to modal realism. True, if one isn't a modal realist, it might be difficult to account for the success of possible world semantics given this linguistic fact. But perhaps the opponent of modal realism could explain this success after careful scrutiny. Similarly, even though the degree approach invokes explicit quantification over degrees, it does not yet follow that adopting it would commit us to the existence of degrees. Whatever maneuver the opponent of modal realism makes to avoid ontological commitment to

[16] For instance, Priest [2005, p. 123] and Fitting [2013, p. 4] argue this.

[17] See Partee [1973, 1984], Cresswell [1990], Stone [1997], Kratzer [1998], King [2003], Schlenker [2006], and Schaffer [2012] for a discussion.

possible worlds, one could imagine making the analogous maneuver to avoid ontological commitment to degrees.

§5.3 Problems with the Degree Approach

With all that said, the degree approach faces five problems (two technical, three conceptual) that make finding a different solution to the problem of cross-world predication worthwhile.

Problem 1: It doesn't solve the problem of cross-world quantification. Consider (Rich) again. The problem we faced when trying to express it in $\mathcal{L}^{1M}$ was that the most natural attempt to formalize it has the quantifier ranging over the domain of the wrong world. But (one can show that) adding degrees doesn't solve that problem. The defender of degrees could respond by arguing that cross-world predication and cross-world quantification are simply distinct phenomena that require distinct solutions. But there are combined solutions to both problems. So if we can kill two birds with one stone, why waste stones?

Problem 2: It only works for cross-world comparisons. Every example of cross-world predication we've encountered so far seems to be an instance of cross-world *comparison*.[18] Thus, it's tempting to think that somehow an approach to cross-world predication should (like the degree approach) make essential use of comparative notions.

But for one thing, the claim that all natural instances of cross-world predication in English are instances of cross-world comparisons is a *linguistic* thesis. However plausible this thesis is, it's far from clear that *logic* should take a stand on it. Moreover, when we look at other forms of modality, this thesis seems plainly false. For instance, an adequate solution to the problem of cross-world predication should be able to analyze ***cross-time predication*** as well. After all, each of our examples of cross-world predication have natural cross-time counterparts, so all the problems that arise for the former arise for the latter. And yet, cross-time predication doesn't always involve a comparison. Consider for example:[19]

(Cherish) I will cherish the person I once was.

Semi-formally, this sentence says that, where t_1 is the present, there's a future time $t_2 > t_1$ and a past time $t_0 < t_1$ such that I in t_2 cherish myself in t_0. But it's unclear how the degree approach could possibly formalize (Cherish) by making use of degrees. It's not as though

[18]Several authors have gone so far as to use the term "cross-world comparatives" as opposed to "cross-world predication", e.g., Lewis [1973, p. 436] and Cantwell [1995]. This is misleading, though, since many cross-world predications don't involve comparat*ives*—e.g., "I could have sat between where Russ and Matt are actually sitting" has no comparative in it. But still, such examples still seem to involve some kind of compar*ison* (in this case, comparing *locations*), and are still formalizable on the degree approach.

[19]Sider [2001, p. 26] has argued that sentences like "Some American philosopher admire some ancient Greek philosopher" express cross-time predications of a similar sort that can't be expressed in first-order temporal logic. The force of this example, however, depends on how we set up first-order temporal logic. If we allow "tenseless" quantifiers (Π and Σ), and if we allow extensions of predicates to contain non-existents, then the sense can easily be formalized as $\Sigma x\, \Sigma y\; (\mathsf{AmPhil}(x) \wedge \mathsf{P}\, \mathsf{GreekPhil}(x) \wedge \mathsf{Admire}(x, y))$. (In defense of Sider, he was in particular discussing presentism, which would most likely not allow for either of these in their preferred temporal logic.) By contrast, (Cherish) is problematic even allowing for tenseless quantifiers and non-existent objects in extensions of predicates.

the object of my cherish is a "degree of youth", for example; I can cherish the person I once was without cherishing the person John once was *at the same age*.

Similarly, when discussing fiction, it seems natural to say:

(Admire) I admire Napoleon from *War and Peace*.[20]

Semi-formally, this says that I (in actuality) admire the Napoleon as in *War and Peace*. Note that this sentence can't just be seen as relating me to a non-existent object: I might admire Napoleon from *War and Peace* without admiring the (now non-existent) Napoleon. Again, it's not clear what the degree approach could appeal to to formalize this sentence. So even if the degree approach works for cross-*world* predication, it can't extend beyond that.

Problem 3: It isn't parsimonious. The degree approach isn't a minimal solution, since it now allows us to express new non-cross-world sentences like "There is an uninstantiated height":

$$\exists h \, \forall x \; \neg \, \mathsf{Height}(x; h). \tag{24}$$

Now, no one is saying this is an inherently bad feature of the degree approach. For some purposes, the ability to express sentences like this might be very important. But our focus is on cross-world predication and quantification: we want to know what's the minimal extension of $\mathcal{L}^{1\mathsf{M}}$ we need to overcome these particular expressive limitations. And the fact that the degree approach can express new non-cross-world sentences suggests that we haven't achieved this goal. So as far as minimality goes, it looks like we can do better.

Problem 4: The presence of degrees seems to be inessential. When I think about the claim "x in w is happier than y is in v", I *could* understand this as claiming that there's a quantity of hedons associated with x in w, and another quantity of hedons associated with y in v, and those two quantities stand in some relation (the greater-than relation). But this seems to be a rather roundabout way of claiming that x and w stand together in some relation to y and v (the happier-than relation). Why should we need to factor our reasoning through *other* objects first in order to understand the relation that x and w stand in to y and v? Why not just reason about this relationship directly?[21]

Problem 5: We still have to worry about ontological commitment. Though we recognized earlier that the debate over ontological commitment is far from settled, some may still not be convinced that the degree approach isn't ontologically committing. Wouldn't it be nice if we could just avoid the debate altogether? Given that such technical machinery brings so much philosophical contention, it would be preferable to find a more conservative solution that is neutral between various philosophical theories.

[20]Based on a sentence from Button [2012, p. 246].

[21]A similar argument is made by Lewis [1986, p. 13].

§6 The Function Approach

In giving semi-formal characterizations of the truth conditions of cross-world predications, I've made use of locutions such as "Arc in w" and "the polar bear in v". And (if I've done my job properly), these semi-formalizations are accurate precisifications of these ordinary English sentences. Thus, a natural solution to the problem of cross-world predication is to mirror these locutions directly in the object language itself.

One way to implement this fuzzy idea is to add in function symbols whose interpretation sends objects "as they are in one world" to objects "as they are in another world". A common candidate in the literature is an "actually" function $\blacktriangleleft$.[22] Thus, the denotation of $\blacktriangleleft \tau$ will be the denotation of τ "as it actually is". Thus, the source of cross-world predication is in the denotation of terms. Call this ***the function approach***.

As we'll see, the function approach avoids many of the worries raised by the degree approach. However, as we'll also see, the function approach is not general enough to solve the problems of cross-world predication and cross-world quantification in full.

§6.1 Formalism

Let's examine how the function approach is supposed to work.[23] The signature for our functional language $\mathcal{L}^{\mathsf{F}}$ is exactly the one for $\mathcal{L}^{1\mathsf{M}}$ in §3.1, with CON, VAR, and PRED^n. The difference now is that we also have a unary function symbol $\blacktriangleleft$ such that if τ is a term, so is $\blacktriangleleft \tau$. Thus, the set of ***terms in $\mathcal{L}^{\mathsf{F}}$*** or ***$\mathcal{L}^{\mathsf{F}}$-terms***, $\mathsf{TERM}^{\mathsf{F}}$, is defined recursively:

$$\tau ::= x \mid c \mid \blacktriangleleft \tau.$$

The set of ***formulas in $\mathcal{L}^{\mathsf{F}}$*** or ***$\mathcal{L}^{\mathsf{F}}$-formulas***, $\mathsf{FORM}^{\mathsf{F}}$, is defined recursively as in §3.1:

$$\varphi ::= P^n(\tau_1, \ldots, \tau_n) \mid \tau_1 \approx \tau_2 \mid \mathsf{E}(\tau_1) \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid @\varphi \mid \forall x\, \varphi.$$

Again, $\mathcal{L}^{\mathsf{F}}_{\mathsf{p}}$ is the result of adding Π to $\mathcal{L}^{\mathsf{F}}$.

As for our models, we need to make one important change.[24]

Definition 6.1 *(Cross-world Models).* A ***cw-model*** is a tuple $\mathcal{M} = \langle W, R, D, \delta, I\rangle$ where W, R, D, and δ are as in **Definition 3.1**, and I is such that:

- for each $c \in \mathsf{CON}$, $I(c, w) \in D$;
- for each $P^n \in \mathsf{PRED}^n$, $I(P^n, w) \subseteq (D \times W)^n$.

[22] E.g., Milne [1992], Cantwell [1995], and Forbes [1994].

[23] The following formalization is, for the most part, based on Forbes [1994] and Wehmeier [2012]. Milne [1992] and Cantwell [1995] both incorporate degrees into their solutions by having a unary degree function symbol $\mathtt{deg}$ in the language. Then cross-world predications involve terms of the form $\blacktriangleleft \mathtt{deg}(\tau)$. For instance, (Tall) just becomes $\Diamond(\mathtt{deg}(\mathtt{me}) >_{\mathsf{Taller}} \blacktriangleleft \mathtt{deg}(\mathtt{me}))$. The function approach as presented here, by contrast, does not use degrees, as they're not needed to solve the problem of cross-world predication; nor are they helpful in avoiding the objections that follow.

[24] This definition is a slightly more general version of the definition in Wehmeier [2012, p. 109].

The crucial difference is that the extension of predicates in cw-models are subsets of $(D \times W)^n$. The idea is that the ordered pair $\langle \text{Arc}, w \rangle$ will represent "me as I am in w". Thus, when we say something like (Tall), what we're saying is that, in some possible world v, me as I am in v is taller than me as I actually am in w—that is, $\langle \text{Arc}, v \rangle$ bears the taller-than relation to $\langle \text{Arc}, w \rangle$.

For many cross-world predicates, the relativity of their extensions to a world is redundant. For example, if $\langle \text{Arc}, v \rangle$ is taller than $\langle \text{Arc}, w \rangle$ in one world, then $\langle \text{Arc}, v \rangle$ is taller than $\langle \text{Arc}, w \rangle$ in all worlds. For some cross-world predicates, however, this need not be the case. Suppose you and I are discussing music in 1980, and I say "By today's standards, the Maria Callas of 1951 sounds better than the Maria Callas of 1954". Later on, in 2014, we're discussing how much the musical tastes of society have changed, and I say "By today's standards, the Maria Callas of 1954 sounds better than the Maria Callas of 1951". *Prima facie*, the first sentence is true in 1980, but not in 2014. If that's so, then "sounds better than" will have a time-relative extension. Now, I don't wish to settle whether this is actually the best way to model such sentences. Still, for the sake of generality, we keep the full relativity to worlds in what follows. If the extension of a predicate P is not relative to worlds, we may write "$I(P)$" instead of "$I(P, w)$".

Definition 6.2 *(Function Denotation)*. Let τ be a term, let $\mathcal{M}$ be a cw-model, let $w, v \in W$, and let $g \in \mathsf{VA}(\mathcal{M})$. The ***denotation of*** τ ***at*** $\langle \mathcal{M}, w, v, g \rangle$, $[\![\tau]\!]^{\mathcal{M},w,v,g}$, is defined as follows (where $(\langle a, w \rangle)_{\mathsf{obj}} := a$):

$$[\![\tau]\!]^{\mathcal{M},w,v,g} = \begin{cases} \langle I(\tau, v), v \rangle & \text{if } \tau \in \mathsf{CON} \\ \langle g(\tau), v \rangle & \text{if } \tau \in \mathsf{VAR} \\ \left\langle \left([\![\sigma]\!]^{\mathcal{M},w,v,g} \right)_{\mathsf{obj}}, w \right\rangle & \text{if } \tau = \blacktriangleleft \sigma. \end{cases}$$

One should not confuse $[\![\blacktriangleleft \tau]\!]^{\mathcal{M},w,v,g}$ with $[\![\tau]\!]^{\mathcal{M},w,w,g}$. That is, the denotation of $\blacktriangleleft \tau$ is *not* the *actual* denotation of τ. Rather, it is the denotation of τ*-as-it-actually-is*. For example, suppose p is a constant for "the president". Say in w the president is Obama, while in v the president is Clinton. Then $[\![\blacktriangleleft p]\!]^{\mathcal{M},w,v,g} = \langle \text{Clinton}, w \rangle$, but $[\![p]\!]^{\mathcal{M},w,w,g} = \langle \text{Obama}, w \rangle$. That is, in $\langle \mathcal{M}, w, v, g \rangle$, $\blacktriangleleft p$ is v's president, as she is in w; but in $\langle \mathcal{M}, w, w, g \rangle$, p is w's president, as he is in w.

