Intentional Identity

William Lanier

University College

Submitted for the degree of D.Phil.

Trinity Term 2013

Intentional Identity

William Lanier, University College

Submitted for the degree of D.Phil.

Trinity Term 2013

Abstract

If Hob and Nob both read the same newspaper article claiming that a witch has come into town, then the following sentence can be true, even if the article is fabricated and there are no witches:

(1) Hob thinks that a witch has blighted Bob's mare, and Nob wonders whether she killed Cob's sow.

This phenomenon is called 'intentional identity', and there is no consensus on the semantics of (1) or similar sentences. Intentional identity is related to important, unsettled topics in the philosophy of language (e.g., anaphora, dynamic semantics) and in metaphysics (e.g., fictional and Meinongian objects). Thus, a correct semantic account of intentional identity is desirable. In this thesis, I argue that 'she' in (1) is behaving semantically like a traditional definite description, and that the truth of sentences like (1) often requires a certain causal connection between the two subjects.

In chapter 1, I explain the difficulty in finding a correct semantic account of intentional identity sentences, and I present new evidence that the phenomenon is broader than previously thought. Chapter 2 explores the idea that (1) involves certain exotic objects—e.g., fictional, Meinongian, or merely possible witches. I show that what I call the 'causal connection problem' affects most versions of this idea, and that even the best version is probably incorrect. In chapter 3, I argue that 'she' in (1) is not being bound

dynamically, and that the 'guise theory' approach suggested by several dynamic semanticists is unhelpful. Chapter 4 contains my proposed solution. With a broader view of the problem, one can see that 'she' is functioning like a traditional incomplete definite description, and that its complete semantic value involves Hob and Nob being causally connected. This solution allows us to avoid an extravagant semantics and ontology.

Contents

Acknowledgements

My principal debt is to my primary supervisor, Cian Dorr. This thesis would certainly not exist without him, for he first proposed the subject, continually suggested ideas and improvements through constant meetings, and discussed countless drafts with me. I am deeply grateful for the thoughtfulness and diligence he showed me, and I am constantly inspired by his example of clear philosophical thinking.

I am also indebted to my secondary supervisor, Daniel Rothschild, and to my earlier teachers: Harold Hodes, Timothy Williamson, Benjamin Morison, and John Hawthorne.

For comments on earlier written and presented versions of this material, thanks to an anonymous confirmation of status reviewer, anonymous journal reviewers, and an Oxford audience. I especially owe Michael Gibb for reading and commenting on the whole manuscript. Thanks to fellow Oxford graduate students for philosophical discussions.

I am very grateful to my examiners Paul Elbourne and John Hawthorne for spending their time reading and thinking about my thesis, and I look forward to incorporating their thoughts into a future, improved version of this material.

For my personal life during the D.Phil., I owe too much to be expressed to Mom, Dad, Becky, my sisters, my grandmothers and grandfathers, my whole family, Dawn Rennie, and my close friends. Thank you Nora Danielson for the last poms.

Chapter 1

Intentional Identity

1.1 The problem

> *Geach's scenario*: The local newspaper has printed an article saying that there is a witch that has come into town. The article is completely made-up, and in fact there are no witches. Hob credulously reads the article, however, and says, 'I think this witch blighted Bob's mare'. Across town, Nob also reads the same article and says, 'Perhaps this witch killed Cob's sow'.

In Geach's scenario, there is a true reading of this sentence:

(G) Hob thinks a witch has blighted Bob's mare, and Nob wonders whether she killed Cob's sow.

Geach (1967) called the phenomenon exhibited in this sentence 'intentional identity'—there is no witch, though in some sense Hob and Nob are 'thinking about the same thing'—and Geach could not find a correct semantic account of such sentences.

Traditionally, singular third-person pronouns have been thought to function either (i) like free variables that pick out a particular entity salient in the context of utterance, or (ii) like bound variables of first-order logic (Heim and Kratzer 1998, c. 9):

(2) He's tall! [Pointing at Obama]

(3) Every boy dreams about the first person that kisses him.

At the context in which (2) is uttered, 'he' picks out a particular person, Obama, and the whole sentence says that Obama is tall. In (3), rather than picking out a particular person, 'him' functions like a variable bound by the quantifier phrase 'Every boy', and the whole sentence says that every boy x is such that x dreams about the first person that kisses x.

In (G) however, 'she' does not seem to function in either of these two ways. The pronoun does not appear to pick out a particular thing, for the article was made-up, and there are no witches, so it seems there is no thing x such that Nob is wondering whether x killed Cob's sow. The pronoun also does not seem to be functioning as a bound variable, for the only quantifier phrase that could plausibly be binding it is 'a witch', and binding 'she' with this quantifier seems to result in a different meaning. 'There is a witch such that: Hob thinks it has blighted Bob's mare, and Nob wonders whether it killed Cob's sow' is not equivalent to (G)—again, there is no such witch, so the latter is false though (G) is true.

Rather than posing an isolated puzzle, intentional identity sentence like (G) are closely connected to important issues. In the philosophy of language, sentences like (G) are part of a broader debate on anaphora, and this debate has caused many to make significant changes to traditional theories of meaning (King 2010).[1] In metaphysics, sentences like (G) have been claimed as

[1]This broader debate involves any anaphoric pronouns that, like in the specific case of intentional identity sentences, do not seem to be properly accounted for by the traditional semantics of pronouns. In this class we have donkey, fish, paycheck, and Bach-Peters sentences, and modal subordination (which I discuss in chapter 3).

evidence that there are certain peculiar objects, for example, non-existent witches (e.g., Parsons 1980).[2]

Though numerous semantic accounts of intentional identity sentences have been proposed since Geach, none have gained widespread acceptance. In this thesis, I present evidence that the currently proposed accounts are incorrect, and then I argue for a new account that treats 'she' semantically as a definite description. The account is conservative in that it does not require radical changes in either our ontology or our semantics.

This chapter performs preliminary work before we discuss proposed accounts. First, I describe some formal semantic tools and choices of notation. Then, these tools are used to show why the most obvious accounts of sentences like (G) are inadequate. Finally, I give an overview of less obvious accounts of (G), and I describe how the problems of intentional identity arise in a much broader class of sentences than previously thought.

1.2 Formal semantics

I assume the basic perspective of most of formal semantics, as exemplified in (Heim and Kratzer 1998). Humans have the ability to independently consider a new scenario and sentence, and give a uniform judgment about whether the sentence is true of the scenario. It seems that this is done mostly outside of our conscious awareness, with our brain computing the judgment from knowledge of the meanings of the smallest meaningful bits of the sentence, the structure in which those bits are put together, and general rules. Semantics tries to model at least one way this computation might go, by

[2] A further reason for the importance of intentional identity: it seems to involve similar issues as the debate on empty names. It is plausible that having an account of intentional identity sentences would suggest an account of sentences like 'Homer believed a god named 'Zeus' threw thunderbolts, and Hesiod believed that Zeus lived on Olympus', which would in turn suggest an account of the meaning of 'Zeus'.

assigning 'semantic values' to the smallest meaningful bits, and then listing general rules that, given meanings of smaller bits and structures in which they are combined, assign semantic values to larger bits. Different ways of modeling give different kinds of semantic values to sentences, but in every way of modeling, the semantic values of the sentences predict particular judgments about whether or not the sentences are true of various scenarios. We try to find a model whose predicted judgments accurately match actual human judgments.

