PROPER NAMES AS COUNTERPART-THEORETIC INDIVIDUAL CONCEPTS

James Ravi Kirkpatrick
University of Oxford & Magdalen College, Oxford

Abstract

Many philosophers and linguists have been attracted to counterpart theory as a framework for natural language semantics. I raise a novel problem for counterpart theory involving simple declarative sentences with proper names. To resolve this problem, counterpart theorists must introduce the notion of a counterpart in the semantics of the non-modal fragment of language. I develop my preferred solution: a novel theory of proper names as counterpart-theoretic individual concepts. The resulting view highlights a hitherto unnoticed fact: counterpart theorists should formulate their theory, not by modifying the standard semantics for modal operators, but by modifying the standard semantics for names and variables.

1 Introduction

A central task of natural language semantics is to assign suitable worldly objects as the meanings of expressions in such a way that the meanings of complex expressions are determined by the meanings of their constituents and the rules that combine them. Life would be very simple if we could agree that the meanings of declarative sentences are truth-conditions, intensions, or propositions, perhaps modelled as (the characteristic functions of) sets of possible worlds. Then we could say things like the following:[1]

(1) a. The sentence 'Kim smokes' expresses the proposition that Kim smokes (i.e., $[\![\text{Kim smokes}]\!] = \lambda w.1$ iff Kim smokes in w).

b. The sentence 'Kim smokes' is true at a world w iff Kim smokes in w (i.e., $[\![\text{Kim smokes}]\!]^w = 1$ iff Kim smokes in w).

All that would remain would be to assign suitable meanings to 'Kim' and 'smokes' and formulate suitable compositional rules to combine them.

Life is simple for many of us. But counterpart theorists have to worry about the simplest things. Counterpart theorists believe that possible individuals are *world-bound*; that is, we each exist in at most one possible world.[2] Counterpart theorists also believe that we have *de re* modal

[1] Some remarks on notation. Here and throughout, the characteristic functions of sets are represented by λ-abstractions of the form $\ulcorner\lambda\alpha : \phi\,.\,\gamma\urcorner$. These expressions involve an *argument variable* α, which is bound by a λ operator; an optional *domain condition* ϕ, which specifies relevant constraints on the domain over which the function is defined; and a *value description* γ, which describes the content of the value of the function (cf. Heim & Kratzer, 1998). The characteristic functions of set-of-worlds propositions map worlds to truth-values, that is, they are functions of semantic type $\langle s, t\rangle$, where s is the semantic type of worlds and t is the semantic type of truth-values. The semantic interpretation function $[\![.]\!]^{\mathcal{M},w,g}$ maps expressions of a language to suitable denotations relative to a model $\mathcal{M} = \langle W, D, I, \ldots\rangle$ consisting of at least a non-empty set of possible worlds W, a domain of individuals D, and an interpretation function I mapping basic expressions of the language to suitable meanings; a world w; and an assignment function g. Indices are omitted when obvious from context. I play fast and loose with the distinction between sets and their characteristic functions; officially, I prefer function talk.

[2] See Lewis (1968, 1986). It is useful to distinguish between the basic theory presented in Lewis (1968) and later revisions, extensions, and various other systems. I use the phrase 'counterpart theory' and its cognates to refer to any

properties, not because we have certain properties at other possible worlds, but rather because we have other-worldly counterparts that have those properties. So, according to counterpart theorists, Kim is world-bound, and so she either smokes in the actual world or she doesn't smoke at all. But regardless of whether she actually smokes, Kim could smoke at least in part because she has other-worldly counterparts who do smoke.

Counterpart theorists have to worry because their theoretical commitments sit uneasily with some central tenets of semantic orthodoxy, namely, that names refer to their bearers, that non-modal intransitive verbs denote ordinary properties, and that semantic meaning is compositional. More specifically, counterpart theory and semantic orthodoxy entail that the set-of-worlds proposition expressed by 'Kim smokes' is either a singleton proposition or the empty proposition, depending on whether Kim actually smokes:

(2) $[\![\text{Kim smokes}]\!] = \{w : \text{ 1 iff Kim smokes in } w\} = \begin{cases} \{@\} & \text{if Kim actually smokes} \\ \varnothing & \text{otherwise} \end{cases}$

This proposition looks like a poor candidate for the meaning of 'Kim smokes'. For one thing, we ordinarily think of such sentences as being truth-evaluable and, moreover, true at multiple worlds. But if a set-of-worlds proposition is true at a world w just in case w is an element of that proposition, then 'Kim smokes' is false at any merely possible world, since Kim herself doesn't exist in those worlds.

No solace is found by dropping the relativisation to worlds in (1) to properly reflect the fact that Kim is world-bound and so doesn't have properties at any other world. Consider the following:

(3)
a. The sentence 'Kim smokes' expresses the proposition that Kim smokes (i.e., $\lambda w.$ 1 iff Kim smokes).
b. The sentence 'Kim smokes' is true at a world w iff Kim smokes (i.e., $[\![\text{Kim smokes}]\!]^w =$ 1 iff Kim smokes).

For then it would follow that the sentence 'Kim smokes' expresses either a necessary truth or a necessary falsehood, depending again on whether Kim actually smokes:

(4) $[\![\text{Kim smokes}]\!] = \{w : \text{1 iff Kim smokes}\} = \begin{cases} W & \text{if Kim actually smokes} \\ \varnothing & \text{otherwise} \end{cases}$

This proposition looks like an equally poor candidate for the meaning of 'Kim smokes'. After all, the sentence ends up expressing either the universal proposition, which is true at every world, or the empty proposition, which is true at no worlds. Such a proposal ends up collapsing the distinction between necessity and contingency.

It didn't take much to formulate this problem. All that was needed was counterpart theory, referentialism about proper names, a standard semantics for intransitive predicates, and the claim that semantic meaning is compositional:

---

view in keeping with the spirit of Lewis's original system, but not necessarily with the details. For other proponents of counterpart theory, see, a.o., Bader (2016); D. G. Fara (2008); Forbes (1982, 1985, 1989, 1990); Ramachandran (1989, 1990a, 1990b); Russell (2013); Stalnaker (1986); Varzi (2020); Wang (2015). For dissidents, see, a.o., M. Fara and Williamson (2005); Hawthorne and Yli-Vakkuri (2023); Hazen (1979, 2012). In addition to modal counterpart theories, temporal counterpart theories can also be developed analogously to how temporal logics can be developed in the framework of modal logic; see Sider (1996, 2001).

Counterpart Theory. Possible individuals are world-bound. Possible individuals are related across worlds by 'counterpart relations'. Possible individuals have modal properties in virtue of the non-modal properties that their counterparts have.

Referentialism. Proper names are individual-denoting expressions that rigidly refer to their bearers.[3]

Intransitive Verb Meanings. Intransitive verbs denote non-counterpart-theoretic properties.[4]

Compositionality. Let μ be the function that maps an expression α to its meaning. Then for every syntactic rule σ there is a semantic operation f_σ such that $\mu(\sigma(\alpha_1, \ldots, \alpha_n)) = f_\sigma(\mu(\alpha_1), \ldots, \mu(\alpha_n))$.[5]

To see how these assumptions generate the truth-conditions in (1), suppose that the meaning of 'Kim' is as in (5), the meaning of 'smokes' is as in (6), and take semantic composition to be Intensional Functional Application:[6]

(5) $[\![\text{Kim}]\!]^{w,g} = \text{Kim}$

(6) $[\![\text{smokes}]\!]^{w,g} = \lambda u_{\langle s,e\rangle}.\ 1$ iff $u(w)$ smokes in w.

(7) *Intensional Functional Application.* (after von Fintel & Heim, 2011)
If α is a branching node and $\{\beta, \gamma\}$ the set of its daughters, then, for any possible world w and any assignment function g, if $[\![\beta]\!]^{w,g}$ is a function whose domain contains $\lambda w'.[\![\gamma]\!]^{w',g}$, $[\![\alpha]\!]^{w,g} = [\![\beta]\!]^{w,g}(\lambda w'.[\![\gamma]\!]^{w',g})$.

The most controversial assumption here is certainly counterpart theory and most of us would be quite happy to reject it. But how should a committed counterpart theorist resolve this puzzle?

The goal of this paper is to argue for a novel way of resolving this puzzle by developing a theory of proper names as *counterpart-theoretic individual concepts*. Counterpart theorists normally combine a metaphysical view — that there are distinct possible worlds which do not share their domains of individuals and that individuals stand in counterpart relations to each other — and a semantic view — that the standard semantics of modal operators should be modified to analyse *de re* modal attributions in terms of these counterpart relations between world-bound individuals. This combination of views allows counterpart theorists to deliver intuitive truth-conditions for *de re* modal claims, even though possible individuals do not exist in multiple worlds. But the puzzle above shows that only modifying the semantics of modal operators is not enough to deliver an intuitively adequate account of non-modal content in a counterpart-theoretic setting. Rather than modifying the semantics of modal operators, I will

[3] See, a.o., Kaplan (1989); Kripke (1980); Mill (1843).
[4] See, e.g., Heim and Kratzer (1998).
[5] For a good overview of the arguments in favour and against compositionality, see Szabó (2022).
[6] Compositionality can be captured with a number of semantic operations; nothing is meant to hinge on my choice of Intensional Functional Application. For Intensional Functional Application to combine (5) and (6), we need the denotation of *smokes* to take individual concepts (i.e., functions from worlds to individuals) as argument. Intransitive verb denotations of semantic type $\langle e, t\rangle$ and Functional Application would do just as well for the non-modal fragment of the language.

argue that counterpart theorists should instead modify the orthodox semantics of proper names and variables.

To provide some initial motivation for this claim, let us consider a second puzzle for counterpart theorists involving sentences with modal expressions.[7] Two additional assumptions are needed to motivate the puzzle.

SYNTAX OF MODAL SENTENCES. Modals are expressions that take full sentences (e.g., complementizer phrases (CPs)) as their semantic arguments.[8]

SEMANTICS OF MODALS. Modals denote functions from propositions to truth-values.[9]

Let us assume that modals like *could* have denotations as in (8):

(8) $[\![\text{could}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.1$ iff there exists a world w' accessible from w and $p(w') = 1$.

Here and throughout, 'accessible' is a placeholder that may be spelled out in different ways depending on how the context-sensitivity of the modal is resolved.[10]

Then, assuming (9) is assigned the syntactic structure in (9a), it is assigned the truth-conditions in (10):

(9) Kim could smoke.

a. [could [$_{\text{CP}}$ Kim smoke]]

(10) $[\![$[could [$_{\text{CP}}$ Kim smoke]]$]\!]^{w,g} = 1$ iff there exists a world w' accessible from w and Kim smokes in w'.

These truth-conditions spell trouble for the counterpart theorist. For if Kim is world-bound, she only actually exists, and so if she doesn't actually smoke, then, according to the above truth-conditions, the sentence 'Kim could smoke' is false. Analogous problems arise for necessity modals. A sentence involving a necessity modal is true at a world just in case its prejacent is true in all accessible worlds; and so, if any merely possible world is actually accessible, then there are no *de re* necessities. After all, Kim doesn't smoke in any merely possible accessible world, since she doesn't exist at any world other than the actual one. These assumptions threaten to collapse the distinction between actuality and possibility, and eradicate the concept of *de re* necessity entirely.

Of course, counterpart theorists will immediately object that nothing we have said so far has made room for counterparts — the defining feature of how counterpart theorists handle modal truths — and so, counterparts must be introduced somewhere. But where? Since counterpart theory was originally conceived as a theory about *de re* modal attributions, a natural suggestion would be to reject the standard semantics for modals in favour of a counterpart-theoretic lexical

[7] To the best of my knowledge, this second puzzle was first notice by Elbourne (2013, p. 33, ft. 21).

[8] See, e.g., von Fintel and Heim (2011).

[9] See, e.g., von Fintel and Heim (2011).