Definition 6.3 *(Function Satisfaction)*. The ***function satisfaction relation*** or the ***F-satisfaction relation***, $\Vdash_{\mathsf{F}}$, is defined as in **Definition 3.4** for cw-models $\mathcal{M}$ with the following modified clauses:

$$\begin{aligned} \mathcal{M}, w, v, g \Vdash_{\mathsf{F}} P^n(\tau_1, \ldots, \tau_n) \quad &\Leftrightarrow \quad \left\langle [\![\tau_1]\!]^{\mathcal{M},w,v,g}, \ldots, [\![\tau_n]\!]^{\mathcal{M},w,v,g} \right\rangle \in I(P, v) \\ \mathcal{M}, w, v, g \Vdash_{\mathsf{F}} \tau \approx \sigma \quad &\Leftrightarrow \quad \left([\![\tau]\!]^{\mathcal{M},w,v,g} \right)_{\mathsf{obj}} = \left([\![\sigma]\!]^{\mathcal{M},w,v,g} \right)_{\mathsf{obj}} \\ \mathcal{M}, w, v, g \Vdash_{\mathsf{F}} \mathsf{E}(\tau) \quad &\Leftrightarrow \quad \left([\![\tau]\!]^{\mathcal{M},w,v,g} \right)_{\mathsf{obj}} \in \delta(w). \end{aligned}$$

Notice $\tau_1 \approx \tau_2$ only requires that τ_1 and τ_2 denote the same *object*, not the same *object-world pair*.[25]

§6.2 Expressive Limitations

To see how the formalism works, let's check that

$$\Diamond \mathsf{Taller}(\mathsf{me}, \blacktriangleleft \mathsf{me}) \tag{25}$$

is an accurate way to formalize (Tall).

$$\begin{aligned} \mathcal{M}, w, w, g \Vdash_{\mathsf{F}} \Diamond \mathsf{Taller}(\mathsf{me}, \blacktriangleleft \mathsf{me}) \quad &\Leftrightarrow \quad \exists v \in R[w]\colon\ \mathcal{M}, w, v, g \Vdash_{\mathsf{F}} \mathsf{Taller}(\mathsf{me}, \blacktriangleleft \mathsf{me}) \\ &\Leftrightarrow \quad \exists v \in R[w]\colon\ \langle [\![\mathsf{me}]\!]^{w,v}, [\![\blacktriangleleft \mathsf{me}]\!]^{w,v} \rangle \in I(\mathsf{Taller}) \\ &\Leftrightarrow \quad \exists v \in R[w]\colon\ \langle [\![\mathsf{me}]\!]^{w,v}, \langle ([\![\mathsf{me}]\!]^{w,v})_{\mathsf{obj}}, w \rangle \rangle \in I(\mathsf{Taller}) \\ &\Leftrightarrow \quad \exists v \in R[w]\colon\ \langle \langle \mathrm{Arc}, v \rangle, \langle \mathrm{Arc}, w \rangle \rangle \in I(\mathsf{Taller}) \end{aligned}$$

which is exactly what we want. Thus, we'll naturally be able to formalize (Tall) as (25).

Furthermore, as one can verify, we can formalize (Smart) as:

$$\forall x\, \Diamond \mathsf{Smarter}(x, \blacktriangleleft x). \tag{26}$$

Unfortunately, (Polar) will resist formalization still. Consider the natural formalization:

$$\Diamond \exists x\ (\mathsf{Polar}(x) \wedge \Box \forall y\ (\mathsf{Grizzly}(y) \to \mathsf{Bigger}(\blacktriangleleft x, y)))\,. \tag{27}$$

This formalization won't do, since $\blacktriangleleft x$ picks out the polar bear-in-*actuality*, not in the new world we shifted to with $\Diamond$. Thus, (27) doesn't quite capture (Polar).

The problem is that the world considered as actual needs to be shiftable, so that $\blacktriangleleft x$ above can shift in denotation with $\Diamond$. One quick fix, then, would be to add a diagonalization operator $\downarrow$ with the following truth condition:[26]

$$\mathcal{M}, w, v, g \Vdash_{\mathsf{F}} \downarrow\varphi \quad \Leftrightarrow \quad \mathcal{M}, v, v, g \Vdash_{\mathsf{F}} \varphi.$$

Thus, $\downarrow$ "resets" the actual world to be the world of evaluation. By adding such an operator, we can shift the denotation of $\blacktriangleleft \tau$ by placing it inside the scope of a $\downarrow$ that is itself inside the scope of a modal operator. With this extra expressive power at hand, (Polar) would be formalized correctly as:

$$\Diamond\downarrow\exists x\ (\mathsf{Polar}(x) \wedge \Box \forall y\ (\mathsf{Grizzly}(y) \to \mathsf{Bigger}(\blacktriangleleft x, y)))\,. \tag{28}$$

But this only pushes the problem one step back. For if we reset the actual world, and then later, inside the scope of more modals we need to refer back to the original actual world before the reset, we're hosed. Thus, consider the sentence:[27]

[25]For a justification of this choice, see footnote **40** in §8.

[26]This appears, e.g., in Lewis [1973, p. 437], although Lewis uses "†" instead of "↓".

[27]Here's a similar example from Kratzer [2007] using different modalities: "Whenever it snowed, some local person dreamed that it snowed more than it actually did, and that the local weather channel erroneously reported that it had snowed less, but still more than it snowed in reality."

(Polar*) There is a polar bear that could be bigger than any grizzly bear could be if the grizzly bear were fatter than the polar bear really is.

Semi-formally, this is true just in case there is a polar bear in the actual world w such that in some possible world v, for every world u, and for every grizzly bear in u, if that grizzly bear in u is fatter than the polar bear is in w, the polar bear in v is bigger than the grizzly bear in u. Note that this can't be captured just using diagonalization. For

$$\exists x\ (\mathsf{Polar}(x) \wedge \Diamond{\downarrow}\Box\forall y\ ((\mathsf{Grizzly}(y) \wedge \mathsf{Fatter}(y, \blacktriangleleft x)) \rightarrow \mathsf{Bigger}(\blacktriangleleft x, y))) \tag{29}$$

fails to capture the intended meaning, since this says that the polar bear as it is in some possibility is bigger than any grizzly bear that's fatter than it is in that possibility. That is, since both instances of $\blacktriangleleft x$ are in the scope of $\downarrow$, their denotations are the same, even though we want the first $\blacktriangleleft x$ to refer to the polar bear *as it really is*, not as it is in this new possibility.

§6.3 Additional Quantifiers

Another issue that the function approach faces is that it doesn't solve the problem of cross-world quantification. Since neither the degree approach nor the function approach have been able to solve this problem by itself, however, we ought to consider some possible solutions that could be added to these approaches. One possible fix would be to add a quantifier $\forall_{@}$ such that:

$$\mathcal{M}, w, v, g \Vdash \forall_{@} x\, \varphi \quad \Leftrightarrow \quad \forall a \in \delta(w)\colon\ \mathcal{M}, w, v, g^{x}_{a} \Vdash \varphi.$$

Then we could express (Rich) as:

$$\Diamond\forall_{@} x\ (@\mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x))\,. \tag{30}$$

But it's not hard to see that this won't help us express (Rich*). For instance,

$$\Box\Diamond\forall_{@} x\ (@\mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x)) \tag{31}$$

doesn't have the right truth conditions, as the $\forall_{@}$ takes us back a world too far. Of course, $\downarrow$ could help us in this case:

$$\Box{\downarrow}\Diamond\forall_{@} x\ (@\mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x))\,. \tag{32}$$

But just as before, $\downarrow$ by itself won't be enough to handle more complicated cases where we need to keep our original world considered as actual. For instance, the most natural formalization of:

(Rich**) Necessarily, the rich could have all been millionaires if they were poor in reality.

is this:

$$\Box{\downarrow}\Diamond\forall_{@} x\ \Big(@\mathsf{Rich}(x) \wedge \underline{@\mathsf{Poor}(x)} \rightarrow \mathsf{Millionaire}(x)\Big)\,. \tag{33}$$

The problem is that we want the underlined @ to go all the way back to reality (i.e., the original actual world). But since it's under the scope of $\downarrow$, this is no longer possible. If we had removed $\downarrow$, the @ without an underline would then go back to the starting world, whereas we wanted @ to go to the world we shifted to by $\Box$.

Another possible approach, taken by Bricker [1989], is to introduce *second-order* quantifiers into the language. Doing so will allow us to talk rigidly about a *group* of objects at different possible worlds. Thus, using $X, Y, Z, \ldots$ for second-order variables, (Rich) could be formalized as:

$$\exists X\ (\forall x\ (Xx \leftrightarrow \mathsf{Rich}(x)) \land \Diamond\Pi x\ (Xx \rightarrow \mathsf{Poor}(x))) \tag{34}$$

and (Rich*) just as the necessitation of (34):

$$\Box\exists X\ (\forall x\ (Xx \leftrightarrow \mathsf{Rich}(x)) \land \Diamond\Pi x\ (Xx \rightarrow \mathsf{Poor}(x)))\,. \tag{35}$$

In general, the second-order approach gets around the problem of cross-world quantification. However, there are a few problems with this approach.

Problem 1: It requires possibilist quantification. If we had replaced Πx with $\forall x$ in (34), for instance, we would obtain an incorrect reading (because again, we want to quantify over every person in the actual world, not every person in the world we shifted to). Perhaps this just means we must abandon actualist quantification. But it would be nice if we could find a solution that doesn't force us to take a stand on this philosophically controversial issue.

Problem 2: It is not parsimonious. Like the degree approach, the second-order approach allows us to express new non-cross-world sentences like "Some critics only admire one another".[28] And while it's true that we may want to eventually obtain this expressive power, in terms of seeking a minimal solution to our problems, second-order logic goes too far.

Problem 3: It treats the problems of cross-world predication and quantification as unrelated. By itself, the second-order approach does nothing to solve the problem of cross-world predication, though it can effectively solve the problem of cross-world quantification. If one combines the second-order approach with, say, the degree approach, it would appear as though the problems are just completely separate problems with completely separate solutions. But the approach presented in the next section can take care of both problems at once. The second-order approach, in this respect, isn't a terribly illuminating solution to the problems we've started with.

§7 The Hybrid Solution

Despite its fallbacks, the function approach seems to come close to what we want. It just needs to be generalized to overcome the limitations above.

[28] From Boolos [1984, p. 432].

Consider the problem it faced with regards to (Polar*). The problem was that, mid-sentence, we needed to remember what our starting world was so that we could refer back to it; but, because of $\downarrow$, we had irretrievably reset that world and had no way of getting it back. Similarly, consider the problem faced with regards to (Rich*). The problem was that, after shifting twice, we needed a way to remember which world we shifted to the first time so that we could go back and quantify over the domain of that world; but @ could only take us to the world considered as actual, which was too far back.

In both cases, it seems like the function approach suffers from a memory problem. As it stands, it doesn't have a fully general method for remembering exactly which worlds we started with or shifted to. If we just had a way for our semantics to remember these worlds, these problems could be avoided.

But there's already a well-studied logic that has this capability: it's called ***hybrid logic***. In hybrid logic, one adds state variables like those in the two-sorted language, but instead of adding state quantifiers, one relativizes @ and $\downarrow$ to state variables. Informally, $@_s$ behaves like @, except it takes you back to whatever world s picks out (rather than the world considered as actual); and $\downarrow s.$ behaves like $\downarrow$, except it sets s (rather than the world considered as actual) to be the current world of evaluation.

The intuition behind this proposal is to treat the newly added state variables as though they were "memory slots". Whenever we access a new world, we can, if we wish, save it (using $\downarrow$) to an unused memory slot. Then later, we might retrieve the information stored in one of these slots (using @) to help us obtain the right truth conditions. Thus, $\downarrow$ could be thought of as "saving" the current world of evaluation to some memory slot, while @ could be thought of as "loading" some previously saved world as the world of evaluation.

In this section, we explore this rather natural idea by expanding the function approach to a ***quantified hybrid logic***. As we'll discuss below, doing so allows one to solve the problems of cross-world predication and cross-world quantification in a fully general yet parsimonious way. Furthermore, as we'll also discuss, moving to quantified hybrid logic brings with it a number benefits, including the solution to several other problems.

§7.1 Formalism

Now to be more precise. The signature for our language $\mathcal{L}^{\mathsf{H}}$ will include CON, VAR, SVAR, and PRED. The set of ***terms in*** $\mathcal{L}^{\mathsf{H}}$ or $\mathcal{L}^{\mathsf{H}}$***-terms***, $\mathsf{TERM}^{\mathsf{H}}$, is defined recursively:

$$\tau ::= x \mid c \mid \blacktriangleleft_s \tau$$

where $x \in \mathsf{VAR}$, $c \in \mathsf{CON}$, and $s \in \mathsf{SVAR}$.