I use structures similar to those in (Heim and Kratzer 1998) and in the beginnings of introductory syntax textbooks like (Carnie 2013). I do not adopt the terminology of the DP-analysis in (Abney 1987), and instead use the traditional term 'noun phrase' (NP) to identify phrases like 'most donkeys', 'Sancho's wine' and 'the sorrowful knight'. When discussing linguistic phrases, I sometimes use Quine's corner quotes '⌜' and '⌝' (Quine 1940). I follow (Heim and Kratzer 1998) in using structures that depart from some introductory syntax textbooks in two particular ways: first, I assume that all non-terminal nodes are double-branching; second, I assume that the structures given to semantics may have undergone Quantifier Raising, which I describe in the semantic details below. I occasionally use the notion of c-commanding: a constituent of a syntax tree c-commands all of its siblings and all of their descendants.

I often use the labels 'anaphor' and 'antecedent'. The rough idea is that noun phrases can be 'linked together' in a vague, pre-theoretic sense:

(4) <u>*A physicist*</u> came late to the party. *He* brought some bongos.

(5) Balbec houses <u>*some Parisians*</u>. *Most* come to escape the city.

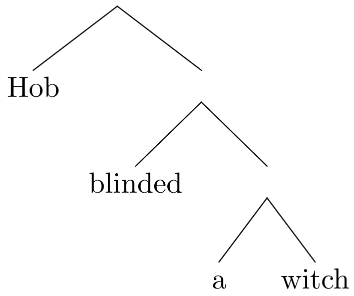
(6) <u>*Sam*</u> stood in the rain in the square where *his* father had once sold books.

(7) That *she* is ruining a thane's life gives *every witch* glee.

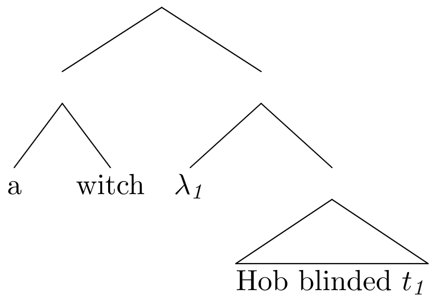
In each of these sentences, the italicized noun phrases are in some sense linked together, and I call the underlined phrase the 'antecedent' and the non-underlined phrase the 'anaphor'. I call this linking phenomena 'anaphora'. Binding theory attempts to explain this linking in more detail (see Büring 2005), but I am not relying on these concepts to do any significant work. I just use these terms so that I can more easily refer to particular phrases of sentences.

When it comes to semantic values, I use the typed lambda notation of (Heim and Kratzer 1998). Type e is the type of entities, t of truth values, s of possible worlds. I drop parenthesis according to the convention that association is from the right, so 'est' is an abbreviation of '(e(st))'. I give particular semantic values when necessary in later chapters, but here I should briefly describe Quantifier Raising (see Heim and Kratzer 1998 for more), since we will use it to see why 'she' is not being bound by 'a witch' in (G).

Quantifier Raising involves quantified phrases like 'a witch' moving from their location in the spoken structure of the sentence to a new location in the structure given to semantics. The moved phrase leaves behind a trace that functions semantically like a first-order-logic variable. A lambda binder which will bind the trace is conjoined to the sentence containing this trace, and then the moved quantified phrase is conjoined to the lambda-bound constituent. For example, the following spoken structure:



is turned into the following structure given to semantics:



We have the following semantic values (which will be more complicated when we turn to analyzing intensional phrases), relativized to variable assignments g:

$[\![$'Hob'$]\!]^g$ = Hob

$[\![$'blinded'$]\!]^g = [\lambda y_{\mathrm{e}}.\lambda x_{\mathrm{e}}.x \text{ blinded } y]$

$[\![$'a witch'$]\!]^g = [\lambda f_{\mathrm{et}}.\exists x(x \text{ is a witch and } f(x) = \text{True})]$

$[\![$'t_1'$]\!]^g = g(1)$

Rather than giving a semantic value to 'λ_1', variable binding is performed by the following rule, where $g^{x/i}$ is the variable assignment that assigns entity x to i, and that agrees with g in its assignment to everything else:

Predicate Abstraction For all variable assignments g, for all indices i, if α is $[\lambda_i\ \beta]$, then

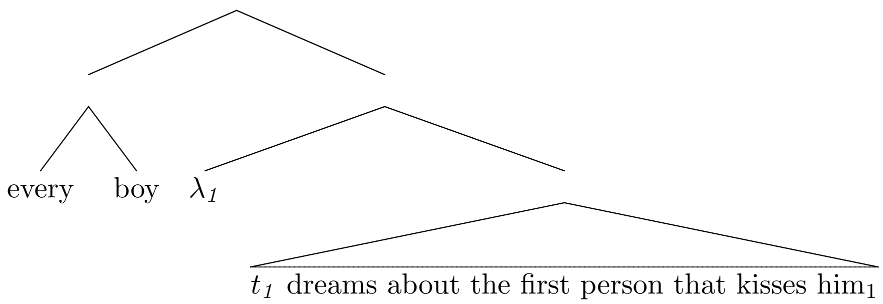
$$[\![\alpha]\!]^g = [\lambda x_{\mathrm{e}}.[\![\beta]\!]^{g^{x/i}}]$$

Our other rule is functional application:

Functional Application For all variable assignments g, if α has β and γ as immediate daughters, and $[\![\gamma]\!]^g$ is in the domain of $[\![\beta]\!]^g$ then,

$$[\![\alpha]\!]^g = [\![\beta]\!]^g([\![\gamma]\!]^g)$$

Pronouns receive the same semantic values as traces; $[\![$'she$_1$'$]\!]^g = g(1)$. If they are functioning as bound variables, Quantifier Raising will raise some quantified phrase, and will create a λ that will bind both the pronoun and the trace left behind by the raised phrase. So 'Every boy dreams about the first person that kisses him' will turn into:



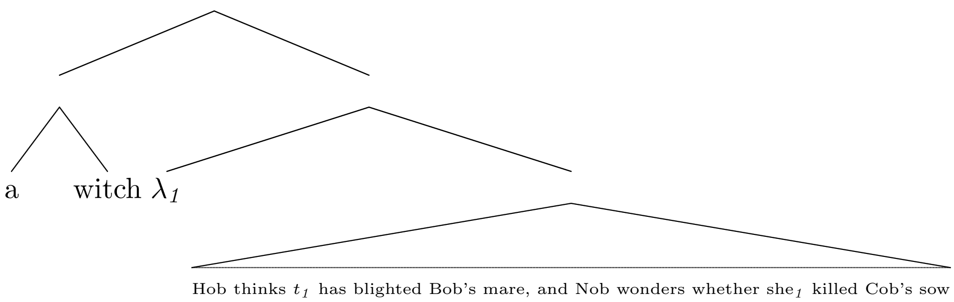
If on the other hand there is a pronoun that is functioning like a free variable, then it will remain unbound, and we model this by saying that the context of the pronoun's utterance provides a variable assignment that assigns the index of the pronoun to the appropriate object. So for 'He's tall$_1$ [Pointing at Obama]', the variable assignment provided by the context takes the index 1 to Obama.