[10] This denotation glosses over important features of modality in natural language, such as the effects of context on which set of worlds or notion of accessibility are relevant for the evaluation of the modal; see Kratzer (2012). These simplifications are inconsequential for present purposes.

entry. But, as we shall soon see, making this idea precise is not a trivial task. The important point to appreciate for now is that this move does nothing to resolve our initial puzzle, which makes trouble for counterpart theory without any mention of modality, and so cannot be resolved by adopting a counterpart-theoretic semantics for modality.

The upshot: to resolve both of our puzzles, it seems that counterpart-theoretic natural language semantics must introduce the notion of a counterpart for non-modal portions of natural language. There are many strategies to pursue. Here are the main options:[11]

DENY COMPOSITIONALITY. Give up or significantly alter the compositionality principle.

DENY INTRANSITIVE VERB MEANINGS. Give up the claim that intransitive verb denotations are non-counterpart-theoretic properties.

DENY REFERENTIALISM. Give up the claim that names denote their bearers.

While it is not my primary aim to argue that the first two options are untenable, I will offer reasons why a counterpart theorist shouldn't adopt them. I will then argue that if one wants to keep certain theoretical commitments, denying referentialism is the preferred option.

The question at the heart of this paper should be of interest to philosophers and linguists alike: can counterpart theory serve as an ontology for natural language semantics? Counterpart theory has humble beginnings as an extensional alternative to standard relational semantics for quantified modal logic, one that provides an account of *de re* and *de dicto* modal attributions without assuming transworld identity. But since its conception, philosophers of language and linguists have argued that counterpart theory is needed to adequately capture certain fragments of natural language. Philosophers have been motivated to accept counterpart theory in order to capture true statements of contingent identity (D. G. Fara, 2008; Gibbard, 1975; Lewis, 1971; Stalnaker, 1986). Kratzer deploys counterpart theory in her seminal work on the semantics of conditionals, as well as in her analysis of partial content and 'lumps of thought' (Kratzer, 1979, 1986, 1989, 2012). Counterpart theory has also been used to develop views about the distribution of reflexive and non-reflexive pronouns in so-called 'split-identity' sentences, such as "I dreamed that I was Brigitte Bardot and that I kissed me" (Anand, 2007; Heim, 1994; Lakoff, 1968; Percus & Sauerland, 2003), as well as reference oneself in photographs and belief-contexts (Jackendoff, 1975; van Rooij, 1997). An impressive array of applications, no doubt, but an inability to handle simple declarative sentences seriously stunts any appeal counterpart theory might enjoy.

The paper is structured as follows. Section 2 introduces counterpart theory against the backdrop of Lewis's 1968 presentation. I argue that this theory can resolve the puzzles above, but it is ill-suited for natural language semantics. Along the way, I consider and reject the idea that standard compositional semantics for modals can be adapted by building in a counterpart relation. Section 3 considers whether the puzzles can be resolved by adopting a counterpart-theoretic version of Predicate Abstraction. Section 4 considers whether the puzzles can be resolved by building counterpart-theoretic notions into the meanings of predicates. Section 5 develops my favoured solution, a novel theory of proper names according to which names

[11] A flat-footed response to the second puzzle rejects the claim that there are any true non-trivial *de re* modal claims (Quine, 1953). This is a non-starter for the counterpart theorist and, in any case, it makes no attempt at solving the first puzzle.

denote counterpart-theoretic individual concepts. Section 6 compares my favoured solution to a related but importantly distinct theory. Section 7 concludes.

2 Counterpart Theory and Translation Schemas

Counterpart theory is a family of theories about *de re* modal (or temporal) attributions according to which an individual possibly (temporarily) has a property just in case there is a *counterpart* of that individual which has that property in some world (time). The basic idea is simple enough. Begin with a class of possible worlds and take the domain of individuals associated with each world to be completely disjoint from each other. In a slogan, individuals are *world-bound*. Individuals are nevertheless 'connected' to others by a *counterpart relation*, a qualitative similarity relation over the set of individuals which assigns to each individual (possibly relative to a world) a set of individuals, their *counterparts*. Counterpart theorists then say that what it is for an individual to possibly be some way is just for it to have a counterpart which is that way. For example, Kim could have been the President of the United States just in case she has a counterpart in some world that is the President of the United States. Similar remarks apply to necessity, as well as temporal properties.

What is a counterpart? According to Lewis, the counterpart relation is a matter of purely qualitative similarity:

> The counterpart relation is our substitute for identity between things in different worlds. Where some would say that you are in several worlds, in which you have somewhat different properties and somewhat different things happen to you, I prefer to say that you are in the actual world and no other, but you have counterparts in several other worlds. Your counterparts resemble you closely in content and context in important respects. They resemble you more closely than do the other things in their worlds. (Lewis, 1968, p. 114)

But counterpart theory can be divorced from Lewis's original views about the nature of the counterpart relation and for what it should be used. Lewis himself revised his counterpart theory so as to allow for multiple counterpart relations to track different aspects in which individuals can be qualitatively similar to each other (Lewis, 1971). Other theorists have argued that counterpart theory must hold between *sequences* of individuals (Dorr, 2010; Hazen, 1979; Lewis, 1986; Russell, 2013).[12] Yet more theorists have extended counterpart theory to *de re* temporal predications (Sider, 2001). In essence, a counterpart theory is any theory of *de re* modal (or temporal) predication according to which an individual possibly (or temporally) has a given property just in case that individual is R-related to something that has that property in some possible world (or time), where R is some relation other than identity. Indeed, there need not be a single R relation for the theory to be a counterpart theory; a family of R relations can partly constitute a counterpart theory, so long as one of them is not the identity relation. But for present purposes, it will be useful to focus on Lewis's original system.

David Lewis's 1968 version of counterpart theory was originally proposed as a meaning-preserving translation schema from sentences of quantified modal logic to sentences of counterpart theory, rather than as a semantics of quantified modal logic directly. It is the result of adding

[12] For other versions of counterpart theory, see ft. 2.

to classical first-order predicate logic with identity and without ineliminable singular terms some meaning postulates governing the interaction of some non-logical primitive predicates, and taking the domain of quantification to be the class of all possible individuals (among them, the possible worlds). Simplifying Lewis's presentation somewhat, consider the following primitive predicates of counterpart theory, together with their intended interpretations:[13]

(11) a. Ixy (x is in a possible world y)
b. Ax (x is actual)
c. Cxy (x is a counterpart of y)

The primitives are to be understood according to their English readings and eight meaning postulates. The first two meaning postulates guarantee that if an individual is in anything at all, then it is in exactly one thing. The second two meaning postulates guarantee that if anything is or has a counterpart, then it is in a possible world. The third two meaning postualtes guarantee that everything has exactly one counterpart in the world that it is in: itself. And the last two meaning postulates guarantee that there is a unique, non-empty possible world in which all and only the actual individuals exist: the actual world, @.[14]

Lewis maintains that counterpart theory is superior to quantified modal logic as a means of understanding modal discourse in natural language. This is achieved with the help of a systematic method of translating sentences of English into sentences of counterpart theory. The translation process takes place in two steps. First, an English sentence is translated into quantified modal logic. Second, the sentence of quantified modal logic is translated into a sentence of counterpart theory via the following translation scheme:

Lewis's Translation Scheme

- (T1) A sentence ϕ of the language of quantified modal logic is translated to a sentence $\phi^{@}$ of the language of counterpart theory [read: 'ϕ holds in the actual world'; in primitive notation, $\phi^{@} =_{\text{df.}} \exists v(\forall x(Ixv \equiv Ax) \,\&\, \phi^{v})$];[15]
- For any sentence ϕ of the language of quantified modal logic, and any world-term v, the sentence ϕ^{v} of the language of counterpart theory is defined recursively as follows:

(T2a) ϕ^{v} is ϕ if ϕ is atomic.
(T2b) $(\neg\phi)^{v}$ is $\neg(\phi^{v})$.
(T2c) $(\phi \wedge \psi)^{v}$ is $(\phi^{v} \wedge \psi^{v})$.

[13] Lewis has an additional primitive predicate intended to hold only of possible worlds. Following M. Fara and Williamson (2005) and Hawthorne and Yli-Vakkuri (2023), I depart from Lewis in favour of sorted variables: $v, w, \ldots$ range over possible worlds; $x, y, z, \ldots$ over other possible individuals. Consequently, we can translate formulae containing sorted variables as $\forall w\phi := \forall x(Wx \supset \phi[x/w])$ and $\exists w\phi := \exists x(Wx \wedge \phi[x/w])$, where $\ulcorner\phi[x/w]\urcorner$ is the result of substituting each free occurrence of x in ϕ with a free w. Nothing of substance hinges on this simplification. I am sloppy, here and throughout, with distinguishing between variables in the object language and meta-variables.

[14] The actual world is itself unique; formally @ abbreviates $\imath x\forall y(Iyx \equiv Ay)$ (read: the actual world).

[15] Here, and throughout, I am sloppy about using quotations and Quine-quotations, favouring the use of no quotations when the meaning is clear.

(T2d) $(\phi \supset \psi)^v$ is $(\phi^v \supset \psi^v)$.

(T2e) $(\forall x\phi)^v$ is $\forall x(Ixv \supset \phi^v)$.

(T2f) $(\exists x\phi)^v$ is $\exists x(Ixv \wedge \phi^v)$.

(T2g) $(\Box\phi)^v$, where the unbound terms in ϕ are $a_1, \ldots, a_n$, is
$\forall w\forall x_1 \ldots \forall x_n((Ix_1w \wedge Cx_1a_1 \wedge \ldots \wedge Ix_nw \wedge Cx_na_n) \supset \phi^v[x_1/a_1, \ldots, x_n/a_n])$.

(T2h) $(\Diamond\phi)^v$, where the unbound terms in ϕ are $a_1, \ldots, a_n$, is
$\exists w\exists x_1 \ldots \exists x_n(Ix_1w \wedge Cx_1a_1 \wedge \ldots \wedge Ix_nw \wedge Cx_na_n \wedge \phi^v[x_1/a_1, \ldots, x_n/a_n])$.

How do we translate sentences containing proper names into the language of counterpart theory? A natural suggestion is to extend the languages of quantified modal logic and counterpart theory with ineliminable constants that translate English proper names. But this is problematic. To see this, suppose that 'Sk' is the formalisation of 'Kim smokes' in the language of quantified modal logic. Then, by (T1), its counterpart-theoretic translation is $Sk^{@}$. We expand '@' to get '$\exists w(\forall x(Ixw \equiv Aw)\ \&\ Sk^w)$'. And from (T2a), we have '$\exists w(\forall x(Ixw \equiv Aw)\ \&\ Sk)$'. In English: there exists an actual world and Kim smokes. This strategy is problematic because such counterpart-theoretic translations just do not express what the English sentence seems to express; the latter but not the former expresses a content that can meaningfully be evaluated at different worlds. Consider, for example, if we try to evaluate the counterpart-theoretic translation of 'Kim smokes' at some merely possible world w. The translation would be Sk^w, and so by (T2a), we get Sk, which is true just in case Kim actually smokes. As we saw in the Introduction, this proposal collapses the distinction between necessity and contingency.

Thankfully, Lewis himself does not opt for this treatment of proper names. Instead, he tacitly rejects Referentialism in favour of a different treatment of proper names according to which singular terms are replaced with semantically equivalent definite descriptions (potentially analysing the definite description into primitive language using Russell's method), which he then assigns scope with respect to any modals (Lewis, 1968, pp. 120–121). For example, taking the singular term 'John' as a description $\ulcorner \imath x\ \phi x \urcorner$ with narrow scope, the English sentence (12) gets translated into the counterpart-theoretic sentence (14) via the sentence (13) of quantified modal logic:

(12) John might be tall.