The set of ***formulas in*** $\mathcal{L}^{\mathsf{H}}$ of $\mathcal{L}^{\mathsf{H}}$***-formulas***, $\mathsf{FORM}^{\mathsf{H}}$, is defined recursively as follows:[29]

$$\varphi ::= P^n(\tau_1, \dots, \tau_n) \mid \tau_1 \approx \tau_2 \mid \mathsf{E}(\tau_1) \mid \neg \varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid @_s\varphi \mid {\downarrow}s.\varphi \mid \forall x\, \varphi$$

where $P^n \in \mathsf{PRED}^n$ for each $n > 0$, $\tau_1, \dots, \tau_n \in \mathsf{TERM}^{\mathsf{H}}$, $s \in \mathsf{SVAR}$, and $x \in \mathsf{VAR}$. Again, $\mathcal{L}^{\mathsf{H}}_{\mathsf{p}}$ will be the result of adding Π to $\mathcal{L}^{\mathsf{H}}$. Any instance of $s \in \mathsf{SVAR}$ in an $\mathcal{L}^{\mathsf{H}}$-formula that occurs in the scope of ${\downarrow}s.$ is said to be ***bound***; any other occurrence of s is ***unbound***. If every

[29] Unlike in ordinary hybrid logic, I haven't allowed for state variables to count as well-formed formulas. This is because such additional expressive power isn't necessary here. See footnote **40** in §8.

state variable is bounded in a $\mathcal{L}^{\mathsf{H}}$-formula φ, then we'll say φ is ***state-closed***; otherwise, it's ***state-open***. Finally, we'll adopt these abbreviations throughout:[30]

- $\Box_s\varphi \coloneqq \Box{\downarrow}s.\varphi$ ("for all accessible worlds s, φ")
- $\Diamond_s\varphi \coloneqq \Diamond{\downarrow}s.\varphi$ ("for some accessible world s, φ")
- $\forall_s x\,\varphi \coloneqq {\downarrow}t.@_s\forall x\,@_t\varphi$, where t occurs nowhere in φ ("for all x in world s, φ")

Models on the hybrid solution are just the cw-models from before. As for the semantics:

Definition 7.1 *(Variable Assignment)*. Let $\mathcal{M}$ be a cw-model. A ***H-variable assignment for*** $\mathcal{M}$ is a function g such that: (i) for each $x \in \mathsf{VAR}$, $g(x) \in D$; and (ii) for each $s \in \mathsf{SVAR}$, $g(s) \in W$. The set of all H-variable assignments for $\mathcal{M}$ will be $\mathsf{VA}^{\mathsf{H}}(\mathcal{M})$.

If a H-variable assignment g for $\mathcal{M}$ agrees with a H-variable assignment g' for $\mathcal{M}$ on every variable except possibly μ, then g and g' are ***μ-variants***, $g \sim_\mu g'$. The H-variable assignment $g[\mu \mapsto \alpha]$ or g^μ_α is the μ-variant of g that sends μ to α.

Definition 7.2 *(Hybrid Denotation)*. Let τ be a $\mathcal{L}^{\mathsf{H}}$-term, $\mathcal{M}$ be a cw-model, $w, v \in W$, and $g \in \mathsf{VA}^{\mathsf{H}}(\mathcal{M})$. The ***denotation of*** τ ***at*** $\langle \mathcal{M}, w, g\rangle$, $[\![\tau]\!]^{\mathcal{M},w,g}$, is defined as follows:

$$[\![\tau]\!]^{\mathcal{M},w,g} = \begin{cases} \langle I(c,w), w\rangle & \text{if } \tau = c \text{ where } c \in \mathsf{CON} \\ \langle g(x), w\rangle & \text{if } \tau = x \text{ where } x \in \mathsf{VAR} \\ \left\langle \left([\![\sigma]\!]^{\mathcal{M},w,g}\right)_{\mathsf{obj}}, g(s)\right\rangle & \text{if } \tau = \blacktriangleleft_s \sigma \text{ where } s \in \mathsf{SVAR}. \end{cases}$$

Definition 7.3 *(Hybrid Satisfaction)*. The ***hybrid satisfaction relation*** or ***H-satisfaction***, $\Vdash_{\mathsf{H}}$, is defined recursively for all cw-models $\mathcal{M} = \langle W, R, D, \delta, I\rangle$, all $w \in W$, and all $g \in \mathsf{VA}^{\mathsf{H}}(\mathcal{M})$:

$$\begin{array}{lcl}
\mathcal{M}, w, g \Vdash_{\mathsf{H}} P^n(\tau_1, \ldots, \tau_n) & \Leftrightarrow & \left\langle [\![\tau_1]\!]^{\mathcal{M},w,g}, \ldots, [\![\tau_n]\!]^{\mathcal{M},w,g}\right\rangle \in I(P^n, w) \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} \tau_1 \approx \tau_2 & \Leftrightarrow & \left([\![\tau_1]\!]^{\mathcal{M},w,g}\right)_{\mathsf{obj}} = \left([\![\tau_2]\!]^{\mathcal{M},w,g}\right)_{\mathsf{obj}} \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} \mathsf{E}(\tau) & \Leftrightarrow & \left([\![\tau]\!]^{\mathcal{M},w,g}\right)_{\mathsf{obj}} \in \delta(w) \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} \neg\varphi & \Leftrightarrow & \mathcal{M}, w, g \nVdash_{\mathsf{H}} \varphi \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} \varphi \wedge \psi & \Leftrightarrow & \mathcal{M}, w, g \Vdash_{\mathsf{H}} \varphi \text{ and } \mathcal{M}, w, g \Vdash_{\mathsf{H}} \psi \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} \Box\varphi & \Leftrightarrow & \forall v \in R[w]\colon\ \mathcal{M}, v, g \Vdash_{\mathsf{H}} \varphi \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} @_s\varphi & \Leftrightarrow & \mathcal{M}, g(s), g \Vdash_{\mathsf{H}} \varphi \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} {\downarrow}s.\varphi & \Leftrightarrow & \mathcal{M}, w, g^s_w \Vdash_{\mathsf{H}} \varphi \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} \forall x\,\varphi & \Leftrightarrow & \forall a \in \delta(w)\colon\ \mathcal{M}, w, g^x_a \Vdash_{\mathsf{H}} \varphi \\
\mathcal{M}, w, g \Vdash_{\mathsf{H}} \Pi x\,\varphi & \Leftrightarrow & \forall a \in D\colon\ \mathcal{M}, w, g^x_a \Vdash_{\mathsf{H}} \varphi.
\end{array}$$

[30]While these abbreviations may make the hybrid solution look much like the two-sorted language, there are subtle but important differences between the two approaches. This will be discussed in §8.

One might wonder why we dropped the unsubscripted ◀ and @ from the language and adopted a plain vanilla one-dimensionsal semantics. One reason is that it will simplify some of the discussion that follows. Another reason is that the extra symbols and dimension are redundant in the sense that there is no real increase in expressive power, at least over diagonal indicies. For if we had adopted a two-dimensional semantics for $\mathcal{L}^{\mathsf{H}}$, then for all φ, where $s \in \mathsf{SVAR}$ does not occur anywhere in φ, $\mathcal{M}, w, w, g \Vdash \varphi$ iff $\mathcal{M}, w, w, g \Vdash {\downarrow} s.\varphi^*$, where φ^* is the result of replacing every ◀ and @ respectively with $◀_s$ and $@_s$. So we can adopt a convention that, say, $g(s_0)$ always denotes the world considered as actual, and that we never bind s_0 when formalizing sentences. Then we can just define $◀\,\tau$ and $@\varphi$ respectively as $◀_{s_0}\,\tau$ and $@_{s_0}\varphi$. No generality is lost, but much is gained in simplicity.

Simpler quantified hybrid logics without ◀ have been previously studied by e.g., Blackburn and Marx [2002] and Areces et al. [2003], though not explicitly in the context of these problems. The version of quantified hybrid logic presented here is similar in spirit to the proposals of Forbes [1989], Cresswell [1990], and Wehmeier [2012], though it is rarely acknowledged that these other proposals are essentially hybrid logics.[31] The first two aren't designed to handle cross-world predication (only cross-world quantification) and are equivalent to the ◀-free fragment of $\mathcal{L}^{\mathsf{H}}$. The third is designed to handle cross-world predication and resembles the hybrid solution in a number of respects. One can think of the hybrid solution as a generalization of Wehmeier's framework (e.g., to allow for a non-universal accessibility relation, different arities of cross-world predicates, etc.). See §B for an elaboration of this point.

§7.2 The Solution

At last, we can see how quantified hybrid logic solves the problems of cross-world predication, cross-world quantification, and the other related problems we saw in §1. I assume for concreteness that all object quantifiers are actualist, though it doesn't really matter which quantifier one prefers to use (another benefit of the hybrid solution!).

Cross-world Predication & Quantification. Here again are some of the key examples of cross-world predication and quantification that have appeared throughout the paper:

(Tall) I could have been taller than I actually am.

$$ {\downarrow} s.\Diamond \mathsf{Taller}(\mathsf{me}, ◀_s \mathsf{me}) \tag{36} $$

(Smart) Everyone could have been smarter than they actually are.

$$ {\downarrow} s.\forall x\, \Diamond \mathsf{Smarter}(x, ◀_s x) \tag{37} $$

(Polar) A polar bear could be bigger than a grizzly bear could be.

$$ \Diamond_s \exists x\, (\mathsf{Polar}(x) \wedge \Box \forall y\, (\mathsf{Grizzly}(y) \rightarrow \mathsf{Bigger}(◀_s x, y))) \tag{38} $$

[31] After completing this paper, it came to my attention that Yanovich [2015] has explicitly used quantified hybrid logic, in its familiar form, for addressing the problem of cross-world quantification. The problem of cross-world predication is not addressed in his paper however.

(Polar*) There is a polar bear that could be bigger than any grizzly bear could be if the grizzly bear were fatter than the polar bear really is.

$$\downarrow s.\exists x\ (\mathsf{Polar}(x) \wedge \Diamond_t \Box \forall y\ ((\mathsf{Grizzly}(y) \wedge \mathsf{Fatter}(y, \blacktriangleleft_s x)) \rightarrow \mathsf{Bigger}(\blacktriangleleft_t x, y))) \quad (39)$$

(Rich) The rich could have all been poor.

$$\downarrow s.\Diamond \forall_s x\ (@_s \mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x)) \quad (40)$$

(Rich*) Necessarily, the rich could have all been poor.

$$\Box_s \Diamond \forall_s x\ (@_s \mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x)) \quad (41)$$

(Rich**) Necessarily, the rich could have all been millionaires if they were all poor in reality.

$$\downarrow t.\Box_s \Diamond \forall_s x\ ((@_s \mathsf{Rich}(x) \wedge @_t \mathsf{Poor}(x)) \rightarrow \mathsf{Millionaire}(x)) \quad (42)$$

To illustrate how the formal system works, let's check that (Rich*) gets the right results. I leave it to the reader to check the other examples.

$$\begin{aligned}
& \mathcal{M}, w, g \Vdash_{\mathsf{H}} \Box_s \Diamond \forall_s x\ (@_s \mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x)) \\
\Leftrightarrow\ & \forall v \in R[w]:\ \mathcal{M}, v, g^s_v \Vdash_{\mathsf{H}} \Diamond \forall_s x\ (@_s \mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x)) \\
\Leftrightarrow\ & \forall v \in R[w]\, \exists u \in R[v]:\ \mathcal{M}, u, g^s_v \Vdash_{\mathsf{H}} \forall_s x\ (@_s \mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x)) \\
\Leftrightarrow\ & \forall v \in R[w]\, \exists u \in R[v]\, \forall a \in \delta(g^s_v(s)):\ \mathcal{M}, u, g^{s,x}_{v,a} \Vdash_{\mathsf{H}} @_s \mathsf{Rich}(x) \rightarrow \mathsf{Poor}(x) \\
\Leftrightarrow\ & \forall v \in R[w]\, \exists u \in R[v]\, \forall a \in \delta(v):\ \mathcal{M}, v, g^{s,x}_{v,a} \Vdash_{\mathsf{H}} \mathsf{Rich}(x) \Rightarrow \mathcal{M}, u, g^{s,x}_{v,a} \Vdash_{\mathsf{H}} \mathsf{Poor}(x) \\
\Leftrightarrow\ & \forall v \in R[w]\, \exists u \in R[v]\, \forall a \in \delta(v):\ \langle g^{s,x}_{v,a}(x), v \rangle \in I(\mathsf{Rich}, v) \Rightarrow \langle g(x), u \rangle \in I(\mathsf{Poor}, u) \\
\Leftrightarrow\ & \forall v \in R[w]\, \exists u \in R[v]\, \forall a \in \delta(v):\ \langle a, v \rangle \in I(\mathsf{Rich}, v) \Rightarrow \langle a, u \rangle \in I(\mathsf{Poor}, u).\ \checkmark
\end{aligned}$$

In addition to being able to solve the problems of cross-world predication and quantification, the hybrid solution can explain how the two are related. What they share in common is their essential use of ***cross-world recollection***, which involves the ability to remember worlds that have previously been encountered in the semantic evaluation process.