A sentence is true at a context iff, relative to the contextually-provided variable assignment, the sentence's semantic value is True. With these definitions, values, structures, and rules, we predict the right judgments for the examples in this section, and we can now try to model the semantics of pronouns in intentional identity sentences.

1.3 Why the simple solutions fail

Now that we have our semantic tools, we can be more specific in why (G) is problematic. The problem is with 'she'. We have two ways in which pronouns function. First, the pronoun 'she' can refer to a particular object. In Geach's scenario, however, it seems that there is no witch—and no anything else—that is the subject of the newspaper article, or of Nob's wonderings. The only way to hold that 'she' functions as a free variable is to postulate some exotic, extraordinary thing to be its value, for example, a 'mythical' witch, a 'merely possible' witch, or a non-existent witch. I look at accounts that use exotic objects in chapter 2. These are the only accounts that could have the pronoun functioning as a free variable.[3]

The second way 'she' can function is by being a bound variable. There is overwhelming evidence, however, that 'she' is not functioning as a bound variable in intentional identity sentences. We might try moving 'a witch' out from its immediate clause, so that we get the following structure:



Sentences with other determiners, however, show that this is not how

[3]In this section, I am not considering accounts that change the type of the pronoun, or add significant implicit material that is processed along with the pronoun, or significantly change the semantics traditionally given to the surrounding material. With such changes, many (and probably all) of the accounts described in the rest of the thesis could have variants where the pronouns themselves function as free or bound variables.

intentional identity sentences are working. As Evans (1977) noted in connection with different sentences, this wide-scoping delivers the wrong truth conditions for sentences involving pronouns anaphoric on noun phrases with determiners like ‘exactly one’, ‘just two’, and ‘few’. Take for example:

(8) Hob thinks exactly one witch teaches logic at Oxford, and Nob thinks she has rooms at All Souls.

This sentence cannot satisfactorily be analyzed as:

(9) exactly one witch [λ_1 [Hob thinks t_1 teaches logic at Oxford and Nob thinks she$_1$ has rooms at All Souls]]

For (8) has a clear reading on which it is true only if Hob thinks no more than one witch teaches logic at Oxford, whereas (9) can be true even if Hob believes more than one witch teaches logic at Oxford. Similar problems arise with the other determiners. This objection is sufficient evidence that intentional identity pronouns are not wide-scope bound, for the correct semantic account of (8) should not be radically different from that of (G), and surely whatever semantic account explains ‘she’ in (8) will also explain ‘she’ in (G), making a wide-scoping account of (G) pointless.[4]

Furthermore, most syntacticians believe that, in general, NPs cannot bind pronouns when the pronoun is not contained in the smallest sentence containing the NP:

(10) Hob believes every witch wishes she could afford a top-of-the-line broomstick.

[4]Moreover, there is a clear reading of (G) where it is only true if Hob thinks that there are witches. Moving ‘witch’ out of the scope of ‘Hob thinks’ will mean we lose this entailment. One might propose instead to transform (G) into some structure like [A thing [λ_1 [Hob thinks t_1 is a witch and blighted Bob’s mare, and Nob wonders whether she$_1$ killed Cob’s sow]]. This looks impossible on any theory considered by contemporary syntacticians.

(11) Hob believes every witch bought a top-of-the-line broomstick, and Nob believes she went to Oxford to learn how to fly.

In (10), 'every witch' can bind the pronoun 'she'; in (11), it cannot. More generally, when most syntacticians have evaluated the data they have concluded that this binding requirement holds true of *all* NPs, and that something else must explain any instances of apparent exceptions to this general rule (Schwarz 2011).

In summary, the pronoun cannot be functioning like a bound variable, and the only way it can be functioning as a free variable is if there are various exotic things like non-existent witches. So either our ontology needs to be enlarged, or our semantics of pronouns needs to change.

1.4 More complex solutions

In large part because of donkey sentences—sentences that, like intentional identity sentences, have anaphoric pronouns whose antecedents do not c-command them—most semanticists have now moved to new accounts of the semantics of pronouns (Elbourne 2008). There is no single consensus, but undoubtedly the two most popular groups of accounts are dynamic semantic theories and description theories.[5] In Chapters 3 and 4 I discuss these two kinds of accounts respectively, and look at whether intentional identity pronouns can be functioning in either of these ways. There are numerous proposals that intentional identity pronouns are functioning dynamically, but surprisingly it seems that there is only one work in the literature (Hawthorne

[5]Many formulations of dynamic semantics do not change the traditional semantic values assigned to pronouns, but they change the values of the surrounding material in such a way that pronouns can now be dynamically bound; this is the sense in which I speak of them as new semantic accounts of pronouns.

and Manley 2012) claiming that the pronouns are functioning descriptively.[6] In chapter 4, I argue for this unpopular position—there is strong evidence that they are functioning descriptively, rather than dynamically.

Beyond those two new proposals for the semantics of pronouns, there are other proposals that have received less discussion, for example, Jacobson's variable-free semantics (1999). Some of these have included proposed accounts of intentional identity sentences: here I am thinking of King's context-dependent quantifiers (1993), Chiercia's pronouns with choice functions as semantic values (2005), and Moltmann's assigning of novel, complex semantic values to sentences (2006). There are also proposed intentional identity accounts that depart more radically from traditional semantics: game-theoretic semantic approaches (Saarinen 1978, Pietarinen 2001), and a truth predicate relativized to theories (Edelberg 1992, 1995). In this thesis, I do not give expositions of these less discussed and more radical proposals, or thoroughly evaluate them, though I do occasionally mention them in footnotes if I am aware of something brief and relevant to say.

1.5 Intentional identity in greater generality

Intentional identity debates tend to focus on (G) and very similar sentences.[7] I think that this is a mistake, that the problems posed by (G) are also posed by a broader class of sentences, and that broadening the data allows us to

[6] Neale (1990, c. 5 fn. 66) expresses some doubt about whether (G) indeed has a true reading in Geach's scenario, and endorses the weaker claim that if it does have such a reading then the pronoun is behaving semantically like a traditional definite description.