(13) $\Diamond\exists x(\forall y(\phi y \equiv y = x) \wedge Tx)$

(14) $\exists w\exists x(Ixw \wedge \forall y(Iyw \supset (\phi^w y \equiv y = x)) \wedge Tx)$

This is the *de dicto* interpretation; the modal operator attaches to the already closed sentence. In English, (14) states that there is a possible world w containing a unique x such that $\phi^w x$, and for any such x, Tx; or, collapsing the definite description, there is a possible world in which John is tall.

Contrastingly, taking the singular term to have wide scope, (12) gets translated into the counterpart-theoretic sentence (16) via the sentence (15) of quantified modal logic:

(15) $\exists x(\forall y(\phi y \equiv y = x) \wedge \Diamond Tx)$

(16) $\exists x(Ix@ \wedge \forall y(Iy@ \supset \phi^{@} y \equiv y = x) \wedge \exists w\exists z(Izw \wedge Czx \wedge Tx))$

This is the *de re* interpretation; the modal operator attaches to an open sentence to form a new open modal sentence which is then attributed to the actual thing denoted by 'John'. In English, it states that the actual world contains a unique x such that $\phi^{@}x$, and there is some counterpart y thereof in some world w, such that Ty; or, collapsing the definite description, there is a possible world in which a counterpart of John is tall. Consequently, this strategy seems to resolve the second of our puzzles.

How does Lewis's strategy treat sentences with proper names and no modal operators? Since there is no modal for a proper name scope over or under, the sentence 'Kim smokes' is straightforwardly translated as '$\exists x(Ix@ \wedge \phi x \wedge \forall y(Iy@ \wedge \phi y \equiv x = y) \wedge Sx)$'. In English: the unique thing in actuality that is ϕ smokes (where 'ϕ' contains some artificial predicate made from a proper name). This proposal is more successful than the referentialist approach, at least insofar as it creates general content when evaluated across different possible worlds. For example, we evaluate 'Kim smokes' at any merely possible world w by considering the counterpart-theoretic translation 'Sk^{w}', we get '$\exists x(Ixw \wedge \phi x \wedge \forall y(Iyw \wedge \phi y \equiv x = y) \wedge Sx)$'. And this is a perfectly general content. More generally, we get $\{w : \exists x(Ixw \wedge \phi x \wedge \forall y(Iyw \wedge \phi y \equiv x = y) \wedge Sx)\}$, as the set-of-worlds-proposition that 'Kim smokes' expresses. So far, so good.

However, trouble still lurks. The interest of philosophers of language and linguists in counterpart theory resides in its potential to provide a more flexible metaphysical foundation to natural language semantics than competing relational semantics for modal logic. But a friend of natural language semantics should not be happy with using Lewis's translation schema for such purposes. While translation schemas were standardly used in the early days of formal semantics (cf. Link, 1983; Montague, 1970), this practice has long since been eschewed in favour of directly assigning suitable compositional semantic values to lexical items of the object language. Natural language semantics now proceeds by, first, syntactically associating a lexically and perhaps structurally disambiguated structure with each occurrence of a natural language expression, which may differ from its apparent phonetic structure, and which is the primary object of semantic interpretation; then, each atomic lexical morpheme is assigned a denotation or intention by semantic interpretation and the meaning of each complex expression is given according to the meanings of its constituent parts, together with semantic composition rules that track syntactic configurations (cf. Heim & Kratzer, 1998). If counterpart theory is to serve as an ontological foundation for natural language semantics, we should depart from Lewis's translation schema approach and instead directly assign counterpart-theoretic lexical meanings to each atomic expression of our natural language together with meaning composition rules.

But it is not clear how Lewis's translation clauses can be adapted for this task. Lewis's translation clauses for sentences of the form $\ulcorner\Box\phi\urcorner$ and $\ulcorner\Diamond\phi\urcorner$ are actually enumerably many translation clauses, one for any n-ary open formula under the scope of a single, initial modal operator $\Box$ or $\Diamond$. For example, a closed sentence of quantified modal logic with a single, sentence-initial modal operator would be $\ulcorner\forall w\phi^{w}\urcorner$ or $\ulcorner\exists w\phi^{w}\urcorner$; the translation for a 1-place open sentence with a single, initial modal operator is $\ulcorner\forall w\forall x_1(Ixw \wedge Cx_1x \supset \phi^{w}[x_1/x])\urcorner$ or $\ulcorner\exists w\exists x_1(Ixw \wedge Cx_1x \wedge \phi^{w}[x_1/x])\urcorner$, where x is an unbound term in ϕ; and so on. The trouble is that the semantic value for the modal operators, as a result of what I just outlined, needs to be sensitive to the number of individual constants or free variables that occur inside its scope, so that it knows how many counterparts (or what length of counterpart sequences) to consider. But that information cannot be read off from the semantic value of the prejacent, since this is meant to denote a proposition and a proposition does not encode the number of individual constants or free variables that occur in the sentence which expresses it.

More carefully, under standard semantic analyses of modal auxiliaries, their meanings are taken to be functions from propositions to truth-values, such as the following:

(17) $[\![\text{might}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.1$ iff there exists a world w' accessible from w such that $p(w') = 1$.

(18) $[\![\text{must}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.1$ iff for every world w' accessible from w, $p(w') = 1$.

These lexical entries make clear that, once the semantic value of the prejacent has been computed, the semantic clause for the modal has no access to the number of individual constants or free variables that occur inside its scope.

One approach in the spirit of Lewis's translation schema would be to posit enumerably many modal auxiliaries, each distinguished only in terms of the arity of its prejacent, and each of which assigned a distinct lexical meaning, again depending only on the arity of its prejacent. In effect, we should augment our modal language with enumerably many $\Box$'s and $\Diamond$'s, each superscripted with an arity to indicate the 'arity' of the sentence which it may select as argument. For example, let $\Box^0$ and $\Diamond^0$ be modal operators that take closed sentences, $\Box^1$ and $\Diamond^1$ be modal operators that take 1-place open sentences, and so on. We may then fully specify meanings for these new modals in the following manner:[16]

(19)
a. $(\Box^0\phi)^v$ is $\forall w\phi^w$;
b. $(\Diamond^0\phi)^v$ is $\exists w\phi^w$;
c. $(\Box^1\phi)^v$ is $\forall w\forall x(Ixw \wedge Cxa \supset \phi^w[x/a])$, where a is an unbound term in ϕ;
d. $(\Diamond^1\phi)^v$ is $\exists w\exists x(Ixw \wedge Cxa \wedge \phi^w[x/a])$, where a is an unbound term in ϕ;
...and so on...

Such an approach would require some explanation for why modals with a certain arity index cannot compose with prejacents of a different arity. But more problematically, any such multiplication of expressions and meanings beyond necessity is an unparisonious extravagance that we should try to do without. After all, natural languages must be learnable, and languages with infinitely many primitive expressions are unlearnable.[17]

How, then, should counterpart theorists resolve the puzzles? Observe that Lewis's translation schemas for modal sentences essentially perform two functions. The first function is to introduce quantification over a special class of possible individuals, namely, the possible worlds. The second function is to introduce quantification over counterparts. I have been arguing that

[16] More generally, we have:

(i) $(\Box^n\phi)^v$ is $\forall w\forall x_1\ldots\forall x_n(Ix_1w \wedge Cx_1a_1 \wedge \ldots \wedge Ix_nw \wedge Cx_na_n \supset \phi^w[x_1/a_1,\ldots,x_n/a_n])$, where $a_1,\ldots,a_n$ are unbound terms in ϕ;

(ii) $(\Diamond^n\phi)^v$ is $\exists w\forall x_1\ldots\forall x_n(Ix_1w \wedge Cx_1a_1 \wedge \ldots \wedge Ix_nw \wedge Cx_na_n \wedge \phi^w[x_1/a_1,\ldots,x_n/a_n])$, where $a_1,\ldots,a_n$ are unbound terms in ϕ.

[17] Similar remarks apply to alternative translation schemas, such as those proposed by Forbes (1982, 1990) and Ramachandran (1989, 1990b), as well as to the syncategorematical semantics, such as that proposed by, e.g., Forbes (1985).

syncategorematical rules are ill-suited for compositional semantics and direct implementing both of these two functions into standard semantic treatment for modals leads to infinitely many clauses for language users to learn. Consequently, I propose that we depart from Lewis's translation schema and move to a more familiar contemporary setting to investigate the possibility of counterpart-theoretic formal semantics. But there is no reason to encoded both functions with the same expression. If we remove the counterpart relation from the semantic values of modals, they can just focus on quantifying over worlds without having to worry about also introducing an appropriate number of counterpart relations and some other expression or rule can do the rest. The question remaining is this: where should we introduce counterparts into the meanings of expressions? Sections 3–5 explore three different places to encode counterpart relations: composition rules, predicates, and proper names.

3 Counterparts in Predicate Abstraction

Let us consider whether we can introduce counterparts in our composition rules without rejecting the spirit of compositionality entirely. As we have seen, Lewis's translation schemas analyse *de re* modal attributions by taking proper names (construed as definite descriptions) to scope over modal operators that embed them. One way to replicate Lewis's distinction between *de re* and *de dicto* modal claims involving proper names is to introduce counterparts into semantic derivations when a subject term scopes over a modal. This amounts to rejecting the claim that the following syntactic structures are semantically equivalent:

(20) [could [Kim smoke]]

(21) [Kim [λ_1 [could [t_1 smoke]]]]

The semantic equivalence of (20) and (21) is standardly established by adopting the following rules governing the semantic behaviour of predicate abstraction operators and traces:[18]

(22) *Predicate Abstraction Rule* (after Heim & Kratzer, 1998)
For all worlds w, indices i, and assignments g, $[\![\lambda_i\ \alpha]\!]^{w,g} = \lambda x_e.[\![\alpha]\!]^{w,g[x/i]}$, where $g[x/i]$ is an assignment exactly like g except in maybe assigning x to i.

(23) *Traces and Pronouns Rule* (after Heim & Kratzer, 1998)
For any world w, if α is a pronoun or a trace, g is a variable assignment, and $i \in \text{dom}(g)$, then $[\![\alpha_i]\!]^{w,g} = g(i)$.

An immediate observation is that, in the present setting, pronouns and traces are rigid in the sense that they designate the same object (relative to a given variable assignment) in every possible world it exists, if they designate anything at all in that world, and their value is determined solely by the assignment function.

The most natural way to establish the non-equivalence of (20) and (21) is by following Lewis's example and drawing a semantic difference from scopal differences. This can be achieved by revising the Predicate Abstraction rule so that it introduces counterparts itself.

[18] Strictly speaking, since Predicate Abstraction is a syncategorematic rule, our system is no longer fully compositional. This can easily be remedied by providing an appropriate lexical entry for λ-operators in the object language (cf. Rabern, 2013, p. 399).

We begin by introducing a counterpart function $\mathcal{C}_w$ that maps any suitable individual to its counterpart in the possible world w:

(24) $\mathcal{C}_w$ maps individual x to $\mathcal{C}_w(x)$, the counterpart of x in world w.

For simplicity, it will be useful to assume that individuals have at most one counterpart at any world in which it has any counterparts at all, though nothing hinges on this simplification.[19] To accommodate multiple counterparts at a world, we can always take $\mathcal{C}_w$ to be a function from individuals to sets of individuals (intuitively, the set of its counterparts at w), modifying other lexical entries as required. We also take $\mathcal{C}_w$ to be a partial function to allow individuals to fail to have a counterpart at a world, and we assume that an individual is its own counterpart in the world it inhabits, that is, if x is in w, then $\mathcal{C}_w(x) = x$. Otherwise, we leave open the question of what kind of counterpart relation is relevant.

We can then state the revised Predicate Abstraction Rule as follows:[20]

(25) *Predicate Abstraction Rule (Revised)*
For all worlds w, indices i, and assignments g, $[\![\lambda_i\ \alpha]\!]^{w,g} = \lambda x_e.[\![\alpha]\!]^{w,g[\mathcal{C}_w(x)/i]}$.