Take (Rich) for instance. Intuitively, what's going on in (Rich) is that, once we've already shifted to another world, we want to jump back to our starting world to quantify over its domain, and then "jump forward" to *the same* shifted world as before. While standard modal logic can't reliably jump back and forward like this,[32] hybrid logic can do this quite easily because it can remember which worlds we've encountered. Similarly, what's going on in (Polar) is that we need the denotation of a term to "reach outside" the scope of a modal to refer to a world we previously accessed. Both examples seem to crucially rely on the power of cross-world recollection obtained in $\mathcal{L}^{\mathsf{H}}$.

[32]Such limitations are discussed in Hazen [1976, p. 39].

Cross-time Predication. The hybrid solution to cross-world predication naturally extends to a solution to cross-time predication in a number of respects. For one thing, since we don't have to deal with degrees, non-comparative sentences like (Cherish) can be easily formalized:[33]

(Cherish) I will cherish the person I once was.

$$\downarrow s.\mathsf{P}\downarrow t.@_s\mathsf{F}\,\mathsf{Cherish}(\mathsf{me},\blacktriangleleft_t\mathsf{me}). \tag{43}$$

Moreover, one might naturally wonder how to formalize sentences like:

(Fat) Toby was fatter in 1980 than William in 1982.[34]

As it stands, $\mathcal{L}^{\mathsf{H}}$ can't formalize this; but it could easily be modified by adding *state constants* (or ***nominals*** as they're called in hybrid logic) whose denotation is fixed by the model, rather than a variable assignment. In that case, $\mathcal{L}^{\mathsf{H}}$ could formalize (Fat) as:

$$\mathsf{Fatter}(\blacktriangleleft_{1980}\mathsf{toby},\blacktriangleleft_{1982}\mathsf{will}). \tag{44}$$

Thus, quantified hybrid logic is well-equipped to generalize as an account of cross-*time* predications.

Cross-fictional Predication. Consider again (Bilbo):

(Bilbo) Bilbo in *The Lord of the Rings* is taller than Thumbelina in *Thumbelina*.

While the degree approach could formalize (Bilbo) with degrees, it can also be formalized fairly easily without degrees in $\mathcal{L}^{\mathsf{H}}$:

$$\downarrow s.\,[\mathsf{LotR}]_r\,@_s\,[\mathsf{Thumb}]_t\,@_s\mathsf{Taller}(\blacktriangleleft_r\mathsf{bilbo},\blacktriangleleft_t\mathsf{thumb}). \tag{45}$$

And unsurprisingly, the hybrid solution can account for a number of other mixed-world sentences like:

(Bilbo*) I am taller than Bilbo in *The Lord of the Rings*.

$$\downarrow s.\,[\mathsf{LotR}]\downarrow t.@_s\mathsf{Taller}(\mathsf{me},\blacktriangleleft_t\mathsf{bilbo}). \tag{46}$$

Now, the hybrid solution won't by itself stand as a complete account of fictional discourse. For one thing, there's the possibility of impossible fictions, which has yet to be discussed. But there's a more fundamental problem. Consider one problematic sentence for $\mathcal{L}^{\mathsf{H}}$:

(Fame) Sherlock Holmes from the Holmes stories is more famous than any actual detective.

[33] Where F is the "it will be the case that" operator, and P is the "it was the case that" operator.
[34] From Butterfield and Stirling [1987].

Suppose we have a predicate Famous for "is more famous than". Can (Fame) be expressed in $\mathcal{L}^{\mathsf{H}}$? The best attempt seems to be the following:

$$\downarrow s.\forall x \; (\mathsf{Detective}(x) \rightarrow [\mathsf{Holmes}] \downarrow t.@_s \mathsf{Famous}(\blacktriangleleft_t \mathsf{sherlock}, x)) \,. \tag{47}$$

The problem with this formalization, however, is that it claims that *every* precisification of the Holmes stories yields a Sherlock more famous than any actual detective. That is, (47) claims that every version of Sherlock Holmes compatible with the Holmes stories is more famous than any actual detective. But this seems much stronger than (Fame)—the claim that a version of Holmes that has a freckle on his back is more famous than any actual detective is quite odd, if not dubious.

However, there is a natural fix. Suppose we replace fictional worlds in our models with fictional *possibilities*—that is, possible states which do not determinately decide on every sentence whether that sentence is true or false.[35] In doing so, we could reinterpret [Holmes] so that this operator (instead of checking every fictional world compatible with the Holmes stories) simply shifts the world of evaluation to the fictional possibility of the Holmes stories. In making this move, (47) doesn't have the problem stated above—only one Sherlock Holmes is picked out by the use of $\blacktriangleleft_t$ sherlock above. Developing this idea in detail is quite tricky, however, and will have to wait for another time.

Counteridenticals. Recall the problematic counteridenticals:

(Horse) If I were you, I wouldn't bet on that horse.

(Bold) I would be bolder if I weren't me.

(Horse*) If you were me, I wouldn't bet on that horse.

(Sally) If I were you and you were Sally, I would be Sally.

(Star) If I were you and you were me, I would be a rock star and you wouldn't.

On the degree approach, one could formalize these by postulating the existence of "shoes" (or rather, positions). But in quantified hybrid logic, we can get away with just the Shoes predicate (from §1) in the following manner:

$$\downarrow s.(\mathsf{Shoes}(\mathsf{me}, \blacktriangleleft_s \mathsf{you}) \mathrel{\Box\!\!\rightarrow} \neg\, \mathsf{Bet}(\mathsf{me}, \mathsf{horse})) \tag{48}$$

$$\downarrow s.(\neg\, \mathsf{Shoes}(\mathsf{me}, \blacktriangleleft_s \mathsf{me}) \mathrel{\Box\!\!\rightarrow} \mathsf{Bolder}(\mathsf{me}, \blacktriangleleft_s \mathsf{me})) \tag{49}$$

$$\downarrow s.(\mathsf{Shoes}(\mathsf{you}, \blacktriangleleft_s \mathsf{me}) \mathrel{\Box\!\!\rightarrow} \neg\, \mathsf{Bet}(\mathsf{me}, \mathsf{horse})) \tag{50}$$

$$\downarrow s.((\mathsf{Shoes}(\mathsf{me}, \blacktriangleleft_s \mathsf{you}) \wedge \mathsf{Shoes}(\mathsf{you}, \blacktriangleleft \mathsf{sally})) \mathrel{\Box\!\!\rightarrow} \mathsf{Shoes}(\mathsf{me}, \blacktriangleleft_s \mathsf{sally})) \tag{51}$$

$$\downarrow s.((\mathsf{Shoes}(\mathsf{me}, \blacktriangleleft_s \mathsf{you}) \wedge \mathsf{Shoes}(\mathsf{you}, \blacktriangleleft_s \mathsf{me})) \mathrel{\Box\!\!\rightarrow} (\mathsf{RockStar}(\mathsf{me}) \wedge \neg\, \mathsf{RockStar}(\mathsf{you}))). \tag{52}$$

Notice that all the problems that arose for the naïve view of counteridenticals, which interpreted the antecedents as literally involving identity statements, don't arise for the hybrid solution. Just as (Bold) isn't trivial, (49) isn't trivial; just as (Horse*) isn't equivalent to

[35] Such possibilities could be *situations* as in Kratzer [2007], though they need not be cashed out in that particular way.

(Horse), (50) isn't equivalent to (48); just as (Sally) isn't necessarily true, (51) isn't necessarily true; and just as (Star) is possibly true, (52) is possibly true.

A similar analysis could be given for beliefs about identity. Consider an example:

(KR) Biron thinks that Katherine is Rosaline.[36]

If names are rigid, this clearly can't mean that Biron believes the negation of an identity statement that Katherine is identical to Rosaline, since that identity statement is true in all worlds. (One can also illustrate it's not an identity statement by showing that all the standard properties of identity fail, as with counteridenticals.) Thus, we need another way to formalize (KR).

The solution can be given in terms of "roles". The idea is that in all of Biron's belief worlds, Katherine is playing a certain role in Biron's cognitive life, e.g., being the person identified as Rosaline. Thus, if we had a predicate like $\mathsf{Role}(a, x, y)$, which means "x plays the same role as y for agent a", then we can formalize sentences like (KR) as:[37]

$$\downarrow s.\mathsf{B}_{\mathsf{biron}}\mathsf{Role}(\mathsf{biron}, \mathsf{kat}, \blacktriangleleft_s \mathsf{rosa}). \tag{53}$$

Again, the details will have to come at another time.[38]

Supervenience. Finally, recall the statement of supervenience:

(Super) The mental supervenes on the physical—there can't be any difference in the mental without a difference in the physical.

This can be formalized as follows:

$$\downarrow r.\Box_s \forall x\, @_r \Box_t \forall y\, @_r\, (\blacktriangleleft_s x \sim_{\mathsf{P}} \blacktriangleleft_t y \rightarrow \blacktriangleleft_s x \sim_{\mathsf{M}} \blacktriangleleft_t y)\,. \tag{54}$$

Notice that such formalization requires the hybrid solution, not just the function approach. At last, all the data is accounted for.

§8 Distinguishing $\mathcal{L}^{\mathsf{H}}$ from $\mathcal{L}^{\mathsf{2S}}$

To sum up the paper so far, we started with the following problem: is there an extension of $\mathcal{L}^{\mathsf{1M}}$ that can express every instance of cross-world predication and quantification? Of course, the two-sorted language $\mathcal{L}^{\mathsf{2S}}$ could, but we wanted to know whether something more minimal would suffice. We saw that neither adding degrees to the model nor adding a function symbol $\blacktriangleleft$ sufficed to express cross-world predication and quantification in full generality. Finally, we found that moving to quantified hybrid logic provides an elegant and parsimonious solution to the problems.

[36] From Cumming [2008, p. 529].

[37] Here, the use of $\blacktriangleleft_s$ as opposed to $\blacktriangleleft$ might be necessary, since there are powerful arguments that doxastic and epsitemic modal operators shift the actual world, e.g., Rabinowicz and Segerberg [1994].

[38] This approach is similar to the ones taken by, e.g., Aloni [2005]; Holliday and Perry [2014]. These authors could be seen as the kind of approach sketched above, except with explicit quantification over "roles", akin to the degree approach.

I would like to conclude by addressing the following concern: how minimal is quantified hybrid logic? Sure, as we've seen above, quantified hybrid logic seems to solve all the problems we've discussed so far in a natural, well-motivated manner. The two-sorted language can also solve these problems, though in §4.3, I claimed that we ought not accept the two-sorted language as the easy solution. But some readers might feel cheated: isn't the hybrid solution just the two-sorted language in disguise?

No. Just like with $\mathcal{L}^{\mathsf{1M}}$, we can translate every $\mathcal{L}^{\mathsf{H}}$-formula into a $\mathcal{L}^{\mathsf{2S}}$-formula, but not *vice versa*. Still, it's easy to see why one would be tempted to say they're notational variants, at least on a conceptual level. For one thing, the presence of state variables in the object language seems to be suspicious, and $\Box{\downarrow}s_k.$ bears a close resemblance to $\forall s_k$. Moreover, the actual difference between $\mathcal{L}^{\mathsf{H}}$ and $\mathcal{L}^{\mathsf{2S}}$ is quite subtle, and the gap between the two languages can be collapsed rather easily. But that doesn't mean these differences aren't important.

§8.1 Differences

There are three main differences between $\mathcal{L}^{\mathsf{2S}}$ and the fragment of $\mathcal{L}^{\mathsf{2S}}$ that $\mathcal{L}^{\mathsf{H}}$ characterizes (henceforth, I'll simply identify the fragment of $\mathcal{L}^{\mathsf{2S}}$ that $\mathcal{L}^{\mathsf{H}}$ characterizes with $\mathcal{L}^{\mathsf{H}}$ itself). First, $\mathcal{L}^{\mathsf{H}}$ requires that all state quantifiers be ***R-bounded***. That is, all universal state quantifiers in a $\mathcal{L}^{\mathsf{H}}$-formula must take the form $\forall t\ (\mathsf{R}(s,t) \to \cdots)$, and all existential state quantifiers must take the form $\exists t\ (\mathsf{R}(s,t) \wedge \cdots)$ (where $s \neq t$). Second, if we use actualist quantifiers, $\mathcal{L}^{\mathsf{H}}$ requires that all object quantifiers be E-bounded—that is, of the form $\forall x\ (\mathsf{E}(x;s) \to \cdots)$ and $\exists x\ (\mathsf{E}(x;s) \wedge \cdots)$. And third, $\mathcal{L}^{\mathsf{H}}$ does not allow one to build up formulas from *any* atomic formula. In particular, it bans building up complex formulas from atomic formulas of the form $\mathsf{R}(s,t)$ and $s \approx t$.[39]

Putting these restrictions together, we can succicintly state the difference between the two languages: $\mathcal{L}^{\mathsf{H}}$ can only be built from certain kinds of atomic formulas—not including those of the form $\mathsf{R}(s,t)$ and $s \approx t$—using negation, conjunction, (E-bounded) object quantifiers, and R-bounded state quantifiers. Put another way: if we assume R is universal, allow ourselves to build from any atomic formula, and use possibilist quantifiers, we get back the full two-sorted language. Thus, it's no surprise that one would suspect the difference between $\mathcal{L}^{\mathsf{H}}$ and $\mathcal{L}^{\mathsf{2S}}$ to be quite small.[40]

But the difference is still important. For one thing, even if we grant that R should be universal for metaphysical modality, it's not clear it should be universal for other kinds of modalities. An obvious example is temporal modalities, which are crucially restricted by the earlier-than relation. But for a less obvious example, it seems very plausible that fiction modal operators are not **S5** operators. For instance, R_{LotR} is certainly not reflexive

[39]It also bans other kinds of atomics if constants are non-rigid designators, and if there are no possibilist quantifiers. For instance, $\mathcal{L}^{\mathsf{H}}$ can't be built from atomics of the form $P(c(s);t,t')$ where $s \neq t$. It is possible to state precisely which atomic formulas are allowed, but it won't be necessary to go into the details here.