[7] The only exception I know of is King (1993) who broadens intentional identity to include non-indefinite NPs as antecedents and definite descriptions as anaphors. I think, however, that some of King's examples are ungrammatical, and that the others are not of intentional identity but rather of the phenomenon linguists call 'telescoping' (see Roberts 1989, pp. 717–8), which King then embeds under propositional attitude verbs. Telescoping is a very peculiar phenomenon. It can only occur in certain rare contexts, and it is prima facie unconnected to intentional identity sentences. Below I give examples of this sort of broadening that do not involve telescoping.

rule out accounts that would be more difficult to rule out by only focusing on sentences very similar to (G).

First, (G) can be broadened to include non-pronoun NPs as anaphors and non-indefinite NPs as antecedents. Sentence (8) above was an example with a non-indefinite antecedent NP. Below is an example with a non-pronoun anaphoric NP:

> Hob and Nob both read the local newspaper saying fifty witches from the west came to town, and they believe what they read. Furthermore, Nob thinks that every witch that the newspaper wrote about flew in on a broomstick, though he thinks a few witches from the east also came to town without using broomsticks. In actuality, the newspaper's article is false, and there are no witches.

In this scenario, the following sentence is true:

(12) Hob thinks fifty witches came to town, and Nob thinks that every witch flew in on a broomstick.

This sentence has a quantifier phrase 'every witch' rather than a pronoun, yet appears to involve the same phenomenon as (G). Somehow Nob and Hob are 'thinking about the same fifty witches', though there are no witches. Proposed intentional identity accounts focus on pronouns, but a correct account should explain the semantics of (12) as well. I propose a particular account of sentences like (12) in chapter 4.

Second, (G) can perhaps be broadened to include not just propositional attitude and intensional transitive verbs, but also other intensional contexts, such as modal expressions:

(13) If I use the MegaLure, I'll catch a fish big enough to eat. If I use the

MiniLure, it won't be big enough.

Here, like in (G), we have two different intensional contexts, each of which quantifies over different possible worlds, with an anaphoric relationship stretching across the two contexts. I think we should probably prefer a unified account of (G) and (13) over separate accounts. I am not aware of any such account, or of any account just of sentences like (13), so in chapter 4 I propose one.[8]

The rest of the thesis consists in looking for a correct account of intentional identity sentences. In chapter 2, I try to find the best account of intentional identity that uses exotic objects. In the process, I present several difficult sentences that are also counter-examples to other non-exotic accounts. I conclude that the correct account of intentional identity sentences does not involve exotic objects. In chapter 3, I look at dynamic semantics, one of the two popular current accounts of pronouns. I show how dynamic semantics is not as deeply different from traditional semantics as is often presumed. Dynamic semantics alone proves unhelpful in explaining (G), so I then evaluate theories proposed by dynamic semanticists that involve what I call 'guises'. It seems like we have convincing evidence against the most developed of these theories as well. Finally, in chapter 4, I turn to the other popular current account of pronouns, description theories. Though it appears that almost no one has ever endorsed a description theory of intentional identity, I argue that a correct description theory of intentional identity is almost inevitable, and then I propose a particular kind of description theory, before extending it to cover other cases discussed in the

[8]Note here that we do not want 'it' in (13) to have to pick out only things that satisfy its whole antecedent N$'$ 'fish big enough to eat' (e.g., by saying that 'it' means 'the fish big enough to eat'), but this would happen if we applied the pronoun semantics of (Neale 1990), (King 1993) or (Elbourne 2005). I think that (Elbourne 2005) could easily be adjusted to accord with (13) though, and I propose such an adjustment in chapter 4.

thesis.

Chapter 2

Exotica Theories

> *Geach's scenario*: The local newspaper has printed an article saying that there is a witch that has come into town. The article is completely made-up, and in fact there are no witches. Hob credulously reads the article, however, and says, 'I think this witch blighted Bob's mare'. Across town, Nob also reads the same article and says, 'Perhaps this witch killed Cob's sow'.

(G) Hob thinks a witch has blighted Bob's mare, and Nob wonders whether she killed Cob's sow.

As discussed in chapter 1, traditional semantics seems unable to give a correct account of (G) at certain contexts. The pronoun cannot be functioning as a bound variable. It also does not appear to be functioning as a free variable, for there seems to be no object that, when assigned to 'she', satisfies 'Nob wonders whether she killed Cob's sow' (the 'anaphoric sentence').

There is only one way to disagree and insist that the pronoun is functioning as a free variable: since there is no ordinary object that satisfies the anaphoric sentence at this context, one needs to posit an 'exotic' object that satisfies the sentence.[1] I will call intentional identity accounts of this sort *exotica theories*. An exotica theory is one that insists that there is some object satisfying the anaphoric sentence at Geach's context, and similarly for similar sentences and contexts. These theories will posit certain exotic objects that assimilate Geach's context for (G) to contexts where the article in the newspaper is actually true, and so where there clearly is something that satisfies the anaphoric sentence. If, for example, there are such things as 'fictional' witches, then perhaps, even if the article in Geach's context is false, there is some fictional witch that both Hob and Nob's thoughts concern, and that satisfies the anaphoric sentence.

The main proponent of an exotica theory is Nathan Salmon (1998, 2002, 2008, Forthcoming). A recent proposal by Ephraim Glick can also be usefully seen as an exotica theory (2012). Despite my awareness of only two philosophers publishing and advocating exotica theories of intentional identity, the idea is a natural, intuitive response to (G), and worthy of detailed discussion, especially since it might aspire to introduce no further complications into the semantics of pronouns. Before we move on to newer semantics of pronouns, it is worth eliminating the remaining way they could be functioning traditionally.

Below, I examine the prospects for exotica theories. First, I discuss in more detail the core exotica theory idea. Then I describe three different kinds of exotic objects and how they might be used in an account of intentional identity. My goal here is to introduce a framework for categorizing different exotica theories, so that the options in the logical space of possible views are clearer. I then argue that the best exotic objects for intentional identity are ones inhabiting a particular chunk of the logical space. Next, I

[1] Though see the qualification discussed in chapter 1 footnote 3, which applies here as well.

show that, even if there are exotica, intentional identity pronouns cannot all be functioning in the traditional way. I also present three different kinds of sentences that exotica theories apparently cannot handle, but that the later theory I will present in chapter 4 can. These are either counterexamples to exotica theories, or proof that we need a hybrid theory, explaining some of the sentences in this chapter with an exotica theory, and others with a different theory. This suggests that, with or without exotica, we need some of my account in chapter 4, and so if this account can be extended to cases like (G), it may be a preferable semantic account of (G). Finally, I present an objection to the core exotic idea, which I take to show that exotic accounts are all incorrect. Throughout the chapter, I remain agnostic about metaphysical arguments for or against the existence of exotica, and instead explore whether, if there are such things, they help in giving the semantics of intentional identity sentences.