In contrast with our original Predicate Abstraction Rule, the revised rule makes certain assignments to variables non-rigid by allowing the counterpart function to determine the value of a pronoun or trace relative to a world. To see this, observe how the revised Predicate Abstraction Rule applies to the λ-expression in (21):

(26) $[\![[\lambda_1$ [could [t_1 smokes]]]$]\!]^{w,g}$
$= \lambda x_e.[\![$[could [t_1 smokes]]$]\!]^{w,g[\mathcal{C}_w(x)/1]}$
$= \lambda x_e.$ there exists a world w' accessible from w such that $\mathcal{C}_{w'}(x)$ smokes in w'.

This denotation allows us to correctly compute the relevant *de re* interpretation for (21):

(27) $[\![(21)]\!]^{w,g} = 1$ iff there exists a world w' accessible from w such that $\mathcal{C}_{w'}$(Kim) smokes in w'.

This proposal also generates intuitively adequate meanings for non-modal declarative sentences with proper names like (28), just so long as we reject syntactic structures like (29) in favour of ones in which any proper name is obligatorily raised out of initial subject position to get structures like (30):

(28) Kim smokes.

---

[19] See Sections 5.2 and 5.3 for discussion of how to lift this requirement.

[20] Alternatively, we might adopt the following Predicate Abstraction Rule:

(i) *Predicate Abstraction Rule (Alternative Revision)*
For all worlds w, indices i, and assignments g, $[\![\lambda_i\ \alpha]\!]^{w,g} = \lambda x_e.\exists y(Iyw \wedge Cyx \wedge [\![\alpha]\!]^{w,g[y/i]})$.

This allows us to take (ii) to be the syntactic structure of 'Kim could smoke' with the final truth-conditions as in (iii):

(ii) [could [Kim [λ_1 [t_1 smoke]]]]

(iii) $[\![(ii)]\!] = 1$ iff there exists a world w' accessible from w such that $\exists y(Iyw' \wedge Cyk \wedge \text{smokes}(y))$.

(29) [Kim smokes]

(30) [Kim [λ_1 [t_1 smokes]]]

By applying the revised Predicate Abstraction rule to introduce a counterpart function into traces bound by λ-terms, (30) has the following meaning:

(31) $[\![$[Kim [λ_1 [t_1 smokes]]]$]\!]^g = \lambda w.1$ iff $\mathcal{C}_w$(Kim) smokes in w.

Such propositions are true at world just in case the counterpart of Kim in that world smokes. This seems general enough to play the role of sentence meaning.

But this strategy is somewhat unnatural. While one could motivate the claim that (20) and (21) are not semantically equivalent on the grounds that the proper name only receives a *de re* reading when it scopes over an intensional operator, such explanations are not available for non-modal cases where there is no intensional operator to which a proper name can stand in scopal relations. The idea that (29) and (30) are semantically different merely on the grounds that the proper name in the latter has been raised is theoretically unsatisfactory.

A more serious problem is that this strategy is not general enough to be an adequate solution for our puzzles. Certain syntactic environments do not allow expressions to scope out of them. For example, it is commonly assumed that the restrictive clause of a conditional is a *scope island*, meaning that expressions normally cannot scope out of it. But consider the following sentence:

(32) It could be that, if John comes, Sally won't come.

There is an epistemic reading of this sentence according to which, for all the speaker knows, the conditional could be true. Supposing that John actually comes and Sally doesn't, the epistemic reading is true only if there is world compatible with what the speaker knows in which either a counterpart of John doesn't come or a counterpart of Sally does. But the Predicate Abstraction strategy is unable to capture this reading. This is because the *if*-clause is a scope island, meaning *John* can't scope out of the clause that it is embedded within through Predicate Abstraction. Since the Predication Abstraction strategy is unable to capture all desired readings of sentences with proper names embedded under modals, we turn now to more other strategies.

4 Counterparts in Predicates

One might reject the standard semantics for intransitive verbs in favour of introducing counterparts in the lexical entries of predicates (Heim, 2001; Percus, 1998; Sauerland, 2014). One way to make this idea precise would be to integrate the previously introduced counterpart function directly into predicate denotations themselves, rather than relying on a separate composition rule to introduce it. The idea would be to apply the counterpart function to the subject argument of intransitive verbs, like in the following:[21]

[21] In fact, one could take advantage of the fact that the counterpart function introduces a bindable world variable and drop the usual relativisation to worlds from the value description of the λ-term. The result would be denotations like the follows:

(33) $[\![\text{smoke}]\!]^w = \lambda x_e.1 \text{ iff } \mathcal{C}_w(x) \text{ smokes in } w.$

Consequently, the sentences 'Kim smokes' and 'Kim could smoke' would have the following truth-conditions:

(34) $[\![\text{Kim smokes}]\!]^w = 1 \text{ iff } \mathcal{C}_w(\text{Kim}) \text{ smokes in } w.$

(35) $[\![[\text{could } [_{\text{CP}} \text{ Kim smoke}]]]\!]^w = 1$ iff there exists a world w' accessible from w such that $C_{w'}(\text{Kim})$ smokes in w'.

This theory provides an empirically adequate solution to both of our puzzles. First, (34) says that 'Kim smokes' is true at a world w just in case the counterpart of Kim at w smokes. If Kim is the counterpart of herself in actuality, then these truth-conditions correctly predict that the sentence is actually true just in case Kim smokes. And, according to this account, the meaning of 'Kim smokes' is the sets-of-worlds proposition $\{w : 1 \text{ iff } \mathcal{C}_w(\text{Kim}) \text{ smokes in } w\}$, which is sufficiently general to be non-trivially evaluated for truth at merely possible worlds. Second, (35) says that 'Kim could smoke' is true at a world w just in case there is some world w' accessible from w such that the counterpart of Kim at w' smokes. These truth-conditions are clearly empirically adequate.

The major concern I have with this proposal is that it seems to turn intrinsic properties into non-intrinsic properties. Intrinsic properties hold of an individual in virtue of how that individual is in themselves, and not because of how they stand in relation to other things. In contrast, extrinsic properties hold of individuals in virtue of how they stand in relation to other individuals. Here's Lewis on the matter:

> A sentence or statement or proposition that ascribes intrinsic properties to something is entirely about that thing; whereas an ascription of extrinsic properties to something is not entirely about that thing, though it may well be about some larger whole which includes that thing as part. (Lewis, 1983, p. 197)

If there are any intrinsic properties at all, the properties postulated by our best theories of fundamental physics are as good of a candidate as any. Being a quark or having spin are intrinsic properties, if anything is. But on the present proposal, such properties are non-intrinsic. Consider the following denotation:

(36) $[\![\text{is a quark}]\!]^w = \lambda x_e.\ 1 \text{ iff } \mathcal{C}_w(x) \text{ is a quark.}$

Here we have essentially turned an intrinsic property into an extrinsic property. In order for the counterpart function to apply to the subject argument in the value description of the predicate denotation, its world variable must either be bound or else be supplied by the context. And so the property of being a quark holds of an individual just in case its counterpart at a world is a quark.

(i) $[\![\text{smoke}]\!]^w = \lambda x_e.1 \text{ iff } \mathcal{C}_w(x) \text{ smokes.}$

Such denotations bring the application of counterpart theory to natural language semantics much closer to Lewis's original treatment of atomic formulae.

One response to this argument would be to distinguish between intrinsic properties as such and the denotations of predicates.[22] That is, one might maintain that intrinsic properties, such as those represented in (37), exist, but disagree that our predicates, such as *has charge* or *is a quark* express them.

(37) a. $\lambda x_e.x$ has charge
b. $\lambda x_e.x$ is a quark

This response allows the counterpart theorist to answer our two puzzles without rejecting the existence of intrinsic properties. While I have no strong objections to this approach, I would prefer my counterpart theory to permit me to actually talk about intrinsic properties with my predicates, which is ruled out by this response. So, without further ado, I turn now to developing my favoured solution.

5 Counterpart-Theoretic Individual Concepts

How can we account for reference in modal contexts from the perspective of counterpart theory? Consider again the sentence, 'Kim smokes', and recall that the referentialist takes 'Kim' to refer to Kim in every world in which she exists. Together with a counterpart-theoretic metaphysics, this semantic theory of proper names spells doom for the prospect of delivering an intuitively acceptable notion of content for non-modal sentences and non-trivially true *de re* modal sentences alike.

However, instead of taking names to rigidly referring to a single individual, we could say that 'Kim' picks out Kim in the worlds she in which she does exist, namely, the actual world. But, for worlds in which Kim doesn't exist, 'Kim' picks out the counterpart of Kim at that world, if there is one. Since Kim is a counterpart of herself in the world in which she exists, we can streamline the proposal and say that the denotation of 'Kim' at a world w is a counterpart of Kim at w.

This proposal is reminiscent of Lewis's preferred view of proper names, at least as stated in his 'General Semantics':[23]

> Common nouns also have different extensions at different worlds; *and so do some names*, at least if we adopt the position (defended in Lewis, [(1968)]) that things are related to their counterparts in other worlds by ties of strong similarity rather than identity. (Lewis, 1970, p. 24; my *emphasis*)

This remark was never developed in any serious detail by Lewis. However, I contend that it provides a general insight about how to reconcile counterpart theory and a compositional semantics for natural language.

This core idea can be made precise in different ways. In what follows, I make use of a particular combination of syntactic and semantic assumptions that I feel are best suited to supplement my hypothesis. These assumptions do not represent the only way to cast the theory I

[22] Thanks to Nicholas K. Jones for pushing me to consider this response.

[23] Kratzer also seems to prefer this view, although she gives no compositional semantics for delivering it; see Kratzer (1989, p. 619). This suggestion is also made, but not endorsed, by Elbourne (2013, p. 33, ft. 21).

develop, and the success of the basic argument is compatible with other alternative assumptions. I begin by sketching out the basic idea (Section 5.1), before developing the idea in ways more in keeping with the main motivations for counterpart theory, as well as general discussion of its relation to rigidity and systems of modal logic.

5.1 The basic idea

A first pass at developing the Lewisian idea suggests the following as a faithful counterpart-theoretic lexical entry for the name 'Kim':

(38) $[\![\text{Kim}]\!] = \lambda w_s.\iota x_e\ x = \mathcal{C}_w(\text{Kim})$

Let's explain this lexical entry in stages. First, according to this theory, proper names denote individual concepts, that is, functions from worlds to individuals. In the language of semantic types, proper names denote functions of type $\langle s, e\rangle$. In a Carnapian sense, individual concepts can be seen as the intensionalisation of an individual expression denotation. Consequently, given Referentialism, the individual concept of 'Kim' maps every world to Kim, that is, it is the function λw.Kim. But the individual concept in (38) is much less trivial.

Second, let us examine the value description of the denotation in (38). The formula in (38) makes use of the iota operator ι. The intended interpretation of $\ulcorner \iota v\ \phi \urcorner$ is the unique v that satisfies ϕ, if one exists; it is undefined if there are no such individuals or there are multiple such individuals. So, $\iota x\ x = \mathcal{C}_w(\text{Kim})$ is the unique individual k such that k is the counterpart of Kim in w, if one exists. If either no individual or more than one individual satisfies the descriptive material, then the ι operator is undefined.

Third, the descriptive material in (38) uses the counterpart function, $\mathcal{C}_w(x)$, as introduced above, which maps an individual x to its counterpart in w, if there is one. So if x is in the domain of C_w, the uniqueness and existence conditions of the ι operator are guaranteed to be satisfied. While this proposal is not neutral about the particularities of the counterpart relation, for present purposes it is the most simple one. Nothing hinges on this assumption: more exotic counterpart relations can be evoked as needed.

An immediate consequence of taking proper names to denote individual concepts is that intransitive verb and noun denotations must be treated as functions that take individual concepts as argument and map them to propositions. For example, we have (39) for *smokes*:

(39) $[\![\text{smokes}]\!] = \lambda u_{\langle s,e\rangle}.\lambda w_s.1$ iff $u(w)$ smokes.