[40]The difference is even smaller if we either allow state variables as formulas, as is oridinarily done in hybrid logic, or if we define an identity relation $\equiv$ such that $\mathcal{M}, w, g \Vdash \tau \equiv \sigma$ iff $[\![\tau]\!]^{\mathcal{M},w,g} = [\![\sigma]\!]^{\mathcal{M},w,g}$. If either of these are added to the syntax, then we can express $\mathsf{R}(s,t)$ either as $@_s \Diamond t$ or $@_s \Diamond (c \equiv \blacktriangleleft_t c)$. We can also express $s \approx t$ either as $@_s t$ or $@_s (c \equiv \blacktriangleleft_t c)$. This is one reason why we use $\approx$ and why we don't allow state variables to also be formulas, apart from the fact that we don't need the extra expressive power of the expanded language to express the sentences we're interested in.

or symmetric, as this world is not a world compatible with *The Lord of the Rings*.[41] Thus, for fictional discourse, the R-boundedness will be an important restriction on the formulas one can construct.[42]

The fact that we can't build complex formulas from $\mathsf{R}(s,t)$ and $s \approx t$ arbitrarily in $\mathcal{L}^{\mathsf{H}}$ is perhaps even more crucial. This allows us to preserve a general feature of most modal logics: duplicating worlds preserves modal truth. That is, in most modal logics, no modal formula can distinguish between a model with two worlds that are exact duplicates of one another (having the same extensions, the same local domains, seeing the same worlds, etc.) and a model with just one of the duplicates. By contrast, a $\mathcal{L}^{\mathsf{2S}}$-formula can do this.

Here's a very simple example. Consider the following $\mathcal{L}^{\mathsf{2S}}$-formula:

$$\exists s\, \mathsf{R}(s,s). \tag{55}$$

This $\mathcal{L}^{\mathsf{2S}}$-formula isn't techincally R-bounded, since the second state variable isn't different from the first. And in fact, we can prove this isn't equivalent to a $\mathcal{L}^{\mathsf{H}}$-formula by considering two different models (where we assume w and v have the same local domains, same extensions for predicates, and same denotations for constants):

$\mathcal{M}_1$ $\mathcal{M}_2$

$\mathcal{L}^{\mathsf{2S}}$ can distinguish $\langle \mathcal{M}_1, w\rangle$ from $\langle \mathcal{M}_2, w\rangle$ by (55). However, both satisfy the exact same $\mathcal{L}^{\mathsf{H}}$-formulas. Thus, no $\mathcal{L}^{\mathsf{H}}$-formula distinguishes these two. And intuitively, there shouldn't be one: there isn't a substantial difference between these two situations.

We can also distinguish $\langle \mathcal{M}_1, w\rangle$ and $\langle \mathcal{M}_2, w\rangle$ with the $\mathcal{L}^{\mathsf{2S}}$-formula:

$$\exists t\ (\mathsf{R}(s,t) \wedge s \not\approx t)\,. \tag{56}$$

Unlike (55), (56) is R-bounded. But it also contains $s \not\approx t$, so it isn't a $\mathcal{L}^{\mathsf{H}}$-formula. And again, since $\langle \mathcal{M}_1, w\rangle$ and $\langle \mathcal{M}_2, w\rangle$ satisfy the same $\mathcal{L}^{\mathsf{H}}$-formulas, (56) isn't equivalent to a $\mathcal{L}^{\mathsf{H}}$-formula. Many more interesting examples can be generated, but one sees the general point: $\mathcal{L}^{\mathsf{2S}}$ makes distinctions where $\mathcal{L}^{\mathsf{H}}$ doesn't. And if one has principled reasons for thinking such distinctions shouldn't be made, then one has reason to refrain from adopting $\mathcal{L}^{\mathsf{2S}}$ over $\mathcal{L}^{\mathsf{H}}$.

[41] Whether it's transitive depends on how one understands fictional discourse. If R_{LotR} is understood linking worlds compatible with *The Lord of the Rings* as written in *this world*, then it's probably transitive. If R_{LotR} is understood as linking worlds compatible with *The Lord of the Rings* as written in *the world of evaluation*, then it's not, as some fictional worlds compatible with *The Lord of the Rings* won't even contain such a fiction.

The former theory seems more plausible, since it seems possible that Bilbo never existed in *The Lord of the Rings* (if Tolkien had decided to write the story differently, for example). But if that's so, then two different worlds of evaluation might disagree on which worlds are compatible with *The Lord of the Rings*. Hence, R_{LotR} would have to be interpreted as compatability with *The Lord of the Rings* as it's written in the world of evaluation, not this world. But I won't settle the matter here due to space constraints.

[42] However, if we add any global modality, then we again lose R-boundedness. Thanks to Balder ten Cate for pointing this out.

§8.2 Minimality

I've argued above that there are good reasons to not go beyond the expressive power of $\mathcal{L}^{\mathsf{H}}$. But can we be more minimal than $\mathcal{L}^{\mathsf{H}}$? That is, is there a language extending $\mathcal{L}^{\mathsf{1M}}$ that can express cross-world predication and quantification without extending $\mathcal{L}^{\mathsf{H}}$? If there is, this could motivate moving away from $\mathcal{L}^{\mathsf{H}}$ to a more minimal language.

This question seems difficult to address, but I wish to conclude by providing a tentative answer. The problem in addressing this question is that we haven't precisely defined what it takes for a formula to express a genuine example of cross-world predication or quantification. We've argued above that $\mathcal{L}^{\mathsf{H}}$ at the very least has enough expressive power to capture every instance of cross-world predication and quantification, just like $\mathcal{L}^{\mathsf{2S}}$. But just like $\mathcal{L}^{\mathsf{2S}}$, it might also add some expressive power that's not essential to solving the problems of cross-world predication and quantification. And until we know precisely which formulas are cross-world and which aren't, we won't have a principled way of deciding the matter.

To fill this gap, I will conjecture that the intuitive properties of expressing genuine cross-world predication or quantification are captured by the following definition:

Definition 8.1 *(Cross-world)*. An $\mathcal{L}^{\mathsf{H}}$-formula is a ***cross-world formula*** if every equivalent $\mathcal{L}^{\mathsf{H}}$-formula either contains an instance of $\blacktriangleleft$ or contains an object quantifier scoping over an instance of @. An $\mathcal{L}^{\mathsf{H}}$-formula is a ***non-cross-world formula*** if it's not a cross-world formula. We'll also say an $\mathcal{L}^{\mathsf{2S}}$-formula is cross-world if it's equivalent to the translation of a cross-world $\mathcal{L}^{\mathsf{H}}$-formula.

Such a conjecture is akin to the Church-Turing thesis that the correct precisification of computability is Turing computability. To refute the conjecture, one would need to provide an intuitive example of a (non-)cross-world predication or quantification that doesn't fit this definition. And such examples are not forthcoming (indeed, all of (36)–(42) are cross-world in the sense of **Definition 8.1**).[43]

But such a definition does make intuitive sense upon reflection. All cross-world predications seem to rely on "as in" locutions, which are directly captured by $\blacktriangleleft$. As for cross-world quantification, the key feature of cross-world quantification is the ability of a quantifier to "reach outside" the scope of a modal it's embedded in, and then "jump back in". To escape the scope of a modal, one only needs to use a modal operator to shift the world of evaluation; but to jump back inside the scope of a modal, one needs to make use of @. So the presence of a quantifier scoping over @ evinces cross-world quantification.

Given that **Definition 8.1** embodies the correct precisification of cross-world predication and quantification, we can now state the question above more precisely. Let $\mathcal{L}^{\mathsf{1M-}}$

[43] Of course, this needs to be qualified. Here's a potential counterexample: "The authors of *Principia* could have written more clearly than they actually did". This sentence cannot be expressed in $\mathcal{L}^{\mathsf{H}}$, yet clearly is cross-world. But the reason this sentence is not expressible is not because of the cross-world part of the sentence, but because it involves plurals; and this is a problem even for the two-sorted language, our most expressive language thus far. So really, what should be said is this: to refute the thesis, one would need to find a sentence that can be formalized into the two-sorted language that isn't expressible in $\mathcal{L}^{\mathsf{H}}$ but still appears to be a cross-world sentence.

be the @-free fragment of $\mathcal{L}^{1M}$ (that is, standard first-order modal logic *without* an actuality operator). According to **Definition 8.1**, $\mathcal{L}^{1M-}$ only contains non-cross-world formulas. Furthermore, according to **Definition 8.1**, $\mathcal{L}^{H}$ contains *all* of the cross-world formulas. But is there a language that extends $\mathcal{L}^{1M-}$ but doesn't extend $\mathcal{L}^{H}$ that captures all the cross-world formulas in $\mathcal{L}^{H}$?

Theorem 8.2 *(The Non-Cross-World Fragment of $\mathcal{L}^{H}$ is Essentially $\mathcal{L}^{1M-}$)*. Every state-closed non-cross-world $\mathcal{L}^{H}$-formula is equivalent to an $\mathcal{L}^{1M-}$-formula. More generally, every non-cross-world formula of the form $\varphi(x_1, \ldots, x_n; t_1, \ldots, t_m)$ is equivalent to a boolean combination of $\mathcal{L}^{H}$-formulas that are either of the form $@_{t_i}\psi$ or of the form θ, where ψ and θ are $\mathcal{L}^{1M-}$-formulas.

In other words, the only non-cross-world state-closed formulas in $\mathcal{L}^{H}$ were already in $\mathcal{L}^{1M-}$: the only state-closed formulas $\mathcal{L}^{H}$ adds are cross-world. The proof is in §C. Combined with **Definition 8.1**, which says $\mathcal{L}^{H}$ adds all state-closed cross-world formulas, it follows that $\mathcal{L}^{H}$ is exactly the minimal extension of $\mathcal{L}^{1M-}$ that captures all cross-world formulas. $\mathcal{L}^{H}$ does add some new state-open non-cross-world formulas (that's why the theorem doesn't just say that the non-cross-world fragment of $\mathcal{L}^{H}$ is $\mathcal{L}^{1M-}$ full stop). But as far as minimality is concerned, this shouldn't be troubling for two reasons.

First, over *diagonal* indices—which, in the hybrid setting, means indices of the form $\langle \mathcal{M}, w, g\rangle$ where $g(s) = w$ for all $s \in \mathsf{SVAR}$—the restriction to state-closed $\mathcal{L}^{H}$-formulas can be dropped. That is, every non-cross-world formulas is equivalent over diagonal indices to an $\mathcal{L}^{1M-}$-formula. This is just because if φ is an $\mathcal{L}^{1M-}$-formula, then $@_{t_i}\varphi$ is equivalent over diagonal indices to φ.

Second, even over non-diagonal indices, $\mathcal{L}^{H}$ doesn't introduce a wholly new *kind* of non-cross-world formulas as previous approaches did. After all, just consider the case where our non-cross-world formula φ only contains our special state variable s_0 that by convention picks out the world considered as actual. Then φ will be equivalent to a boolean combination of non-cross-world formulas in $\mathcal{L}^{1M}$ (with @). In general, the new kind of non-cross-world formulas that $\mathcal{L}^{H}$ introduces simply assert that old kinds of non-cross-world formulas hold elsewhere. By contrast, the new non-cross-world formulas in the degree approach or the two-sorted language will be of a wholly different sort not found in anything like $\mathcal{L}^{1M}$.

So modulo state-closure, one will not find a language that extends $\mathcal{L}^{1M-}$ and can express all cross-world formulas, but doesn't extend $\mathcal{L}^{H}$. Hence, it's reasonable to conclude that $\mathcal{L}^{H}$ does get at the *heart* of cross-world phenomena.