2.1 The core exotica idea

To see the core idea of exotica theories of intentional identity, it is helpful to look at sentences like (G) but where there is obviously an object satisfying the anaphoric sentence:

(D) Hob thinks a duchess is coming into town today, and Nob thinks she's coming into town tomorrow.

Say that the local newspaper contains a truthful article that the Duchess of Cambridge is going to come to town either today or tomorrow. Hob reads the article and optimistically thinks it will be today. Nob reads it and thinks it will be tomorrow. In this context, (D) is true. It is possible that here 'she' works like a free variable assigned at this context to the Duchess of

Cambridge, and that (D) is true in part because the Duchess of Cambridge is thought by Nob to be coming into town tomorrow. With (D), there is a traditional semantic account that could work, whereas with (G) there is none. This is why the (G) case is much more puzzling than the (D) case.

I will call cases like (D) and its context—cases where there is an obvious object satisfying the anaphoric sentence—'duchess cases'. I will call cases like (G) and its context—where there is no such obvious object—'witch cases'. Of course, one could have fictitious articles about duchesses, and in at least one sense of 'witch' there clearly are witches. So duchess cases could be constructed using sentences with 'witch', and witch cases like (G) could be constructed using 'duchess'. In general, however, 'witch' tends more than nouns like 'duchess' to fit naturally with cases like (G), which is surely why Geach used it in his original example.

The core idea of exotica theories is to assimilate witch cases to duchess cases, by claiming that in the witch cases there is indeed something that satisfies 'Nob thinks she's coming into town tomorrow'. This core idea should be carefully distinguished from specific proposals about the semantics of duchess cases (and so also of witch cases, if the latter are assimilated to the former), since specific semantic proposals are logically independent from the assimilation idea.

2.2 Three different kinds of exotica

Before I look at the semantics, I will first describe particular kinds of exotic objects one might try to use, and give general criteria for sorting them. In duchess cases, Hob and Nob each have beliefs about one particular thing—a duchess. Turning to the witch case of (G), we can sort exotic object theories by their answers to three questions: (i) How many objects do Hob and Nob have these specific beliefs ascribed in (G) about—one, several, many? (ii) Are these objects indeed witches? (iii) Do we causally interact with these objects? I will describe three different kinds of exotic objects—fictional, Meinongian, and merely possible.[2] Each kind fits more naturally with different answers to the three questions, and I will argue that we have evidence telling against some of the answers.

2.2.1 Fictional objects

In the philosophical literature, Kripke (1973) and van Inwagen (1977) are the first well-known advocates of the claim that there are fictional objects. Fictional objects come into being when authors create works of fiction. These fictional objects are not made up of the fundamental physical building blocks of the universe (if such there be), and do not themselves physically interact with other objects. They are not conscious, living creatures, and so they do not have properties like being a witch or being clever. Instead, they can have properties like being one-dimensional or being frequently discussed in English departments. In general, their properties are in some robust sense dependent on the activity of intelligent beings, and if no works of fiction had been created, there would have been no fictional objects. Kripke holds that there are similar mythical objects that come into being in certain cases where people falsely believe that there is a thing that does such-and-such.

Salmon (1998, 2002) agrees with Kripke that these mythical and fictional objects exist, but he departs from Kripke by attributing more properties to them. They are still not made of atoms, or capable of physical interaction, or of being a witch or being clever. On Salmon's view, though, fictional objects are the subjects of the corresponding works of fiction, and mythical

[2] Arbitrary objects (see Fine 1983 75–76, 1985) are another kind of exotica one could try to use. See (Breckenridge and Magidor) for some objections to arbitrary objects.

objects are the subjects of the corresponding false beliefs. So there is a fictional object that Doyle pretended lived on Baker Street. Likewise, there is a mythical object that the Greeks believed lived on Mt. Olympus and threw thunderbolts. Moreover, these fictional and mythical objects are the referents of names when used in the original fiction or by the ignorant believers. When Hesiod said (the Greek equivalent of) 'And the will of great Zeus was fulfilled', 'Zeus' refered to a particular mythical object that exists thanks to the beliefs of the Greeks, and the sentence falsely asserts of this object that its will was fulfilled—falsely, since mythical objects do not have wills to fulfill. In general, Salmon holds that people's false beliefs give rise to these mythical objects, and their beliefs are about the mythical objects. If one holds that there are these fictional objects (I will use 'fictional' to cover both fictional and mythical objects), then it seems reasonable to follow (Salmon 2002) and think that Hob and Nob's thoughts about such an object are the subject of (G).

A theory of (G) with fictional objects naturally gives the following answers to our three questions: How many objects do Hob and Nob have specific beliefs ascribed in (G) about? One. Are these objects indeed witches? No. Do we causally interact with these objects? Yes—we can for example cause them to come into existence.

2.2.2 Meinongian objects

Meinongians, like mythical object theorists, postulate an abundance of objects. The distinctive idea of a Meinongian ontology is the endorsement of some sort of comprehension principle for objects, for example, that given any plurality X of properties (on an abundant picture of what properties there are), there is an object x that bears exactly those properties that are one of the Xs. When stating such principles, however, care must be taken to avoid contradictions. I will use Terence Parsons' (1980) version of Meinongianism as a specific example, since it is the most well-known.[3]

Here is Parsons' view. There are lots of things you might not have thought there were. To state more informatively what things there are, Parsons distinguishes between two distinct kinds of properties, 'nuclear' and 'extranuclear'. Most properties corresponding to predicates containing only extensional items are nuclear. The properties of swimming, of driving something, and of putting something somewhere are all nuclear. Most properties corresponding to predicates containing non-extensional items are extranuclear. Here are some extranuclear properties: being such that possibly one is a pig, being believed to be a pig.

Now we endorse the above comprehension principle when it is restricted to properties of the first kind—nuclear properties. Given any plurality X of *nuclear* properties, there is an object x that bears exactly those properties that are one of the Xs, and in fact there is *exactly* one such object.[4] Objects are also divided into two sorts: 'real' and 'unreal' (and the properties of being real and unreal are extranuclear). Typically we only interact with real objects. In fact, if we extend the concepts of nuclear and extranuclear to relations, the only objects to which we bear nuclear relations are real objects. So, for example, everything we drive is a real object. We can bear extranuclear relations to unreal objects, however. So we can think about them, and we can believe of them that they are such-and-such. Given that

[3]See (Fine 1984) for a description of the general space of Meinongian views and for where Parsons' view is located in that space.