Note that to respect the considerations about intrinsic properties from the previous section, we resist the temptation to include a relativisation to worlds in the value description of the λ-term.

Let us now see whether this proposal resolves our puzzles. First, let us plug in our new lexical entry for 'Kim' into the denotation for 'smokes' and calculate the truth-conditions for the non-modal sentence 'Kim smokes':

(40) $[\![\text{Kim smokes}]\!]^w = 1$ iff $\iota x\ x = \mathcal{C}_w(\text{Kim})$ smokes.

According to this view, the set-of-worlds proposition expressed by 'Kim smokes' is $\{w :$ 1 iff $\iota x\ x = \mathcal{C}_w(\text{Kim})$ smokes$\}$, which is suitably general enough to be assessed non-trivially at merely possible worlds. Intuitively, 'Kim smokes' expresses a proposition that is true at any world in which Kim has a unique counterpart at that world and that counterpart smokes. And this is exactly what this theory predicts. The upshot is that non-modal sentences still expresses a proposition that is non-trivially true or false at non-actual worlds.

This account also accounts for the truth-conditions of *de re* modal attributions like 'Kim could smoke':

(41) $[\![\text{Kim could smoke}]\!]^w = 1$ iff there exists a world w' accessible from w such that $\iota x\ x = \mathcal{C}_{w'}(\text{Kim})$ smokes.

According to this view, 'Kim could smoke' is true just in case Kim has a counterpart in some accessible world who smokes. This adequately captures the fact even if Kim does not actually smoke, the sentence 'Kim could smoke' can be true, all without modifying the standard semantics for modal operators.

This strategy has none of the flaws of the previous accounts. First, the semantics is fully compositional. Second, it provides suitably general content as the proposition expressed by non-modal declaratives, like 'Kim smokes'. Third, it provides intuitively adequate truth-conditions for *de re* modal attributions, like 'Kim might smoke'. Fourth, this theory preserves the intuition that there are intrinsic properties and we can speak about them. Fifth, and most remarkably, this version of counterpart theory is broadly a theory about our practice of proper names and reference, rather than about *de re* modality per se. According to this theory, proper names refer to us in actuality and to our counterparts in other worlds, and it is through this feature of reference that we can correctly attribute modal properties. My theory stands in contrast with how counterpart theory has been standardly developed. Counterpart theorists are typically committed to a metaphysical claim, that there are distinct possible worlds which do not share their domains of individuals, and a semantic claim, that *de re* modal claims are analyzed in terms of a counterpart relation between individuals. This semantic story is standardly implemented by modifying the semantic clauses for modal operators. In contrast, the version of counterpart theory developed here implements the semantic story by modifying the semantics of terms. I have argued that this delivers a more intuitively acceptable account of content. This concludes the exposition of the basic idea.

5.2 Multiple counterparts, one relation

My framework is remarkably flexible and can be adapted in a variety of different ways. For example, we can drop the restriction that an individual has at most one counterpart in a world. For what reason? Here's Lewis's original motivation:

> It would not have been plausible to postulate that nothing in any world had more than one counterpart in any other world. Suppose x_{4a} and x_{4b} in w_4 are twins; both resemble you closely; both resemble you far more closely than anything else in w_4 does; both resemble you equally. If so, both are your counterparts. (Lewis, 1968, p. 116)

Good point. There are some tricky technical difficulties in how to best implement the idea that an individual can have more than one counterpart at a world, and there are several ways to do this. In what follows, I have tried to develop the most simple way that accounts for all the relevant facts.

Let's fix ideas. Suppose Kim's zygote might have split, which would have resulted in twinning, but didn't. Then Kim might have been twins, but not actually. Let w_4 be a world in which Kim's zygote split; the twins x_{4a} and x_{4b} her resulting counterparts. Now is the sentence 'Kim smokes' true at w_4, false at w_4, or neither? To my ear, this sentence sounds odd, though if both x_{4a} and x_{4b} smoke, I am happy enough to say it's true. Equally, if both x_{4a} and x_{4b} don't smoke, I'd probably say that 'Kim smokes' is false at w_4. Otherwise, it's neither true nor false. Taking these judgements as verdical, the truth-conditions of 'Kim smokes' at w_4 are as follows:[24]

(42) $$[\![\text{Kim smokes}]\!]^{w_4} = \begin{cases} 1 & \text{if both } x_{4a} \text{ and } x_{4b} \text{ smoke} \\ 0 & \text{if neither } x_{4a} \text{ and } x_{4b} \text{ smoke} \\ \# & \text{otherwise} \end{cases}$$

How does this sentence embed under modal operators? To my ear, modal operators seem to remove the trivalent effects we just observed in their prejacents. For example, *de re* possibility claims like 'Kim might smoke' do not require there to be some accessible world in which all of Kim's counterparts smoke; a single counterpart smoker at some accessible world is enough. But *de re* necessity claims like 'Kim must smoke' seem to be true only if all of Kim's counterparts at any accessible world to smoke.[25]

(43) $[\![\text{Kim might smoke}]\!]^{w} = 1$ iff there exists a world w' accessible from w such that there is at least one counterpart of Kim at w' who smokes

(44) $[\![\text{Kim must smoke}]\!]^{w} = 1$ iff, for every world w' accessible from w, every counterpart of Kim at w' smokes.

To build up to these truth-conditions compositionally, let's first start by modifying our counterpart functions to allow for multiple counterparts at a world. Rather than taking $\mathcal{C}_w$ to be a function from possible individuals to possible individuals, let us utilize a function $\mathbb{C}_w$ from possible individuals to *sets of possible individuals*, namely, the set of counterparts of the input individual at w. So, for example, the set of Kim's counterparts at w_4 is $\mathbb{C}_{w_4}(\text{Kim}) = \{x_{4a}, x_{4b}\}$, as prescribed by Lewis.

There are actually several ways of analysing names using this function, but let's try the following idea: a name denotes a function from worlds to the set of counterparts of its actual bearer in that world. (Not quite a counterpart-theoretic *individual* concept; more like a counterpart-theoretic *sets-of-individuals* concept. But no matter.) Then 'Kim' will have the following denotation:

---

[24] Intuitions may vary. Some may think that no sentence is true nor false at worlds in which the individual in question has more than one counterpart. But I'm happy enough to concede truth or falsity where all counterparts agree. Others might think that a sentence is true at such worlds just in case at least one counterpart has the property in questions. But I'm squeamish about the violations of the law of non-contradiction that result.

[25] Again, tastes may vary. Some might think that *de re* necessity attributions require that, for every accessible world, at least one counterpart has the property in question. I prefer my truth-conditions because they align better with Lewis's original counterpart-theoretic translation clause for modal operators.

(45) $[\![\text{Kim}]\!] = \lambda w.\mathbb{C}_w(\text{Kim})$

This function takes a world and returns the set of Kim's counterparts at that world; its semantic type is $\langle s, E\rangle$, where E is the type of sets of entities.

With the denotations of proper names and sentences in hand, we should be able to figure out the denotation of predicates. Intransitive verbs will denote functions from sets-of-individuals concepts to propositions, since names denote sets-of-individuals concepts and declarative sentences denote propositions. But note that the following doesn't work:

(46) $[\![\text{smokes}]\!] = \lambda u_{\langle s,E\rangle}.\lambda w.1$ iff for every $x \in u(w)$, x smokes.

This lexical entry works perfectly when there is at most one counterpart x at the world of evaluation. And it correctly predicts the truth-conditions of 'Kim smokes' at w_4 when both of Kim's counterparts at w_4 smoke. But it doesn't get things right if one or more of them don't smoke. Better to build trivalence directly into the meaning of predicates as follows:

(47) $$[\![\text{smokes}]\!] = \lambda u_{\langle s,E\rangle}.\lambda w. \begin{cases} 1 & \text{if } x \text{ smokes, for every } x \in u(w) \\ 0 & \text{if } x \text{ doesn't smoke, for every } x \in u(w) \\ \# & \text{otherwise} \end{cases}$$

This correctly predicts the truth-conditions in (42). And, assuming Strong Kleene truth-tables for the propositional connectives, this idea also delivers the right result for compound sentences. Consider, for example, the truth-conditions for 'Kim doesn't smoke' below:

(48) $$[\![\text{Kim doesn't smoke}]\!]^{w_4} = \begin{cases} 1 & \text{if } x \text{ doesn't smoke, for every } x \in \mathbb{C}_w(\text{Kim}) \\ 0 & \text{if } x \text{ smokes, for every } x \in \mathbb{C}_w(\text{Kim}) \\ \# & \text{otherwise} \end{cases}$$

Similar remarks apply to the other propositional connectives.

But the denotation in (45) overlooks the nuances that arise when something has no counterparts at a world. These are cases that we cannot reasonably ignore:

> It would not have been plausible to postulate that, for any two worlds, anything in one had some counterpart in the other. Suppose whatever thing x_6 in world w_6 it is that resembles you more closely than anything else in w_6 is nevertheless quite unlike you; nothing in w_6 resembles you at all closely. If so, you have no counterpart in w_6. (Lewis, 1968, p. 116)

To see the problem, suppose that Kim has no counterparts in w_6; then $\mathbb{C}_w(\text{Kim}) = \varnothing$. It follows that 'Kim smokes' is false at w_6, since for every $x \in \varnothing$, it is not the case that x smokes. After all, there is no such x. Equally, since for every $x \in \varnothing$, it is not the case that x doesn't smoke, 'Kim doesn't smoke' is also false at x_6. Some may be happy with these verdicts, but I hear such sentences as neither true nor false.

So I modify (45) to allow failure of reference:

(49) $[\![\text{Kim}]\!] = \lambda w : \exists x \in \mathbb{C}_w(\text{Kim}) \,.\, \mathbb{C}_w(\text{Kim})$

Now if Kim has no counterparts at the world in question, the evaluation of 'Kim' is undefined, and semantic composition grinds to a halt.

Moving on to modal attributions, we can analyse *de re* necessities with lexical entries like the following:

(50) $[\![\text{must } \phi]\!]^w = 1$ iff for every world w' accessible from w, $[\![\phi]\!]^{w'} = 1$

This adequately captures our judgement in (44).

Things are trickier with *de re* possibilities. Note that the standard semantics for 'might' doesn't work:

(51) $[\![\text{might } \phi]\!]^w = 1$ iff there is at least some world w' accessible from w such that $[\![\phi]\!]^{w'} = 1$

To see the issue, suppose that no counterpart of Kim smokes *except* for one of the twins in w_4. Intuitively, 'Kim might smoke' should come out true, as she has a counterpart that does smoke, but given (51), the sentence is predicted to be false as there is no world in which all of Kim's counterparts smoke:

(52) $[\![\text{Kim might smoke}]\!]^w = 1$ iff there is at least some world w' accessible from w such that, for every $x \in \mathbb{C}_{w'}(\text{Kim})$, x smokes

An analogous point holds for 'Kim might not smoke': suppose that Kim and all her counterparts except one of the twins at w_4 smoke; then the above semantics predicts that 'Kim might not smoke' is false, even though she is a counterpart who doesn't smoke.

A better way to handle *de re* possibilities is to use the following semantic analysis for 'might':

(53) $[\![\text{might } \phi]\!]^w = 1$ iff not all worlds w' accessible from w are such that $[\![\phi]\!]^{w'} = 0$

For the bivalent fragment of our language, this lexical entry is equivalent to the above; the difference between (51) and (53) does not show up. But it makes a difference in the trivalent fragment of our language, since truth and falsity aren't mutually exhaustive. To see that (53) delivers the correct predication in these cases, observe:

(54) $[\![\text{Kim might smoke}]\!]^w = 1$ iff not all worlds w' accessible from w are such that, for every $x \in \mathbb{C}_{w'}(\text{Kim})$, x doesn't smoke

Moving around the negations and flipping the quantifiers, we can see that 'Kim might smoke' is true at a world w just in case there is some accessible world w' such that a counterpart of Kim's at w' smokes. This is intuitively correct.