§A Two-Sorted Logic

In this section, we review the standard semantics for two-sorted first-order logic. We then state precisely in what sense the translation from **Definition 4.1** from $\mathcal{L}^{1M}$ to $\mathcal{L}^{2S}$ is in fact accurate.

Throughout, we adopt the following convention:

Notation: If $\alpha_1, \ldots, \alpha_n$ is any sequence (of variables, terms, objects, etc.), we may write "$\overline{\alpha}$" in place of "$\alpha_1, \ldots, \alpha_n$". $\overline{\alpha}$ is assumed to be of the appropriate length in any given context. When f is some unary function, we may write "$f(\overline{\alpha})$" in place of "$f(\alpha_1), \ldots, f(\alpha_n)$". (Context will always distinguish between $f(\alpha_1), \ldots, f(\alpha_n)$ and $f(\alpha_1, \ldots, \alpha_n)$.) Where $\overline{\alpha}$ is a sequence, we'll let $|\overline{\alpha}|$ be the length of $\overline{\alpha}$.

§A.1 Models

First, we review standard two-sorted first-order models.

Definition A.1 *(Two-Sorted Models)*. A ***$\mathcal{L}^{2S}$-model*** or ***two-sorted model*** is an ordered tuple $\mathcal{M}^{2S} = \langle W, D, V \rangle$ where W and D are sets, and V is a function (the ***valuation function***) such that:

- for each $c \in \mathsf{CON}$, $V(c)\colon W \to D$;
- for each $P^{n/m} \in \mathsf{PRED}^{n/m}$, $V(P^{n/m}) \subseteq D^n \times W^m$;
- $V(\mathsf{E}) \subseteq D \times W$;
- $V(\mathsf{R}) \subseteq W \times W$.

Usually, we are interested in $\mathcal{L}^{2S}$-models which *correspond* to some $\mathcal{L}^{1M}$-model.

Definition A.2 *(Model Correspondents)*. Let $\mathcal{M} = \langle W, R, D, \delta, I \rangle$ be an $\mathcal{L}^{1M}$-model. An ***$\mathcal{L}^{2S}$-correspondent of $\mathcal{M}$*** is a $\mathcal{L}^{2S}$-model $\mathcal{M}^{2S} = \langle W, D, V \rangle$ such that:

- for all $c \in \mathsf{CON}$, $V(c)(w) = I(c, w)$;
- for all $P^n \in \mathsf{PRED}^{n/1}$, $V(P^n) = \{\langle a_1, \ldots, a_n; w \rangle \mid \langle a_1, \ldots, a_n \rangle \in I(P^n, w)\}$;
- $V(\mathsf{E}) = \{\langle a; w \rangle \in D \times W \mid a \in \delta(w)\}$;
- $V(\mathsf{R}) = R$.

An ***$\mathcal{L}^{2S}$-correspondent*** is just an $\mathcal{L}^{2S}$-correspondent of some $\mathcal{L}^{1M}$-model.

Notice in particular that **Definition A.2** doesn't pick out a unique $\mathcal{L}^{2S}$-correspondent for any given $\mathcal{L}^{1M}$-model. This will be important below.

§A.2 Semantics

Next, we review the standard semantics for two-sorted first-order logic.

Definition A.3 *(Two-Sorted Variable Assignment)*. Let $\mathcal{M}^{2S}$ be an $\mathcal{L}^{2S}$-model. A ***variable assignment for $\mathcal{M}^{2S}$*** is a function assigning members of D to object variables, and members of W to state variables. The other definitions are as they are in **Definition 3.2**. For any $\mathcal{L}^{2S}$-correspondent $\mathcal{M}^{2S}$ of $\mathcal{M}$, and any g on $\mathcal{M}$, we'll say that

g^{2S} for $\mathcal{M}^{2S}$ is a ***$\mathcal{L}^{2S}$-correspondent for*** g if $g(x) = g^{2S}(x)$ for all $x \in \mathsf{VAR}$.

Definition A.4 *(Two-Sorted Denotation)*. Let $\tau \in \mathsf{TERM}^{2S}$, let $\mathcal{M}^{2S}$ be an $\mathcal{L}^{2S}$-model, and let $g^{2S} \in \mathsf{VA}(\mathcal{M}^{2S})$. The ***denotation of τ at $\langle \mathcal{M}^{2S}, g^{2S} \rangle$***, $[\![\tau]\!]^{\mathcal{M}^{2S}, g^{2S}}$, is defined as follows:

$$[\![\tau]\!]^{\mathcal{M}^{2S}, g^{2S}} = \begin{cases} V(c)(g^{2S}(s)) & \text{if } \tau = c(s) \text{ where } c \in \mathsf{CON} \text{ and } s \in \mathsf{SVAR} \\ g^{2S}(x) & \text{if } \tau = x \text{ where } x \in \mathsf{VAR}. \end{cases}$$

Definition A.5 *(Two-Sorted Satisfaction)*. The ***two-sorted satisfaction relation***, $\vDash$, is defined recursively, for all $\mathcal{L}^{2S}$-models $\mathcal{M}^{2S} = \langle W, D, V \rangle$ and all variable assignments $g^{2S} \in \mathsf{VA}(\mathcal{M}^{2S})$:

$$\begin{array}{lcl}
\mathcal{M}^{2S}, g^{2S} \vDash P^{n/m}(\overline{\tau}; \overline{s}) & \Leftrightarrow & \left\langle [\![\,\overline{\tau}\,]\!]^{\mathcal{M}^{2S}, g^{2S}}; g^{2S}(\overline{s}) \right\rangle \in V(P^{n/m}) \\
\mathcal{M}^{2S}, g^{2S} \vDash \tau_1 \approx \tau_2 & \Leftrightarrow & [\![\tau_1]\!]^{\mathcal{M}^{2S}, g^{2S}} = [\![\tau_2]\!]^{\mathcal{M}^{2S}, g^{2S}} \\
\mathcal{M}^{2S}, g^{2S} \vDash s_1 \approx s_2 & \Leftrightarrow & g^{2S}(s_1) = g^{2S}(s_2) \\
\mathcal{M}^{2S}, g^{2S} \vDash \mathsf{E}(\tau; s) & \Leftrightarrow & \langle [\![\tau]\!]^{\mathcal{M}^{2S}, g^{2S}}; g^{2S}(s) \rangle \in V(\mathsf{E}) \\
\mathcal{M}^{2S}, g^{2S} \vDash \mathsf{R}(s_1, s_2) & \Leftrightarrow & \langle g^{2S}(s_1), g^{2S}(s_2) \rangle \in V(\mathsf{R}) \\
\mathcal{M}^{2S}, g^{2S} \vDash \neg \varphi & \Leftrightarrow & \mathcal{M}^{2S}, g^{2S} \nvDash \varphi \\
\mathcal{M}^{2S}, g^{2S} \vDash \varphi \wedge \psi & \Leftrightarrow & \mathcal{M}^{2S}, g^{2S} \vDash \varphi \text{ and } \mathcal{M}, g \vDash \psi \\
\mathcal{M}^{2S}, g^{2S} \vDash \forall x\, \varphi & \Leftrightarrow & \forall a \in D\colon\ \mathcal{M}^{2S}, (g^{2S})^x_a \vDash \varphi \\
\mathcal{M}^{2S}, g^{2S} \vDash \forall s\, \varphi & \Leftrightarrow & \forall w \in W\colon\ \mathcal{M}^{2S}, (g^{2S})^s_w \vDash \varphi.
\end{array}$$

§A.3 Expressivity

Recall the translations from $\mathcal{L}^{1M}$ into $\mathcal{L}^{2S}$ from **Definition 4.1**. Given the definitions above, the following is easy to prove by induction on the complexity of $\mathcal{L}^{1M}$-terms and $\mathcal{L}^{1M}$-formulas.

Lemma A.6 *(Adequacy of Translation)*. Let $\mathcal{M}$ be an $\mathcal{L}^{1M}$-model, $\mathcal{M}^{2S}$ an $\mathcal{L}^{2S}$-correspondent for $\mathcal{M}$, $w, v \in W$, $g \in \mathsf{VA}(\mathcal{M})$, g^{2S} an $\mathcal{L}^{2S}$-correspondent variable assignment of g for $\mathcal{M}^{2S}$, $s, t \in \mathsf{SVAR}$, τ an $\mathcal{L}^{1M}$-term, and φ an $\mathcal{L}^{1M}_{\Pi}$-formula.

(a) $[\![\tau]\!]^{\mathcal{M}, w, v, g} = [\![\mathsf{st}_{s,t}(\tau)]\!]^{\mathcal{M}^{2S}, (g^{2S})^{s,t}_{w,v}}$

(b) $\mathcal{M}, w, v, g \Vdash \varphi$ iff $\mathcal{M}^{2S}, (g^{2S})^{s,t}_{w,v} \vDash \mathsf{ST}_{s,t}(\varphi)$.

One can now prove more rigorously that cross-world predication is in general not expressible in $\mathcal{L}^{\mathsf{1M}}$. First, let's say what it means for a $\mathcal{L}^{\mathsf{1M}}$-formula to be expressible in $\mathcal{L}^{\mathsf{2S}}$:[44]

Definition A.7 *(Expressivity)*. An $\mathcal{L}^{\mathsf{1M}}$-formula $\varphi(\overline{x})$ ***expresses*** an $\mathcal{L}^{\mathsf{2S}}$-formula $\varphi^{\mathsf{2S}}(\overline{x}; s, t)$ if φ^{2S} is equivalent (in $\mathcal{L}^{\mathsf{2S}}$) to $\mathsf{ST}_{s,t}(\varphi)$. Similarly for $\mathcal{L}^{\mathsf{1M}}_{\Pi}$-formulas.

Now, recall that in **Definition A.2**, no constraints are placed on how $\mathcal{L}^{\mathsf{2S}}$-correspondents are to interpret n/m-predicates where $m > 1$. Thus, the extensions of the relevant 2/2-place predicates occurring in (16)–(18) could be anything—nothing about the $\mathcal{L}^{\mathsf{1M}}$-model will tell us what they must be in its $\mathcal{L}^{\mathsf{2S}}$-correspondents. Such arbitrariness makes it fairly easy to create two $\mathcal{L}^{\mathsf{2S}}$-correspondents of an $\mathcal{L}^{\mathsf{1M}}$-model which disagree on one of (16)–(18). But then it follows by **Lemma A.6** that none of (16)–(18) can be equivalent to the (possibilist or actualist) translation of a $\mathcal{L}^{\mathsf{1M}}$-formula. Thus:

Corollary A.8 *(Inexpressibility of Cross-World Predication)*. There is no $\mathcal{L}^{\mathsf{1M}}_{\Pi}$-formula that expresses any of (16)–(18).

The proof that cross-world quantification—in particular, (19)–(20)—are inexpressible requires more work. The more complicated proof is in Kocurek [2015].

§B Wehmeier's Subjunctive Logic

In this section, we'll briefly examine the framework of Wehmeier [2012], which was also designed to solve the problems of cross-world predication and quantification, and explain how it relates to $\mathcal{L}^{\mathsf{H}}$.[45]

The key idea behind Wehmeier's proposal is idea of *mood*. Essentially, Wehmeier suggests that what needs to be added to $\mathcal{L}^{\mathsf{1M}}$ to overcome these expressivity issues is not some new operators, but rather some way of distinguishing indiciative and subjunctive moods in the syntax. Wehmeier does this by introducing "mood markers" $i, s, s_1, s_2, s_3, \ldots$ (i for "indicative", s for "subjunctive"), which are basically state variables. Predicates will then be decorated with mood markers to indicate the worlds relevant for their evaluation, i.e., the world relative to which we calculate the extension of that predicate.

Formally, Wehmeier's language $\mathcal{L}^{\mathsf{Weh}}$ is built as follows:

$$\tau ::= c \mid x$$
$$\varphi ::= P^{n/t}(\tau_1, \ldots, \tau_n) \mid C^{n/t_1,\ldots,t_n}(\tau_1, \ldots, \tau_n) \mid \tau \approx \sigma \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid \Box^t_k \varphi \mid \forall^t x\, \varphi$$

where $k \geqslant 1$, and $t, t_1, \ldots, t_n$ are mood markers. In $\mathcal{L}^{\mathsf{Weh}}$, we distinguish two kinds of predicates: "ordinary" predicates P^n that are only decorated with one mood marker, and "cross-world" predicates C^n that are decorated with n-mood markers. While $\mathcal{L}^{\mathsf{H}}$ does not make this distinction, it could in principle by introducing a class of "ordinary" predicates

[44]This is not the most general definition of expressivity, but for our purposes, we only need this particular instance of the more general definition.

[45]I again allow myself the flexibility of re-writing Wehmeier's notation, for the sake of continuity.

are essentially insensitive to $\blacktriangleleft$-terms. We will keep this distinction to help ease the comparison betweeh $\mathcal{L}^{\mathsf{H}}$ and $\mathcal{L}^{\mathsf{Weh}}$.