[4]This requires us to deny that there are various sorts of nuclear properties, on pain of contradiction. For example, there is no nuclear property had by exactly those things that do not have the property of swimming, for otherwise Parsons' theory asserts that there is an object that both has this property and also has the property of swimming (Parsons 1980 §5.3).

we can believe of them that they are such-and-such, it's natural to wonder how heavily intentional identity talk involves these Meinongian objects.

Parsons himself holds that false beliefs typically do not involve believing that an unreal object is such-and-such (1980, pp. 207, 228). In general he would not hold that intentional identity frequently involves some unreal thing that is believed by Nob to be such-and-such. Nevertheless, one could try and see what happens if one abandons this claim. One would need a story of when exactly some person believes something of a particular unreal object. The simplest one would be that when someone falsely believes there is a real thing with such-and-such nuclear properties, they mistakenly believe of the unreal thing with exactly such-and-such nuclear properties that it is real. This, however, will make it too hard for (G) to be true, because it will require Hob and Nob to believe that there is a witch with *exactly* the same nuclear properties if Hob and Nob are to have a common relevant object of belief, when clearly (G) can be true even if this condition is not met, for example, if Nob but not Hob believes that the witch has some insignificant nuclear property. Moreover, it gives no guidance as to beliefs involving only extranuclear properties. Say Hob constructs a completely false theory that there is a real thing believed by Michael to be a fairy, and Nob hears Hob propounding the theory, and then believes that the thing is also believed by Luke to be a fairy. Which thing or things do Hob and Nob's beliefs concern, given that Hob and Nob have only been attributing extranuclear properties? Constructing a story of belief of unreal objects that correctly predicts the truth conditions of intentional identity discourse is non-trivial.

As regards our questions, the natural answers for a Meinongian theory are: How many objects do Hob and Nob have specific beliefs ascribed in (G) about? Unclear, and it depends on our theory of belief. Are these objects indeed witches? Yes. Do we causally interact with these objects? No—causal interactions look like relations that we only bear to real objects.

2.2.3 Merely possible objects

Merely possible objects (or 'mere possibilia') are a third category of exotica. There are different proposals of what exactly they are. On Lewis' modal realism (1986), what Lewis calls our 'world' is the totality of everything spatio-temporally related to us. Lewis believes there are an infinite number of these totalities—these things united by space-time (or space-time-like) relations, and he calls them all 'worlds'. Many things are not spatio-temporally related to us and are parts of other worlds. In some sense there is a correspondence between the possibilities for our world, and the other worlds there are. The things that are not part of our world but part of another world are Lewis' merely possible objects.

According to Williamson (1998, 2002), there are objects that do not have 'interesting' properties such as being a witch, causing this or that, or occupying a location in space. They do however have modal properties such as possibly being a witch, or possibly occupying a location in space. These are Williamson's merely possible objects. Other philosophers have discussed views involving quantification over what they called 'possible but not actual' objects, and these views in many ways resemble Williamson's (Williamson 2013).

As with Meinongian objects, a theory of intentional identity using merely possible objects needs a rough story of when exactly people have beliefs of merely possible objects. One immediate difficulty is that, for any particular merely possible object, there are an infinite number of other merely possible things that are very similar to that object. For ordinary objects, we presumably have beliefs about them in part because of causal contact with them. We have no causal contact with merely possible objects, however, so this cannot discriminate between the infinities of possible candidates for the subjects of our beliefs. And unlike in Meinongian theories, there are no 'incomplete' objects that, say, only have two nuclear properties: that of being a witch and of being tall. Instead, given any thing that could possibly be a witch with a certain qualitative profile (including its extrinsic properties), there are an infinite number of things such that, possibly, their qualitative properties are exactly the same, except that a distinct particle light-years away has a slightly different velocity. In a scenario where Hob constructs a false theory about a recent blighting, which of these merely possible things are involved in his beliefs? Discriminating between the vast sums of these mere possibilia looks difficult, so there are two options: one could bite the bullet and accept that people frequently have beliefs about an infinite number of very similar merely possible objects, or one could look for some way to cut down on the number of merely possible objects featuring in these beliefs.

Ephraim Glick (2012) proposes an account that could be seen as a 'bite the bullet' approach.[5] However, if instead of biting the bullet, we want

[5] Glick (2012) uses Lewisian metaphysics and a Lewisian counterpart relation between objects to give his intentional identity account, translating English sentences into counterpart theory talk. Each context provides a particular counterpart relation; given a particular context c, I will use $\ulcorner x$ is a counterpart$_c$ of $y\urcorner$ to say that x is a counterpart of y according to c's counterpart relation. Glick's proposal: 'Hob believes a witch blighted Bob's mare, and Nob believes she killed Cob's sow' is true at a context c iff there is a (perhaps merely possible) thing that, at every possible world compatible with Hob's beliefs, has a counterpart$_c$ that is a witch that blighted a counterpart$_c$ of Bob's mare, and at every possible world compatible with Nob's beliefs, has a counterpart$_c$ that killed a counterpart$_c$ of Cob's sow. Due to the counterpart relations, his account is contextualist about 'believes'. Like contextualist accounts of 'knows', Glick's account is hard to describe without semantic ascent. So instead of talking about when on his account there is a (perhaps merely possible) thing such that one believes certain things of it, it is more informative to talk about when 'there is a (perhaps merely possible thing) such that one believes certain things of it' is true. He does not explicitly address when these sentences are true, but numerous parts of his article suggest the following claim, which would be

to insist that typically there are a fairly small number of merely possible objects one has beliefs about, here is a rough sketch of a way one might go about doing this.[6] The following is a standard form of a sufficient condition for 'belief-of': one believes of x that it is P_1 if both (i) x has some 'suitable' property P_2 and (ii) one believes that the unique thing that has P_2 has P_1.[7] For merely possible objects, we could assume that there is no other way of having belief-of of them (since we do not causally interact with them), and strengthen this to a necessary and sufficient condition for having beliefs of mere possibilia. The substance of such a condition lies in specifying exactly which properties are suitable. We could try cutting down the number of mere possibilia involved in one's beliefs by using suitable properties involving counterfactuals. For example, it is presumably metaphysically possible for Hob's beliefs to be generally accurate and for there to be a witch wreaking havoc and bearing various other sorts of properties in accordance with Hob's actual beliefs (call this 'playing the witchy role for Hob'). There are no actual witches playing the witchy role for Hob, but there could have been things playing that role. With mere possibilia, this means that there are things such that they might have been playing the witchy role. We could say that some of the suitable properties for one's beliefs-of are those of the form: being a thing x such that, if one's witchy role had been filled, x would have filled it. Perhaps the things (or maybe just *thing*) bearing these properties are the mere possibilia featuring in Hob's beliefs.

a natural addition to his explicit account: at a particular context c, 'Hob believes x is tall' is satisfied by a variable assignment assigning x to a iff a has, at every possible world compatible with Hob's beliefs, a counterpart$_c$ that is tall. If so, then, given Glick's account of counterpart relations, at Geach's contexts 'Hob believes x blighted Bob's mare' will be satisfied by infinite numbers of mere possibilia.