5.3 Multiple counterparts, multiple relations

So much for having multiple individuals at a world by standing in the same counterpart relation to them. But you can also have multiple counterparts at a world by standing in *different* counterpart relation to them.

A prominent motivation for counterpart theory involves puzzles of material constitution like that of the exactly coincident statue Goliath and lump of clay Lumpl (Gibbard, 1975; Lewis, 1971). Counterpart theorists can maintain that there is only one actual material object (i.e., Goliath = Lumpl), while also maintaining that the statue could have been distinct from the lump, by proposing that the actual object has a statue counterpart and a distinct lump counterpart at certain worlds. This explains why Goliath and Lumpl seem to have different modal properties (i.e., Lumpl could have survived being squashed, but Goliath could not have survived being squashed), even though they are actually identical, that is, there are worlds in which the lump counterpart survives being squashed, but the statue counterpart does not.

Such accounts have been explored in other contexts, and so the comparative advantages and disadvantages won't be extolled here. But it is important to show that my theory can be extended to allow for a multiplicity of counterpart relations. Rather than having only a single counterpart function that is a matter of overall comparative similarity, let us have indefinitely many counterpart functions, each of which focus on the relative importances of different respects of similarity and dissimilarity. Two important respects of similarity and dissimilarity among material objects are form and matter. Assigning great weight to form, we get the *form counterpart* relation. Only something with the same form, or very much like the same form, can resemble something enough to be its form counterpart. And assigning great weight to matter, we get the *matter counterpart* relation. Only something with the same matter, or very much the same matter, can resemble something enough to be its matter counterpart.

If Goliath is Lumpl, then in many worlds there are things that are both form and matter counterparts of that object. But in other worlds, that object has form counterparts that aren't matter counterparts; or matter counterparts that aren't form counterparts. Since these counterparts are not identical, they can differ in their properties.

We can formalise these observations by letting $\mathcal{C}^{\text{form}}$ and $\mathcal{C}^{\text{matter}}$ be the form counterpart function and the matter counterpart function, respectively. Now, as Lewis once noted, multiple counterparts per relation lead to bothersome distractions in cases where you might be twins; 'Goliath might not have been Lumpl' comes out trivially true in such cases, since there are two counterparts under the same relation of the object that are not identical (Lewis, 1971, p. 206). For present purposes, it will simplify discussion if we can ignore the distraction and assume as before that counterpart functions map individuals to their unique counterpart at a world, if there is such an individual, rather than to the set of their counterparts at a world. I shall proceed under this assumption, since it is clear enough how to generalise discussion to cases with multiple counterparts per relation; simply switch to the counterpart functions $\mathbb{C}^{\text{form}}$ and $\mathbb{C}^{\text{matter}}$ and adapt the lexical entries in the obvious ways. We can always appeal to contextual restrictions on accessible worlds to allow us to properly ignore worlds in which my counterparts are twins.

First, we need to specify appropriate semantic clauses for names. The rough idea is that the sense of a term selects a counterpart relation that is used to find the relevant counterparts across modal space. For example, the meanings of 'Goliath' and 'Lumpl' are as follows:

(55) a. $[\![\text{Goliath}]\!] = \lambda w.\mathcal{C}_w^{form}(\text{Goliath})$

b. $[\![\text{Lumpl}]\!] = \lambda w.\mathcal{C}_w^{\mathit{matter}}(\text{Lumpl})$

To get a sense of how these denotations work, suppose that the meanings of 'Goliath' and 'Lumpl' agree at the actual world, as well as some, but not all, merely possible worlds, as demonstrated in Table 1.

	Worlds			
Individual concepts	@	w_1	w_2	...
$\lambda w.\mathcal{C}_w^{\mathit{form}}(\text{Goliath})$	a	b	c	...
$\lambda w.\mathcal{C}_w^{\mathit{matter}}(\text{Lumpl})$	a	b	d	...

Table 1: Multiple counterpart-theoretic individual concepts

In actuality, the object that is Goliath is the object that is Lumpl, and in worlds like w_1 that object has form counterparts that are matter counterparts, but not necessarily, since in worlds like w_2 that object has a form counterpart that is not its matter counterpart.

Next, let us assume that the ' "*is*" of identity' has the following denotation:

(56) $[\![\text{is}]\!] = \lambda u_{\langle s,e\rangle}.\lambda v_{\langle s,e\rangle}.\lambda w.u(w) = v(w)$

This correctly predicts the truth of 'Goliath is Lumpl' at the actual world:

(57) $[\![\text{Goliath is Lumpl}]\!]^{@} = 1$ iff $\mathcal{C}_{@}^{\mathit{form}}(\text{Goliath}) = \mathcal{C}_{@}^{\mathit{matter}}(\text{Lumpl})$

Finally, assuming the obvious denotation for the predicate 'survives being squashed', we can make the following predications about the relevant *de re* possibility claims:

(58) $[\![\text{Goliath might not have been Lumpl}]\!]^{w} = 1$ iff there is some world w' accessible from w such that $\mathcal{C}_{w'}^{\mathit{form}}(\text{Goliath}) \neq \mathcal{C}_{w'}^{\mathit{matter}}(\text{Lumpl})$

(59) $[\![\text{Lumpl might have survived being squashed}]\!]^{w} = 1$ iff there is some world w' accessible from w such that $\mathcal{C}_{w'}^{\mathit{matter}}(\text{Lumpl})$ survives being squashed.

(60) $[\![\text{It is not the case that Goliath might have survived being squashed}]\!]^{w} = 1$ iff for all worlds w' accessible from w, $\mathcal{C}_{w'}^{\mathit{form}}(\text{Goliath})$ doesn't survive being squashed.

More needs to be said about how the sense of a term selects a counterpart relation that is used to find the relevant counterparts across modal space. But, for now, I hope to have shown how multiple counterpart relations can be incorporated in the present theory.

5.4 Reference and rigidity

Are proper names as counterpart-theoretic individual concepts referential? Much depends on the particulars of the counterpart relation. On the one hand, proper names might directly refer to their referents in much the same way as variables directly refer to their referents. So understood, 'Kim' is a directly referential term that picks out whomever is the unique counterpart of Kim at their world. On the other hand, the theory is compatible with a descriptivist view that takes the counterpart function to be descriptive content that mediates the way that names pick out counterparts at worlds.

Are proper names as counterpart-theoretic individual concepts rigid? A term is rigid only if it picks out the same individual is every world that individual exists, and nothing in any world that individual does not exist. On the present theory, proper names refer to different individuals at different worlds, and so they are not rigid. But, in a counterpart-theoretic setting, this notion of rigidity has little appeal. Better to use an alternative conception of rigidity – call it 'quasi-rigidity' – that is closer to the original intention in a counterpart-theoretic setting.

QUASI-RIGIDITY A term is *quasi-rigid* if and only if it "names at another world the counterpart there of what it names here" (Lewis, 1986, p. 256)

It is easy to see that the lexical entry for the proper name 'Kim' in (38) is quasi-rigid. For while it refers to different individuals in different worlds, it actually refers to Kim and otherwise refers to her counterparts in other worlds, if any exist.[26]

5.5 Relations to Quantified Modal Logic

Let us consider whether the version of counterpart theory developed here can attain any of the desirable features that have motivated the development of other forms of counterpart theory. There are many modal schemas that are considered desirable properties of many readings of 'necessarily' and 'possible'. For example, the following schemas are elementary axioms of many normal modal logics:

B. $\phi \to \Box\Diamond\phi$

4. $\Box\phi \to \Box\Box\phi$

5. $\Diamond\phi \to \Box\Diamond\phi$

In systems of modal logic, these axioms are guaranteed just in case the accessibility relation over worlds is symmetric, transitive, and Euclidean, respectively.[27] My version of counterpart theory separates the semantic work done by modal operators and counterpart theory, retaining the usual possible worlds semantics for the former, and requiring a theory of reference to take care of the latter. For the propositional fragment of my language, the usual constraints on the accessibility relation over worlds do all the work in ensuring that the usual modal axioms are satisfied.

But as soon as we consider a fragment involving proper names, we get potential failures of such axioms. For a countermodel to the 4-axiom schema, consider a set of three worlds $W = \{w_1, w_2, w_3\}$, each with their own disjoint domains; an accessibility relation $R =$

[26] Actually, the notion of quasi-rigidity is more complicated than the above suggests. Since the counterpart relation is context-dependent, a name may refer to different counterparts in different contexts. Consequently, 'quasi-rigidity' must be suitably relativised: names refer to the counterparts-relative-to-context-c of the same individual in every possible world. For further discussion on 'quasi-rigidity', see Lewis (1986, p. 256ff).

[27] A relation R is symmetric iff for every x, y, if xRy, then yRx; a relation R is transitive iff for every x, y, z, if xRy and yRz, then xRz; a relation R is Euclidean iff for every x, y, z, if xRy and xRz, then yRz.

$\{\langle w_1, w_2\rangle, \langle w_2, w_3\rangle, \langle w_1, w_3\rangle\}$ over W; and let a be an individual in w_1, such that (i) a is F, (ii) a's counterpart in w_2 is b who is also F (i.e., $\mathcal{C}_{w_2}(a) = b$ and b is F), (iii) b's counterpart in w_3 is c, who is not F (i.e., $\mathcal{C}_{w_3}(b) = c$ and c is not F), and (iv) a doesn't have a counterpart in w_3 (i.e., $\mathcal{C}_{w_3}(a) = \#$). Then the following instance of the 4-axiom schema fails: $\Box Fa \rightarrow \Box\Box Fa$. Consequently, suitable constraints on just the accessibility relation is not enough to capture certain modal principles.

How can the present framework be adapted to capture these truths? Whether my version of counterpart theory validates certain modal axioms seems to depend only on what general constraints should be imposed on the accessibility relation over worlds and the counterpart function. The question of which constraints could be imposed on the counterpart function to rescue certain modal axioms has been explored in great depth by Dorr (2010), and much of what he says can be adapted to the present setting. In Table 2, I list a number of constraint on counterpart functions that generate some common modal axioms.

Name	**Axiom Schema**	**Constraint**
K	$\Box(\phi \rightarrow \psi) \rightarrow (\Box\phi \rightarrow \Box\psi)$	$\forall w \forall x \exists y \mathcal{C}_w(x) = y$
D	$\Box\phi \rightarrow \Diamond\phi$	$\forall w \forall x \exists y \mathcal{C}_w(x) = y$
T	$\Box\phi \rightarrow \phi$	$\forall w \forall x (Ixw \rightarrow \mathcal{C}_w(x) = x)$
B	$\phi \rightarrow \Box\Diamond\phi$	$\forall w \forall x \forall y (\mathcal{C}_w(x) = y \rightarrow \mathcal{C}_w^{-1}(y) = x)$
4	$\Box\phi \rightarrow \Box\Box\phi$	$\forall w \forall w' \forall x \forall y \forall z (\mathcal{C}_w(x) = y \wedge \mathcal{C}_{w'}(y) = z \rightarrow \mathcal{C}_{w'}(x) = z)$
5	$\Diamond\phi \rightarrow \Box\Diamond\phi$	$\forall x \forall y \forall z \forall w \forall w' (\mathcal{C}_w(y) = x \wedge \mathcal{C}_{w'}(y) = z \rightarrow \mathcal{C}_{w'}(x) = z)$

Table 2: Some modal axioms and constraints on counterpart functions that generate them (after Dorr, 2010)

Lastly, let us consider how the present theory should treat variables, pronouns, and traces. It is natural to assume a close connection between our treatments of names and variables, one that mirrors the connection between referentialist theories of proper names and standard treatments of variables.[28]

I treat pronouns, traces, and variables as denoting counterpart-theoretic individual concepts relative to variable assignments, that is, functions from worlds to counterparts of the individual assigned to the index by the variable assignment at the context. Consider the following lexical entry for the pronoun 'she$_3$':

(61) $[\![\text{she}_3]\!]^g = \lambda w.\iota x\ x = \mathcal{C}_w(g(3))$

This clause is just like our semantic analysis of proper names, except that it allows the individual concept to vary with the variable assignment. More generally, I adopt the following revised version of the Traces and Pronouns Rule to handle traces, pronouns, and variables:

[28] Indeed, such treatment is required if there is any hope of the resulting theory being empirically adequate; see Williamson (2013, Chapter 2.4).