The quantifiers are decorated with mood markers to determine the domain of quantification: $\forall^{s_n} x\, \varphi$ says that every x that exists *in* w_n satisfies φ. Modal operators are both decorated by mood markers and subscriped with numbered indicies. The mood marker indicates world relative to which accessibility is determined, while the subscript indicates where we save the shifted world for reference: $\Box_k^{s_n} \varphi$ says that at every world v accessible to w_n, φ is true assuming we save v as the k^{th} reference world.

With regards to models and semantics, Wehmeier essentially uses cw-models with a number of constraints. For one thing, he assumes that constants are rigid designators (so $I(c, w) = I(c, v)$ for all $w, v \in W$) and that the extension of cross-world predicates are rigid in a similar sense (so $I(C, w) = I(C, v)$ for all $w, v \in W$). We will follow Wehmeier and assume these constraints for ease of comparison to $\mathcal{L}^{\mathsf{H}}$, noting that they are not essential to any of the results that follow and may be dropped if so desired. In what follows, we'll simply write "$I(c)$" and "$I(C)$" for brevity. He also assumes that R is universal (in which case, he can drop the mood marker decorating $\Box$) and that only binary predicates (in particular *comparatives*) are cross-world. However, we will not impose these additional constraints, again for ease of comparison to $\mathcal{L}^{\mathsf{H}}$.[46]

The last major difference in Wehmeier's models is that he defines the extension of ordinary predicates to be an ordered n-tuple of *objects*, not *object-world pairs*. In terms of cw-models, this means that the extension of predicates is insensitive to the world coordinates of object-world pairs. Thus, we can for our purposes define the class of models for Wehmeier as follows:

Definition B.1 *(weh-Models)*. A cw-model $\mathcal{M}$ is a ***weh-model*** if it meets the following two constraints:

(i) for all $c \in \mathsf{CON}$ and all $w, v \in W$, $I(c, w) = I(c, v)$;

(ii) for all ordinary predicates P, all $w, v_1, \ldots, v_n \in W$, and all $a_1, \ldots, a_n \in D$, $\langle\langle a_1, v_1\rangle, \ldots, \langle a_n, v_n\rangle\rangle \in I(P, w)$ iff $\langle\langle a_1, w\rangle, \ldots, \langle a_n, w\rangle\rangle \in I(P, w)$;

(iii) for all cross-world predicates C and all $w, v \in W$, $I(C, w) = I(C, v)$.

The semantics relativizes satisfaction to indices of the form $\langle \mathcal{M}, \overline{w}, v, g\rangle$, where $\mathcal{M}$ is a weh-model, and where $\overline{w} = w_0, w_1, w_2, \ldots$. The denotation of terms only needs to be relativized to a model and variable assignment, since the constants rigidly designate:

$$[\![\tau]\!]^{\mathcal{M},g} = \begin{cases} I(c) & \text{if } \tau = c \in \mathsf{CON} \\ g(x) & \text{if } \tau = x \in \mathsf{VAR}. \end{cases}$$

Finally, the interesting semantic clauses are given as follows:

$$\mathcal{M}, \overline{w}, v, g \Vdash_{\mathsf{Weh}} P^{t_1}(\tau_1, \ldots, \tau_n) \quad \Leftrightarrow \quad \left\langle\left\langle [\![\tau_1]\!]^{\mathcal{M},g}, u_1\right\rangle, \ldots, \left\langle [\![\tau_n]\!]^{\mathcal{M},g}, u_1\right\rangle\right\rangle \in I(P, u_1)$$

[46] Wehmeier has indicated in personal communication that something like the proposal given here is the proposal he would adopt were he to drop the various semantic constraints made in Wehmeier [2012].

$$\begin{array}{lcl}
\mathcal{M}, \overline{w}, v, g \Vdash_{\mathsf{Weh}} C^{t_1,\dots,t_n}(\tau_1, \dots, \tau_n) & \Leftrightarrow & \left\langle \left\langle [\![\tau_1]\!]^{\mathcal{M},g}, u_1 \right\rangle, \dots, \left\langle [\![\tau_n]\!]^{\mathcal{M},g}, u_n \right\rangle \right\rangle \in I(C) \\
\mathcal{M}, \overline{w}, v, g \Vdash_{\mathsf{Weh}} \tau \approx \sigma & \Leftrightarrow & [\![\tau]\!]^{\mathcal{M},g} = [\![\sigma]\!]^{\mathcal{M},g} \\
\mathcal{M}, \overline{w}, v, g \Vdash_{\mathsf{Weh}} \Box_k^{t_1} \varphi & \Leftrightarrow & \forall u \in R[u_1]\colon\ \mathcal{M}, \overline{w}[k \mapsto u], u, g \Vdash_{\mathsf{Weh}} \varphi \\
\mathcal{M}, \overline{w}, v, g \Vdash_{\mathsf{Weh}} \forall^{t_1} x\, \varphi & \Leftrightarrow & \forall a \in \delta(u_1)\colon\ \mathcal{M}, \overline{w}, v, g_a^x \Vdash_{\mathsf{Weh}} \varphi
\end{array}$$

where:

$$u_m = \begin{cases} w_0 & \text{if } t_m = i \\ v & \text{if } t_m = s \\ w_j & \text{if } t_m = s_j. \end{cases}$$

Thus, for instance, (Rich*) is formalized as:

$$\Box_1^i \Diamond_2^s \forall^{s_1} x\ (\mathsf{Rich}^{s_1}(x) \to \mathsf{Poor}^s(x)). \tag{57}$$

With Wehmeier's framework laid out, we now consider the question of whether $\mathcal{L}^{\mathsf{Weh}}$ is intertranslateable with $\mathcal{L}^{\mathsf{H}}$, at least over the class of weh-models. First, notice that it's easy to give a translation $\mathsf{W2H}(\varphi)$ from $\mathcal{L}^{\mathsf{Weh}}$ to $\mathcal{L}^{\mathsf{H}}$ (assume for simplicity that we replaced the mood marker i with s_0):

$$\begin{array}{rcl}
\mathsf{W2H}(P^s(\tau_1, \dots, \tau_n)) & = & P(\tau_1, \dots, \tau_n) \\
\mathsf{W2H}(P^t(\tau_1, \dots, \tau_n)) & = & @_t P(\tau_1, \dots, \tau_n) \\
\mathsf{W2H}(C^{t_1,\dots,t_n}(\tau_1, \dots, \tau_n)) & = & C(\blacktriangleleft_{t_1} \tau_1, \dots, \blacktriangleleft_{t_n} \tau_n) \\
\mathsf{W2H}(\tau \approx \sigma) & = & \tau \approx \sigma \\
\mathsf{W2H}(\neg \varphi) & = & \neg \mathsf{W2H}(\varphi) \\
\mathsf{W2H}(\varphi \wedge \psi) & = & \mathsf{W2H}(\varphi) \wedge \mathsf{W2H}(\psi) \\
\mathsf{W2H}(\Box_k^s \varphi) & = & \Box_{s_k} \mathsf{W2H}(\varphi) \\
\mathsf{W2H}(\Box_k^t \varphi) & = & @_t \Box_{s_k} \mathsf{W2H}(\varphi) \\
\mathsf{W2H}(\forall^s x\, \varphi) & = & \forall x\, \mathsf{W2H}(\varphi) \\
\mathsf{W2H}(\forall^t x\, \varphi) & = & \forall_t x\, \mathsf{W2H}(\varphi)
\end{array}$$

where $t \neq s$ is a mood marker, and $t_1, \dots, t_n$ are mood markers (possibly including s). For example, the translation of (57) is:

$$@_{s_0} \Box_{s_1} \Diamond_{s_2} \forall_{s_1} x\ (@_{s_1} \mathsf{Rich}(x) \to \mathsf{Poor}(x)) \tag{58}$$

which is essentially how we formalized (Rich*) in $\mathcal{L}^{\mathsf{H}}$, except the s_2 in $\mathcal{L}^{\mathsf{H}}$ isn't necessary.

Theorem B.2 *(Adequacy of W2H).* Let $\mathcal{M}$ be a weh-model, $w \in W$, and g an H-variable assignment over $\mathcal{M}$. Then for any $\mathcal{L}^{\mathsf{Weh}}$-formula φ:

$$\mathcal{M}, g(s_0), g(s_1), g(s_2), \dots, w, g \Vdash_{\mathsf{Weh}} \varphi \quad \Leftrightarrow \quad \mathcal{M}, w, g \Vdash_{\mathsf{H}} \varphi.$$

The reverse translation is not as straightforward. The issue we need to deal with is $\downarrow$, which allows one to reset a reference world at will. To deal with this, we need to first manipulate all $\mathcal{L}^{\mathsf{H}}$-formulas into a nice and manageable form.

Definition B.3 *(Nice $\mathcal{L}^{H}$-formulas).* A $\mathcal{L}^{\mathsf{H}}$-formula φ is ***nice*** if φ is of the form $\downarrow s_n.\psi$ where:

(i) every term has at most one occurrence of $\blacktriangleleft$

(ii) s_0 is never bound in φ (and thus, $s_0 \neq s_n$)

(iii) there is no occurrence of $\downarrow s_n.$ in ψ

(iv) for all $s \in \mathsf{SVAR}$ where $s \neq s_n$, there is at most one occurrence of $\downarrow s.$ in ψ

(v) for all $s \in \mathsf{SVAR}$, if s has a free occurrence in ψ, then it does not also have a bound occurrence in ψ

(vi) for all $s \in \mathsf{SVAR}$, if $\downarrow s.$ occurs in ψ, then its single occurrence is prefixed by a $\Box$

(vii) every occurrence of $\Box$ prefixes an occurrence of $\downarrow s.$ for some $s \in \mathsf{SVAR}$

In other words, nice $\mathcal{L}^{\mathsf{H}}$-formulas are those such that (a) the state variables are nicely organized, (b) irrelevant stackings of $\blacktriangleleft$ are removed, (c) every $\Box$ is followed by exactly one unique $\downarrow t.$, and (d) apart from the beginning of the formula, that's the only place where $\downarrow t.$ show up.

The following is easy to prove, but requires some tedious details and is simply a matter of reorganizing and rewriting variables appropriately.

Lemma B.4 *(Nice Normal Form).* Every $\mathcal{L}^{\mathsf{H}}$-formula is equivalent to a nice $\mathcal{L}^{\mathsf{H}}$-formula. Furthermore, there's a recursive procedure for transforming each $\mathcal{L}^{\mathsf{H}}$-formula into one that's nice.

So to show that $\mathcal{L}^{\mathsf{H}}$ can be translated into $\mathcal{L}^{\mathsf{Weh}}$, it suffices to show that the nice $\mathcal{L}^{\mathsf{H}}$-formulas can be translated into $\mathcal{L}^{\mathsf{Weh}}$. The first step is to extract the "object" and "mood" parts of a given term as follows:

$$\mathsf{ob}(\tau) = \begin{cases} c & \text{if } \tau = c \in \mathsf{CON} \\ x & \text{if } \tau = x \in \mathsf{VAR} \\ \mathsf{ob}(\sigma) & \text{if } \tau = \blacktriangleleft_t \sigma \end{cases}$$

$$\mathsf{mo}(\tau) = \begin{cases} s & \text{if } \tau \in \mathsf{CON} \cup \mathsf{VAR} \\ s_k & \text{if } \tau = \blacktriangleleft_{s_k} \sigma. \end{cases}$$

Now, since every nice $\mathcal{L}^{\mathsf{H}}$-formula is of the form $\downarrow s_n.\varphi$, we define a translation function $\mathsf{H2W}(\varphi)$ from $\mathcal{L}^{\mathsf{H}}$ to $\mathcal{L}^{\mathsf{Weh}}$ by induction on φ.

$$\begin{array}{rcl}
\mathsf{H2W}(P(\tau_1,\ldots,\tau_n)) & = & P^s(\mathsf{ob}(\tau_1),\ldots,\mathsf{ob}(\tau_n)) \\
\mathsf{H2W}(C(\tau_1,\ldots,\tau_n)) & = & C^{\mathsf{mo}(\tau_1),\ldots,\mathsf{mo}(\tau_n)}(\mathsf{ob}(\tau_1),\ldots,\mathsf{ob}(\tau_n)) \\
\mathsf{H2W}(\tau \approx \sigma) & = & \mathsf{ob}(\tau) \approx \mathsf{ob}(\sigma) \\
\mathsf{H2W}(\mathsf{E}(\tau)) & = & \exists^s y\,(y \approx \mathsf{ob}(\tau)) \\
\mathsf{H2W}(\neg\varphi) & = & \neg\,\mathsf{H2W}(\varphi) \\
\mathsf{H2W}(\varphi \wedge \psi) & = & \mathsf{H2W}(\varphi) \wedge \mathsf{H2W}(\psi) \\
\mathsf{H2W}(\Box_{s_k}\varphi) & = & \Box_k^s \mathsf{H2W}(\varphi) \\
\mathsf{H2W}(@_{s_k}\varphi) & = & \mathsf{H2W}(\varphi)\,[s/s_k] \\
\mathsf{H2W}(\forall x\,\varphi) & = & \forall_s x\,\mathsf{H2W}(\varphi)
\end{array}$$

where y does not occur in τ and $\mathsf{H2W}(\varphi)\,[s/t]$ is the result of replacing every instance of s *that's not* within the scope of a modal with t.[47]

Theorem B.5 *(Adequacy of H2W)*. Let $\mathcal{M}$ be a weh-model, $w \in W$, and g an H-variable assignment over $\mathcal{M}$. Then for any nice $\mathcal{L}^{\mathsf{H}}$-formula $\downarrow s_n.\varphi$, we have that $\mathcal{M}, w, g \Vdash_{\mathsf{H}} \downarrow s_n.\varphi$ iff $\mathcal{M}, g_w^{s_n}(s_0), g_w^{s_n}(s_1), g_w^{s_n}(s_2), \ldots, w, g_w^{s_n} \Vdash_{\mathsf{Weh}} \mathsf{H2W}(\varphi)$.