[6] Another way one might go about doing this is by trying to adapt some of the ideas developed originally for arbitrary reference in (Fine 1983, 1985) or (Breckenridge and Magidor).

[7] This is formally related to Lewis's account in (1979), while jettisoning his complications meant to accommodate irreducibly *de se* belief.

This will severely cut down the numbers. Robert Stalnaker (1968, 1981) and others have held that, for counterfactuals, 'if A then $(B_1 \vee \ldots \vee B_n)$' entails '(if A then B_1) $\vee \ldots \vee$ (if A then B_n)'. Given all the things (mere possibilia included), if the witchy role had been filled, either that thing or that thing or (and so on through all the things) would have played it, so by Stalnaker's entailment (assuming it extends to sentences with infinite numbers of disjuncts), the counterfactual 'if the witchy role had been filled, that thing would have played it' is true for at least one of the things. So at least one thing bears the suitable property. Moreover, Hob believes a single thing fills the witchy role, so if the witchy role had been filled there would not have been two things filling it. So there aren't two things with the property, because otherwise if the witchy role had been filled the first thing would have filled it, and if the witchy role had been filled the second thing would have filled it, but then (by the logical properties of counterfactuals) if the witchy role had been filled, the first and second thing would both have filled it. So there is only one thing bearing this property, and perhaps this is the thing that Hob's beliefs are about.[8] The same reasoning will apply to Nob—there will be only one thing that, if Nob's-witchy-role had been filled, would have filled it. With Geach's context and this setup, it seems quite plausible that there is one thing that, if both Hob and Nob's witchy roles had been filled, would have filled it; so it seems quite plausible that their beliefs both concern the same thing.

This proposal faces some trouble, however, given the standard picture of the semantics of counterfactuals (Stalnaker 1968, Lewis 1973). Standardly, counterfactuals are contextually sensitive to all sorts of facts about the world

[8]Lewis (1973) rejects Stalnaker's entailment. On his semantics, there might not necessarily be one thing bearing the property, and one might want to play around with the formulation of the properties (perhaps by changing 'would' to 'might'), but there would certainly be significantly fewer things than on a bite-the-bullet account.

that seem irrelevant to the truth-value of certain intentional identity sentences such as (G). For example, suppose that in the actual world there are no witches in Oxford. Nevertheless, unbeknownst to Hob and Nob, there is a global coven that controls the movements of all witches and that has detailed plans for, if there ever is a witch sent to Oxford, which witch to send. Now take a scenario where two distinct newspapers entirely coincidentally happen to publish two stories, each of which is completely fabricated by people with no knowledge of witches or the coven, and each of which claims that a witch is coming to Oxford. If Hob reads one source, and Nob reads another, then 'Hob believes a witch is coming to Oxford, and Nob believes she is coming too' seems false. On the counterfactual-witchy-role proposal, however, Hob and Nob both have beliefs about the particular witch that will be sent if ever one is sent to Oxford, for if either of their witchy roles were filled, that would be the witch that would fill it. So the proposal wrongly suggests that they have beliefs about the same object, which seems to be enough for the sentence to be true.[9]

This is not a precise counterexample, because we have not made the account very precise, and we haven't given precise truth conditions for the intentional identity sentences. For one thing, we need more details about what a witchy role is, and what the full notion of suitability is. Nevertheless, it looks like the truth or falsity of intentional identity sentences will be sensitive to these intuitively irrelevant facts about the world. We must either move to a non-standard semantics of counterfactuals or modify the particular counterfactuals used.

As regards our three questions, the natural answer for a mere possibilia

[9]One might also have the converse worry that (G) can be true in some contexts where two different things would have filled Hob's and Nob's witchy roles respectively, had they been filled.

theory are: How many objects do Hob and Nob have specific beliefs ascribed in (G) about? The most straightforward theories seems to say, an infinite number. Are these objects indeed witches? On Lewis' view, yes. On Williamson's view, no—though they are possibly witches. Do we causally interact with these objects? No.

2.3 The best exotica account and the causal connection problem

Which of the three kinds of exotic objects are most suitable for intentional identity? Fictional objects. In this section, I argue for the particular answers that fictional objects give to our questions. The best exotica view is that (i) there is just one thing such that Nob wonders whether it killed the sow, (ii) this thing is not a witch, and (iii) we causally interact with this thing.

First, the best view says that there is just one such exotic object that Nob is wondering about, because such a view has an easier time getting counting right. Take an intentional identity sentence like, 'Hob thinks exactly three witches cursed the crops, and Nob thinks the three witches defaced the cathedral' and a context where it is true. It will be much easier to explain the semantics of this sentence if there are indeed exactly three things such that Hob thinks they cursed the crops. We can then say that the whole sentence is true when Nob thinks those things defaced the cathedral. If instead there are an infinite number of things such that Hob thinks they cursed the crops, what about the cases where Nob thinks some, but not all, of them defaced the cathedral? There are more moves one could make, but it seems much more natural and straightforward to avoid these enormous numbers.

Second, the best view says that the things are not witches, and that, in general, there can be true readings of $\ulcorner$Hob believes that an F is such-and-such, and Nob believes that she (the F) is so-and-so$\urcorner$ that do not entail that there is an F (or even that there is something that is possibly an F). This is because, for practically any F, one can construct a context where there seem to be true readings of such a sentence. So if such sentences did entail that there is an F, it would be very difficult to avoid contradictions within one's ontology. For example, Parsons' theory entails that there is no nuclear property had by exactly those things that do not have the property of swimming (cf. footnote 4). And yet we can construct contexts in which there will be true readings of intentional identity sentences starting with 'Hob thinks that a nuclear property had by exactly those things that do not have the property of swimming is . . . '. Again, more moves could be made here, but it seems much less difficult to avoid this problem by answering 'No' to question (ii).

Third, the best view says that we causally interact with the things. To argue for this, I introduce what I call the 'causal connection problem'. Some intentional identity sentences require a sort of causal connection between the two subjects of the sentence, and this connectedness does not seem to enter into the truth conditions in the way it affects the subjects' beliefs about ordinary objects. Consider the following two scenarios:

> *Scenario 1*: Al and Bud both suspect that some witch or other has come to town and is wrecking havoc. They get together, share their theory, and decide to warn the town, each from a separate location. They each go to their designated location and give passersby their planned speech: 'There's a witch in town! She's wreaking havoc! Be on your guard!' Hob hears Al.

> Nob hears Bud. Hob and Nob know nothing about Al or Bud before hearing them. They both believe what they hear. Later a person who observes both Hob and Nob muttering about witches speculates,
>
> (14) Hob thinks there is a witch in town, and Nob thinks she's wreaking havoc.

In such a scenario, the utterance of (14) seems clearly true.