(62) *Traces and Pronouns Rule (revised)*
If α is a pronoun, a trace, or a variable, g is a variable assignment, and $i \in \text{dom}(g)$, then $[\![\alpha_i]\!]^g = \lambda w.\iota x\ x = \mathcal{C}_w(g(i))$.

To get a sense of how this idea works, consider what this theory predicts for the truth-conditions of unembedded and embedded open formulae like 'she$_3$ smokes' and 'she$_3$ might smoke':

(63) $[\![\text{she}_3 \text{ smokes}]\!]^{w,g} = 1$ iff $\iota x\ x = \mathcal{C}_w(g(3))$ smokes.

(64) $[\![\text{she}_3 \text{ might smoke}]\!]^{w,g} = 1$ iff there exists a world w' accessible from w and
$\iota x\ x = \mathcal{C}_{w'}(g(3))$ smokes

In English, (63) says that 'she$_3$ smokes' is true at a world w and variable assignment g iff the counterpart at w of what g assigns to the index 3 smokes; (64) says that 'she$_3$ might smoke' is true at a world w and variable assignment g iff there is a world w' accessible from w such that the counterpart at w' of what g assigns to index 3 smokes. Both of these truth-conditions seem intuitively adequate.

This treatment also adequately accounts for the interactions between quantifiers and modals. To see this, let us first consider how the present theory accounts for variations in quantifier scope, such as inverse scope readings of 'Everyone loves someone' (i.e., someone is such that everyone loves them). A standard way of account for inverse scope readings is to suppose that the existential quantifier expression *someone* raises out of its initial position to the top of the sentence via a covert movement called *Quantifier Raising* (May, 1977, 1985). This leaves behind a trace that is then abstracted over by introducing a lambda operator immediately after the quantifier expression. So, the resulting syntactic structure is as follows:

(65) Everyone loves someone.

a. [someone [λ_3 [everyone loves t_3]]]

Let us assume that *everyone*, *someone*, and *loves* have the following denotations:

(66) a. $[\![\text{everyone}]\!]^g = \lambda f_{\langle\langle s,e\rangle,\langle s,t\rangle\rangle}.\lambda w.1$ iff for every $u(w) \in D_w$, $f(u(w))(w) = 1$

b. $[\![\text{someone}]\!]^g = \lambda f_{\langle\langle s,e\rangle,\langle s,t\rangle\rangle}.\lambda w.1$ iff there is a $u(w) \in D_w$ such that $f(u(w))(w) = 1$

c. $[\![\text{loves}]\!]^g = \lambda v_{\langle s,e\rangle}.\lambda u_{\langle s,e\rangle}.\lambda w.1$ iff $u(w)$ loves $v(w)$

We also need to modify *Predicate Abstraction* to take into account that the fact that the arguments of transitive and intransitive predicate denotations are now individual concepts:

(67) *Predicate Abstraction (second revision)*
For all indices i and assignments g, $[\![\lambda_i\ \alpha]\!]^g = \lambda u_{\langle s,e\rangle}.\lambda w'.[\![\alpha]\!]^{g[u(w')/i]}(w')$.

Putting this together, we can see that (65) has the following intuitively correct truth-conditions:

(68) $[\![(65)]\!]^{w,g} = 1$ iff there is a $v(w) \in D_w$ such that for every $u(w) \in D_w$,
$u(w)$ loves $\iota y\ y = \mathcal{C}_w(v(w))$

We are finally in a position to see that the present theory delivers intuitively correct results for interactions between quantifier expressions and modal operators. Consider, for example, the sentence ‘Someone might be loved by everyone’, which we assume to have the following syntactic structure:

(69) Someone might be loved by everyone

a. [someone [λ_3 [might [everyone loves t_3]]]]

Then we can easily see that (69) has the following truth-conditions, again building them up step-by-step:

(70) $[\![(69)]\!]^{w,g} = 1$ iff there is a $v(w) \in D_w$ such that there is a world w' accessible from w such that, for every $u(w') \in D_{w'}$, $u(w')$ loves $\iota y\ y = \mathcal{C}_{w'}(v(w))$

These truth-conditions are intuitively adequate. It is easy to verify that the semantics under consideration adequately account for the truth-conditions of other scopal possibilities.

6 Related views

I want now to compare my theory with a similar theory: *variabilsm about proper names*.[29] A number of theorists have argued that names and pronouns are terms that undergo a shifted interpretation under the scope of epistemic modals and conditionals (Cumming, 2008; Mackay, 2023; Ninan, 2012, 2013, 2018; Rabern, 2021; Santorio, 2012; Schoubye, 2020). These views all fall within the family of counterpart semantics, so it is worth comparing them to my theory explicitly.

A central motivation for variabilism is shifted uses of names. Consider the following example from Schoubye (2020). Banksy is a psuedononyous England-based street artist whose real name and identity are unconfirmed and subject to much speculation. Del Naja is an artist, musician, and founding member of Massive Attack who has been subject to speculation that he is Banksy. Schoubye is concerned to provide a semantic view where simple epistemic modal sentences such as ‘Del Naja might be Banksy’ are intuitively true, even if, unbeknownst to us, Del Naja is not in fact Banksy. After all, Banksy’s identity is currently unknown and we may have positive reasons for thinking that Del Naja is responsible for Banksy’s work. But standard theories of proper names like Millianism, the view according to which the meaning of a name is exhausted by its reference and a name contributes only its reference to the semantic content of sentences in which it is contained, cannot accommodate this datum, which is contrary to standard judgements.

[29] Thanks to an anonymous reviewer for encouraging me to discuss the differences between my theory and variabilism. Another related family of views worth mentioning explicitly are the Contingent Identity Systems, like those developed by Carnap (1947), Gibbard (1975), Aloni (2001, 2005), and Hughes and Cresswell (1996, pp. 330–347). Contingent Identity Systems are based on the framework of modal predicate logic, except that variables are taken to range over *total* individual concepts, that is, total functions from possible worlds to possible individuals. For Contingent Identity Systems to be applicable to the two puzzles central to this paper, we would first need to generalise this treatment of variables to proper names. Even then, because of the insistence that the relevant individual concepts are total, Contingent Identity Systems are not able to capture the full range of motivations central to counterpart theory, such as the considerations discussed at length in Sections 5.2 and 5.3.

To account for shifted uses of names, variabilists first analyse proper names as assignment-dependent singular terms akin to variables. For example, on one version of variabilism, 'Del Naja' has the following lexical entry:

(71) $[\![\text{Del Naja}_i]\!]^{c,g,w} = \begin{cases} g(i) & \text{if } g(i) \text{ is called Del Naja in } w_c \\ \text{undefined} & \text{otherwise} \end{cases}$

In English, the semantic value of 'Del Naja' is a partial function that takes a variable assignment as input and returns the individual that is determined by the assignment and the name's numerical index, if that individual has the property of being called Del Naja, and is undefined otherwise.

Next, variabilists argue that epistemic modals are operators that shift not only the world parameter, but also shift the assignment parameter. More specifically, epistemic modals both quantifier over a set of worlds compatible with the epistemic state of the agents in question and align the representation of the relevant referential relations with the information presupposed by the agents to reflect how they are thinking of certain objects and individuals. Here's one way of making this idea precise. Let the semantics for epistemic 'might' be as follows:

(72) $[\![\text{might } \phi]\!]^{c,g,w} = \lambda p_{\langle s,t \rangle}.1$ iff there is a world–assignment pair $\langle g', w' \rangle$ accessible from $\langle g, w \rangle$, $[\![\phi]\!]^{c,g',w'} = 1$.

What makes one world–assignment pair accessible from another? In addition to determining an assignment function g from variables $x_1, \ldots, x_n$ to individuals, the context also determines a sequence of counterpart functions $F = \langle f_1, \ldots, f_n \rangle$, one for each variable $x_1, \ldots, x_n$. Formally, a counterpart function f_i is a function from worlds to the counterpart of $g(x_i)$ at that world. Then we can say that, where the domain of g is includes $x_1, \ldots, x_n$, a world–assignment pair $\langle w', g' \rangle$ is accessible from $\langle w, g \rangle$ iff:

(73) (i) w' is consistent with the speaker's information at w;
(ii) $f_1(w) = g(1)$ and $f_2(w) = g(2)$ and … and $f_n(w) = g(n)$;
(iii) $g' = \{\langle 1, f_1(w') \rangle, \langle 2, f_2(w') \rangle, \ldots, \langle n, f_n(w') \rangle\}$.

These two ideas allow us to accommodate the intuition that the sentence 'Del Naja might be Banksy' is true even if, unbeknownst to us, Del Naja isn't actually Banksy.

(74) $[\![\text{Del Naja}_i \text{ might be Banksy}_j]\!]^{c,g,w} = 1$ iff there is a world–assignment pair $\langle g', w' \rangle$ accessible from $\langle g, w \rangle$ such that $[\![\text{Del Naja}_i \text{ might be Banksy}_j]\!]^{c,g',w'} = 1$ (i.e., iff there is a world–assignment pair $\langle g', w' \rangle$ accessible from $\langle g, w \rangle$ such that $g'(i) = f_{\text{Del Naja}}(w') = f_{\text{Banksy}}(w') = g'(j)$).

So even if Del Naja is not Banksy (i.e., $g(i) = f_{\text{Del Naja}}(@) \neq f_{\text{Banksy}}(@) = g(j)$), Del Naja might be Banksy.

There are other ways to develop variabilism, but what is common to the vast majority of them is that they introduce counterparts in the semantics of epistemic modals, and not in the semantics of proper names or metaphysical modals.[30] In contrast, I take the counterpart relation to be baked into the semantics of proper names, rather than being introduced by special kinds of modal operators. This difference means that the two theories have remarkably different applications. For one thing, variabilists who are only tempted to make sense of shifted readings of proper names under the scope of *epistemic* (and perhaps doxastive) modal operators, are unable to make sense of a standard motivation for counterpart theory, namely, that split identity sentences like 'I might have been Marilyn Monroe' have a true reading where the modality is interpreted metaphysically. That said, variabilitsts are free to adopt a shifty semantics for metaphysical readings of modal operators as well, should they be so inclined to accommodate metaphysically shifted readings.

More seriously, extant variabilist semantics for proper names are unable to give an intuitively adequate account of the truth-conditions for non-modal declaratives like 'Del Naja$_i$ smokes', at least against a background counterpart-theoretic metaphysics. For example, according to the version of variabilism above, an utterance of the sentence 'Del Naja$_i$ smokes', at a context c, world w, and variable assignment g, expresses the proposition that $g(i)$ smokes, if $g(i)$ is called Del Naja in w_c. But if $g(i)$ is world-bound, variabilism runs into trouble. There are two ways to handle the presuppositional constraint that $g(i)$ must be called Del Naja in the world of the context of utterance. Either $[\![$'Del Naja$_i]\!]^{c,g,w}$ is undefined at any non-actual context as $g(i)$ would not be called anything there; then 'Del Naja$_i$ smokes' only expresses a proposition at the actual world; Or if $g(i)$ is called Del Naja in actuality, he may be called Del Naja everywhere; then 'Del Naja$_i$ smokes' expresses a singleton proposition or an empty proposition depending on whether he actually smokes. It's hard to say which option is correct. But either way, the resulting content does not capture our intuitions about what proposition 'Del Naja$_i$ smokes' expresses. In other words, most extant variabilists don't have a solution to my first puzzle, though in fairness they don't need one; they're not real counterpart theorists.