Thus, in a very strong sense, we can think of $\mathcal{L}^{\mathsf{H}}$ as a generalization of Wehmeier's original framework that lifts various restrictions he placed on the models and the syntax.

§C Characterization of Cross-World Formulas

The goal of this section is to prove that **Theorem 8.2**, viz., that every non-cross-world formula of the form $\varphi(x_1,\ldots,x_n;t_1,\ldots,t_m)$ is equivalent to a boolean combination of $\mathcal{L}^{\mathsf{H}}$-formulas that are either of the form $@_{t_i}\psi$ or of the form θ, where ψ and θ are $\mathcal{L}^{1\mathsf{M}-}$-formulas.

Definition C.1 *(Cross-world)*. An $\mathcal{L}^{\mathsf{H}}$-formula is ***explicitly non-cross-world*** if it neither contains an instance of $\blacktriangleleft$ nor an object quantifier scoping over an instance of @. Thus, non-cross-world formulas are those that are equivalent to some explicitly non-cross-world formula.

[47] In the case of $\Box_k^s$, we do not count the s here as being within its own scope, so such an instance of s would also be replaced by t if its not in the scope of other modals. So for example, $\mathsf{H2W}(@_{s_k}\Box_{s_n}\varphi) = \mathsf{H2W}(\Box_{s_n}\varphi)\,[s/s_k] = (\Box_{s_n}^s \mathsf{H2W}(\varphi))[s/s_k] = \Box_n^{s_k}\mathsf{H2W}(\varphi)\,[s/s_k]$.

Definition C.2 *(Isolation)*. An ***isolated atom*** is any $\mathcal{L}^{\mathsf{H}}$-formula either of the form $@_t\psi$ or of the form θ where ψ and θ are $\mathcal{L}^{\mathsf{1M-}}$-formulas. An $\mathcal{L}^{\mathsf{H}}$-formula φ is in ***isolated form*** if it is a boolean combination of isolated atoms.

Clearly, every $\mathcal{L}^{\mathsf{H}}$-formula in isolated form is (explicitly) non-cross-world.

Theorem C.3 *(Non-Cross-World is Isolation)*. Every non-cross-world $\mathcal{L}^{\mathsf{H}}$-formula is equivalent to a $\mathcal{L}^{\mathsf{H}}$-formula in isolated form.

Theorem C.3 is just a more concise statement of **Theorem 8.2**.

Proof: It suffices to show the claim for explicitly non-cross-world $\mathcal{L}^{\mathsf{H}}$-formulas φ. We proceed by induction. Clearly this holds for atomics and boolean combinations of non-cross-world formulas. Furthermore, if $\forall x\,\psi$ is an explicitly non-cross-world formula, then ψ must not contain @, and hence $\forall x\,\psi$ is already an $\mathcal{L}^{\mathsf{1M-}}$-formula (and so *a fortiori* in isolated form). So the interesting cases are the modals.

Necessity. $\varphi = \Box\psi$. Since φ is a non-cross-world formula, ψ must be too. By inductive hypothesis, suppose ψ is in isolated form. Using standard rewrite rules (and since $@_s$ and $\neg$ commute), WLOG, we can suppose ψ is of the form:

$$\psi = \bigwedge_{i=1}^{k} \left(@_{t^i_1}\alpha^i_1 \vee \cdots \vee @_{t^i_{n_i}}\alpha^i_{n_i} \vee \beta^i\right)$$

where each α and β is an $\mathcal{L}^{\mathsf{1M-}}$-formula. Since $\Box$ distributes over conjunction, it suffices to check that formulas of the form:

$$\Box\,(@_{t_1}\alpha_1 \vee \cdots \vee @_{t_n}\alpha_n \vee \beta)$$

where $\alpha_1, \ldots, \alpha_n, \beta$ are $\mathcal{L}^{\mathsf{1M-}}$-formulas, can be written as a boolean combination of isolated atoms. But it's easy to check that this is equivalent to:

$$@_{t_1}\alpha_1 \vee \cdots \vee @_{t_n}\alpha_n \vee \Box\beta,$$

which is a disjunction of isolated atoms. ✓

Actuality. $\varphi = @_s\psi$. Again, WLOG, write ψ as:

$$\psi = \bigwedge_{i=1}^{k} \left(@_{t^i_1}\alpha^i_1 \vee \cdots \vee @_{t^i_{n_i}}\alpha^i_{n_i} \vee \beta^i\right)$$

It's easy to check that $@_s\psi$ is equivalent to:

$$\bigwedge_{i=1}^{k} @_s \left(@_{t^i_1}\alpha^i_1 \vee \cdots \vee @_{t^i_{n_i}}\alpha^i_{n_i} \vee \beta^i\right)$$

which is equivalent to:

$$\bigwedge_{i=1}^{k} \left(@_{t_1^i} \alpha_1^i \vee \cdots \vee @_{t_{n_i}^i} \alpha_{n_i}^i \vee @_s \beta^i \right)$$

which is in isolated form. ✓

Saving. $\varphi = \downarrow s.\psi$. Again, WLOG, write ψ as:

$$\psi = \bigwedge_{i=1}^{k} \left(@_{t_1^i} \alpha_1^i \vee \cdots \vee @_{t_{n_i}^i} \alpha_{n_i}^i \vee @_s \gamma_1^i \vee \cdots \vee @_s \gamma_{k_i}^i \vee \beta^i \right)$$

where also each γ is an $\mathcal{L}^{1\mathsf{M}-}$-formula, and where none of $t_1^i, \ldots, t_n^i$ are s. Then $\downarrow s.\psi$ is equivalent to:

$$\bigwedge_{i=1}^{k} \left(@_{t_1^i} \alpha_1^i \vee \cdots \vee @_{t_{n_i}^i} \alpha_{n_i}^i \vee \gamma_1^i \vee \cdots \vee \gamma_{k_i}^i \vee \beta^i \right)$$

which is in isolated form. ✓ ■

It follows that, up to equivalence, the non-cross-world $\mathcal{L}^{\mathsf{H}}$-formulas are exactly the $\mathcal{L}^{\mathsf{H}}$-formulas in isolated form. Notice that in the proof above, once we've rewritten φ into isolated form, no bound state variables are left: only the free state variables of φ remain after being transformed into isolated form. In particular, if φ doesn't contain any free state variables, then the result of this procedure will be to rewrite φ as an $\mathcal{L}^{1\mathsf{M}-}$-formula. As a result, **Theorem C.3** shows that all state-closed non-cross-world $\mathcal{L}^{\mathsf{H}}$-formulas are equivalent to $\mathcal{L}^{1\mathsf{M}-}$-formulas. Hence, up to equivalence, the state-closed non-cross-world $\mathcal{L}^{\mathsf{H}}$-formulas are exactly the $\mathcal{L}^{1\mathsf{M}-}$-formulas.

References

Aloni, Maria. 2005. "Individual Concepts in Modal Predicate Logic." *Journal of Philosophical Logic* 34:1–64.

Areces, Carlos, Blackburn, Patrick, and Marx, Maarten. 2003. "Repairing the Interpolation Theorem in Quantified Modal Logic." *Annals of Pure and Applied Logic* 124:287–299.

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

Blackburn, Patrick and Marx, Maarten. 2002. "Tableaux for Quantified Hybrid Logic." In *Automated Reasoning with Analytic Tableaux and Related Methods*, 38–52. Springer.

Boolos, George. 1984. "To Be is to Be a Value of a Variable (or to Be Some Values of Some Variables)." *The Journal of Philosophy* 81:430–449.

Bricker, Phillip. 1989. "Quantified Modal Logic and the Plural De Re." *Midwest Studies in Philosophy* 14:372–394.

Butterfield, Jeremy and Stirling, Colin. 1987. "Predicate Modifiers in Tense Logic." *Logique et Analyse* .

Button, Tim. 2012. "Spotty Scope and Our Relation to Fictions." *Noûs* 46:243–258.

Cantwell, John. 1995. "Comparatives." *Theoretical Linguistics* 21:145–158.

Cresswell, Maxwell J. 1990. *Entities and Indices*. Kluwer Academic Publishers.

Cumming, Samuel. 2008. "Variabilism." *Philosophical Review* 117:525–554.

Davies, Martin and Humberstone, Lloyd. 1980. "Two Notions of Necessity." *Philosophical Studies* 38:1–30.

Fitting, Melvin. 2013. "On Height and Happiness." Manuscript.

Fitting, Melvin and Mendelsohn, Richard L. 1998. *First-Order Modal Logic*. Kluwer Academic Publishers.

Forbes, Graeme. 1989. *Languages of Possibility*. Blackwell Publishers.

—. 1994. "Comparatives in Counterpart Theory: Another Approach." *Analysis* 54:37.

Hazen, Allen P. 1976. "Expressive Completeness in Modal Language." *Journal of Philosophical Logic* 5:25–46.

Hodes, Harold T. 1984. "Some Theorems on the Expressive Limitations of Modal Languages." *Journal of Philosophical Logic* 13:13–26.

Holliday, Wesley H and Perry, John. 2014. "Roles, Rigidity, and Quantification in Epistemic Logic." In Alexandru Baltag and Sonja Smets (eds.), *Johan van Benthem on Logic and Information Dynamics*, 591–629. Cham: Springer International Publishing.

King, Jeffrey C. 2003. "Tense, Modality, and Semantic Values." *Philosophical Perspectives* 17:195–246.

Kocurek, Alexander W. 2015. "On the Expressivity of First-Order Modal Logic with "Actually"." In *Logic, Rationality, and Interaction*, 207–219. Berlin, Heidelberg: Springer Berlin Heidelberg.

Kratzer, Angelika. 1998. "More structural analogies between pronouns and tenses." *Semantics and Linguistic Theory* 8:92.

—. 2007. "Situations in Natural Language Semantics."

Lewis, David K. 1973. *Counterfactuals*. Blackwell Publishing.

—. 1986. *On the Plurality of Worlds*. Blackwell Publishers.

Mackay, John. 2013. "Quantifying over Possibilities." *Philosophical Review* 122:577–617.

Melia, Joseph. 2003. *Modality*. Acumen Press.

Milne, Peter. 1992. "Modal Metaphysics and Comparatives." *Australasian Journal of Philosophy* 70:248–262.

Partee, Barbara H. 1973. "Some structural analogies between tenses and pronouns in English." *The Journal of Philosophy* 70:601.

—. 1984. "Nominal and Temporal Anaphora." *Linguistics and Philosophy* 7:243–286.

Pollock, John L. 1976. *Subjunctive Reasoning*. D. Reidel Publishing Company.

Priest, Graham. 2005. *Towards Non-Being: The Logic and Metaphysics of Intentionality*. Oxford University Press.

Rabinowicz, Wlodek and Segerberg, Krister. 1994. "Actual Truth, Possible Knowledge." *Topoi* 13:101–115.

Schaffer, Jonathan. 2012. "Necessitarian Propositions." *Synthese* 189:119–162.

Schlenker, Philippe. 2006. "Ontological Symmetry in Language: A Brief Manifesto." *Mind* 21:504–539.

Sider, Theodore. 2001. *Four-Dimensionalism: An Ontology of Persistence and Time*. Oxford University Press.

Stone, Matthew. 1997. "The anaphoric parallel between modality and tense." .

von Stechow, Arnim. 1984. "Comparing Semantic Theories of Comparison." *Journal of Semantics* 3:1–77.

Wehmeier, Kai F. 2001. "World Travelling and Mood Swings." In *Trends in Logic*, 257–260. Dordrecht: Springer Netherlands.

—. 2012. "Subjunctivity and Cross-world Predication." *Philosophical Studies* 159:107–122.

Yanovich, Igor. 2015. "Expressive Power of "Now" and "Then" Operators." *Journal of Logic, Language and Information* 24:65–93.