Variabilism about proper names and my version of counterpart theory can be unified, however, by taking proper names to be assignment-dependent, counterpart-dependent singular terms. On the simplest way to combine these ideas, 'Del Naja' has the following lexical entry:

(75) $[\![\text{Del Naja}_i]\!]^{c,g} = \lambda w.\iota x\ x = \mathcal{C}_w(g(i))$

In effect, this account generalises my treatment of pronouns, traces, and variables to proper names. For considerations of space, I shall not compare the non-variabilist counterpart-theoretic treatment of names with the variabilist treatment, but many of the same considerations that motivate standard variabilism may motivate counterpart theorists to adopt analogous versions of my theory.

7 Concluding remarks

This paper submits to scrutiny the implications of adopting counterpart theory as an ontological foundation for natural language semantics. It does so by introducing two puzzles about how

[30] An important exception here is Mackay (2023), who is a variabilist in the sense that he takes terms that to undergo epistemic shift to be sensitive to assignments, but, unlike other variabilist theories, argues that "all modals and conditionals, not just epistemic ones, are formally assignment shifters" (Mackay, 2023, p. 199). Once this point is taken into account, it is not obvious that Mackay's view is really all that different from the view I describe at the end of the present section. Thanks to an anonymous reviewer for drawing my attention to the nuances of Mackay's view

counterpart theory handles proper names. The first puzzle involves non-modal contexts and poses a general problem for a theory of content. The second puzzle involves proper names in intensional contexts and threatens counterpart-theoretic treatments of modality. I then evaluate several proposals for how counterpart theorists can resolve these puzzles.

First, I argued that counterpart theorists should not follow Lewis and other theorists in adopting translation clauses or syncategramatical rules to analyse modal terms, as this would lead to a proliferation of lexical entries beyond necessity. Instead, counterpart theorists should separate the role that the counterpart relation plays from the role that modal expressions play: modal expressions involve quantification over (contextually-accessible) possible worlds, while the counterpart relation is introduced through other means. There are different ways of doing this. Some theorists build the counterpart relation into noun and verb denotations, but I argue that doing so loses a Lewisian commitment to intrinsic properties. Instead, I argue that the counterpart relation should be introduced into the semantics of proper names. This is broadly in a Lewisian spirit and leads to a smooth resolution of our puzzles. However, it is surprising that adopting counterpart theory forces one's hand when it comes to the semantics of proper names. The ontological framework in which we carry out semantic analyses should be general enough to compare the empirical adequacy of different lexical entries for certain expressions. Furthermore, adopting my favoured resolution to the puzzle leads one to substantial commitments in one's counterpart theory, depending on how you implement it.[31]

References

Aloni, M. (2001). *Quantification under Conceptual Covers* (Unpublished doctoral dissertation). University of Amsterdam, Amsterdam.

Aloni, M. (2005). Individual Concepts in Modal Predicate Logic. *Journal of Philosophical Logic*, *34*(1), 1–64. doi: 10.1007/s10992-004-4065-8

Anand, P. (2007). Dream Reports Pronouns, Local Binding, and Attitudes *De Se*. In T. Friedman & M. Gibson (Eds.), *SALT XVII* (pp. 1–18). Ithaca, NY.: Cornell University.

Bader, R. M. (2016). Contingent identity and counterpart theory. *Philosophical Perspectives*, *30*(1), 7–20. doi: 10.1111/phpe.12073

Carnap, R. (1947). *Meaning and Necessity*. Chicago, IL.: University of Chicago Press.

Cumming, S. (2008). Variabilism. *Philosophical Review*, *117*(4), 525–554. doi: 10.1215/00318108-2008-015

Dorr, C. (2010). *How to be a modal realist* [Manuscript].

Elbourne, P. (2013). *Definite Descriptions*. Oxford: Oxford University Press.

Fara, D. G. (2008). Relative-sameness counterpart theory. *The Review of Symbolic Logic*, *1*(2), 167–189. doi: 10.1017/S1755020308080155

Fara, M., & Williamson, T. (2005). Counterparts and Actuality. *Mind*, *114*(453), 1–30. doi: 10.1093/mind/fzi001

[31] I should like to thank Sam Carter, Paul Elbourne, Katherine Hong, Annina Loets, Alexander Roberts, Timothy Williamson, the audience members of a meeting of the Metaphysics & Epistemology Group at the University of Oxford, and especially Nicholas K. Jones and an anonymous reviewer for this journal for helpful comments and discussion on previous drafts of this paper.

Forbes, G. (1982). Canonical Counterpart Theory. *Analysis*, *42*, 33–37.

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

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

Forbes, G. (1990). Counterparts, Logic, and Metaphysics: Reply to Ramachandran. *Analysis*, *50*, 167–173.

Gibbard, A. (1975). Contingent Identity. *Journal of Philosophical Logic*, *4*(2), 187–221.

Hawthorne, J., & Yli-Vakkuri, J. (2023). Counterpart Theory and Counterfactuals. In K. Bennett & D. W. Zimmerman (Eds.), *Oxford Studies in Metaphysics* (Vol. 13). Oxford University Press.

Hazen, A. (1979). Counterpart-Theoretic Semantics for Modal Logic. *The Journal of Philosophy*, *76*(6), 319–338. doi: 10.2307/2025472

Hazen, A. (2012). Reflections on counterpart theory. In K. Bennett & D. W. Zimmerman (Eds.), *Oxford Studies in Metaphysics* (Vol. 7). Oxford: Oxford University Press.

Heim, I. (1994). *Puzzling reflexive pronouns in* de se *reports.* Bielefeld.

Heim, I. (2001). *Semantics and morphology of person and logophoricity* [Handout of a Talk given at University of Tübingen]. Tübingen.

Heim, I., & Kratzer, A. (1998). *Semantics in Generative Grammar*. Oxford: Blackwell Publishers.

Hughes, G. E., & Cresswell, M. J. (1996). *A New Introduction to Modal Logic*. London: Routledge.

Jackendoff, R. (1975). On belief-contexts. *Linguistic Inquiry*, *6*, 53–93.

Kaplan, D. (1989). Demonstratives. In J. Almog, J. Perry, & H. Wettstein (Eds.), *Themes from Kaplan* (pp. 481–563). Oxford: Oxford University Press.

Kratzer, A. (1979). Conditional necessity and possibility. In R. Bäuerle, U. Egli, & A. von Stechow (Eds.), *Semantics From Different Points of View* (pp. 117–147). Berlin: Springer.

Kratzer, A. (1986). Conditionals. In A. Farley, P. Farley, & K. E. McCollough (Eds.), *Proceedings of the Chicago Linguistics Society* (Vol. 22, pp. 115–135). Chicago, IL.: Chicago Linguistics Society.

Kratzer, A. (1989). An investigation of the lumps of thought. *Linguistics and Philosophy*, *12*(5), 607–653.

Kratzer, A. (2012). *Modals and Conditionals*. Oxford: Oxford University Press.

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

Lakoff, G. (1968). *Counterparts, or the problem of reference in transformational grammar* [Manuscript]. Harvard University.

Lewis, D. (1968). Counterpart theory and quantified modal logic. *The Journal of Philosophy*, *65*(5), 113–126. doi: 10.2307/2024555

Lewis, D. (1970). General semantics. *Synthese*, *22*, 18–67.

Lewis, D. (1971). Counterparts of Persons and Their Bodies. *The Journal of Philosophy*, *68*(7), 203–211. doi: 10.2307/2024902

Lewis, D. (1983). Extrinsic Properties. *Philosophical Studies: An International Journal for*

Philosophy in the Analytic Tradition, *44*(2), 197–200.

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

Link, G. (1983). The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze, & A. von Stechow (Eds.), *Meaning, Use, and Interpretation of Language* (pp. 303–323). Berlin: Walter De Gruyter & Co.

Mackay, J. (2023). Counterfactual epistemic scenarios. *Noûs*, *57*(1), 188–208. doi: 10.1111/nous.12403

May, R. (1977). *The Grammar of Quantification* (Unpublished doctoral dissertation). MIT, Cambridge, MA.

May, R. (1985). *Logical Form: Its Structure and Derivation*. Cambridge, MA.: The MIT Press.

Mill, J. S. (1843). *A System of Logic, Ratiocinative and Inductive*. London: John W. Parker.

Montague, R. (1970). English as a Formal Language. In B. Visentini (Ed.), *Linguaggi Nella Societa e Nella Tecnica* (pp. 188–221). Edizioni di Communita.

Ninan, D. (2012). Counterfactual attitudes and multi-centered worlds. *Semantics and Pragmatics*, *5*, 1–57. doi: 10.3765/sp.5.5

Ninan, D. (2013). Self-Location and Other-Location. *Philosophy and Phenomenological Research*, *87*(2), 301–331. doi: 10.1111/phpr.12051

Ninan, D. (2018). Quantification and Epistemic Modality. *The Philosophical Review*, *127*(4), 433–485. doi: 10.1215/00318108-6973010

Percus, O. (1998). Some instructions for the wordly. In K. Shahin, S. Blake, & E.-S. Kim (Eds.), *Proceedings of WCCFL 17.* Stanford: CSLI Publications.

Percus, O., & Sauerland, U. (2003). Pronoun Movement in Dream Reports. In *Proceedings of NELS 33.* University of Massachusetts, Amherst.

Quine, W. V. O. (1953). *From a Logical Point of View*. Cambridge, MA.: Harvard University Press.

Rabern, B. (2013). Monsters in Kaplan's logic of demonstratives. *Philosophical Studies*, *164*, 393–404.

Rabern, B. (2021). Binding bound variables in epistemic contexts. *Inquiry*, *64*(5–6), 533–563.

Ramachandran, M. (1989). An alternative translation scheme for counterpart theory. *Analysis*, *49*(3), 131–141. doi: 10.1093/analys/49.3.131

Ramachandran, M. (1990a). Contingent Identity in Counterparty Theory. *Analysis*, *50*(3), 163–167.

Ramachandran, M. (1990b). Unsuccessful revisions of CCT. *Analysis*, *50*(3), 173–177. doi: 10.1093/analys/50.3.173

Russell, J. S. (2013). Actuality for Counterpart Theorists. *Mind*, *122*(485), 85–134. doi: 10.1093/mind/fzt037

Santorio, P. (2012). Reference and Monstrosity. *Philosophical Review*, *121*(3), 359–406. doi: 10.1215/00318108-1574427

Sauerland, U. (2014). Counterparts block some 'de re' readings. In L. Crnič & U. Sauerland (Eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim, Volume 2* (pp.

65–85).

Schoubye, A. J. (2020). Names Are Variables. *The Philosophical Review*, *129*(1), 53–94. doi: 10.1215/00318108-7890468

Sider, T. (1996). All the world's a stage. *Australasian Journal of Philosophy*, *74*(3), 433–453. doi: 10.1080/00048409612347421

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

Stalnaker, R. (1986). Counterparts and identity. *Midwest Studies In Philosophy*, *11*, 121–140.

Szabó, Z. G. (2022). Compositionality. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (Fall 2022 ed.). Metaphysics Research Lab, Stanford University.

van Rooij, R. (1997). *Attitudes and Changing Contexts* (Unpublished doctoral dissertation). University of Stuttgart, Stuttgart.

Varzi, A. C. (2020). Counterpart theories for everyone. *Synthese*, *197*(11), 4691–4715. doi: 10.1007/s11229-020-02720-1

von Fintel, K., & Heim, I. (2011). *Intensional semantics* [Manuscript]. MIT.

Wang, J. (2015). Actualist Counterpart Theory. *The Journal of Philosophy*, *112*(8), 417–441. doi: 10.5840/jphil2015112826

Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford: Oxford University Press.