> *Scenario 2*: Everything is like Scenario 1, except Al and Bud have never met in their life and have no coordinated plan. Al is delusional. Bud is bored and decides to start a witch-hunt by spreading a rumor that there is a witch. By coincidence, they happen to say the same thing.

In this scenario, the utterance of (14) seems clearly false.[10]

What is going on here? If we causally interact with the exotica, there is a straightforward account. For example, a fictional objects account gives an obvious difference between Scenarios 1 and 2—in the first scenario, Al and Bud collectively create one fictional object, and Hob and Nob then have beliefs about this object. In Scenario 2, however, Al and Bud do not collaborate. Al's delusions give rise to one fictional object, and when

[10] Others have noticed that sometimes a change in causal facts can accompany a change in the truth-value of an intentional identity sentence (e.g., van Rooy 2000, Cumming 2007, Glick 2012). These authors' discussions, however, have several other changes accompanying the change in the causal facts, in such a way that it is easy to attribute the change in truth-value to the other changes. For example, Glick (ibid) considers something like Scenario 2 (his 'Independent Newspaper Scenario') along with a scenario like one where Hob and Nob both listen to just Bud (Glick's 'Newspaper Scenario'), and Glick attributes the change in truth value to the change in Hob and Nob's beliefs about where they get their information from. With my Scenarios 2 and 1, however, there is no analogous change in Hob and Nob's beliefs that could explain the switch in truth-values. Glick's account predicts that (14) should receive the same truth-value in Scenarios 1 and 2, and so they are counterexamples to his account.

Bud spreads his rumor, a different fictional object is created. Hob's beliefs concern the former, Nob's the latter, and so there is no single object which their beliefs both concern. The results of the causal connection scenarios are exactly what we should expect, if we thought that they involve fictional objects, which is corroborating evidence for such an account.

If we do not causally interact with the exotica, however, the causal connection problem is very mysterious. With Meinongian objects and mere possibilia, for example, whether Al and Bud collaborate or not has no effect on which exotica there are. Given this, and the extreme similarity between Hob and Nob's experiences in both scenarios, it looks like the exotic objects that satisfy 'Nob thinks she's wreaking havoc' in the first scenario will thus be the exact same ones that satisfy it in the second scenario. But then what explains (14)'s change in truth-value?

It is not that I have a proof that there is no plausible semantic account using non-causally-interacting exotica that would explain (14); it is just that I have not seen such an account yet, cannot think of one, and find this difficulty compared to the ease with which fictional objects handle the causal connection problem striking. For example, on one of the simplest plausible semantic accounts available to exotica theorists, there is some exotic object satisfying 'Hob thinks x is a witch in town', and it is this object that gets assigned to the later pronoun 'she'. So, at each of these scenarios, (14) is true iff the exotic object thought by Hob to be a witch in town is also thought by Nob to be wreaking havoc. Since (14) is true in Scenario 1, the exotic object thought by Hob to be a witch in town is indeed also thought by Nob to be wreaking havoc. Given the extreme similarity of Scenarios 1 and 2, however, if there are no causal interactions with exotica, it is hard to see how this could fail to be true in scenario 2 as well, but this incorrectly predicts that (14) is also true in Scenario 2.

Because of these reasons, I think that fictional objects are the best objects to use in an exotica theory. I gave later objections that apply to exotica theories using any kind of object, but I take this section to be evidence against any exotica theory that does not answer the three questions in the way fictional objects do.

2.4 Exotica and traditional semantics

In this section I argue that, even with exotica, the intentional identity pronouns cannot be functioning in the traditional way.[11]

Even if there are exotica, the pronouns cannot be functioning as bound variables, for the objections from chapter 1 still apply. Furthermore, the pronouns also cannot be functioning as free variables, assigned to a particular exotic object. We can see this by embedding intentional identity sentences in other constructions. Say that the New York Times has published a false article positing a witch. Say also that the Wall Street Journal has published an entirely independently-conceived, fictitious article positing a witch with completely different properties. Michael sees Hob and Nob both thinking out loud that 'the witch' is such-and-such and so-and-so, and he knows they either both read the NYT (and so are both 'talking about the NYT witch') or both read the WSJ. He's not sure which paper they read, but he can still truly say:

(15) No matter what, it must be the case that: Hob believes that a witch is such-and-such, and Nob believes that she is so-and-so.

Here we have what looks like a quantifier over the different scenarios, and

[11]Again, with the qualification of footnote 3 in chapter 1.

‘she’ needs to pick out the NYT witch at the scenario in which they both read the NYT, and it needs to pick out the WSJ witch at the scenario in which they both read the WSJ. The pronoun ‘she’ cannot be assigned a single object simpliciter, but rather needs to covary with the quantification over scenarios. Similar sentences and contexts can be constructed with other embeddings, for example, propositional attitudes or temporal modalities. ‘Michael is wondering whether (Hob believes a witch is such-and-such and Nob believes she is so-and-so)’ can be true in a context where (i) there is no particular exotic object such that Michael is wondering whether Nob believes that exotic object is so-and-so, but rather (ii) Michael is wondering whether Hob and Nob are ‘thinking about the same thing’, regardless of what that thing is. Similar facts hold for the sentence ‘Every Sunday, Hob believes a witch is coming into town to cause havoc and Nob believes she’s coming into town to do good’. One can construct analogous sentences with NPs with ‘exactly one’ as determiners, and these sentences are counterexamples both to free-variable and wide-scoped bound-variable semantic accounts.

Because of these facts, even if there are exotica, intentional identity pronouns cannot all be functioning in one of the two traditional ways. In the next two chapters I survey other ways the pronouns could be functioning. Before I move on to the other ways, however, I wish to discuss two problems facing any exotica theory. First, there are intentional-identity-like cases that do not seem to involve exotica at all, and second, the core exotica theory idea seems problematic, even applied to the standard cases like (G).

2.5 Cases that do not involve exotica

One natural way of identifying a class of intentional-identity-like discourses in fairly precise terms, is that they are ones where an anaphoric NP in one intensional context is linked to an antecedent NP in another intensional context, and where the set of worlds quantified over by the anaphora's intensional context is not a subset of the set of worlds quantified over by the antecedent's. Some of these discourses, however, seem to have no role for exotica to play. In this section, I present three such discourses: Oneski-Twoski sentences, role sentences, and apartment sentences.

2.5.1 Oneski-Twoski sentences

Oneski-Twoski sentences were originally discovered by Edelberg (1986), and I use Glick's (2012) helpfully clear setup of the scenario. Person A has killed himself, but when the two detectives arrive on the scene, they think he has been murdered. Later, Person B kills himself, and when the two detectives arrive there, Detective Oneski thinks that the same person that murdered A murdered B. Detective Twoski disagrees, and thinks a different person murdered B. Now consider:

(16) Twoski thinks someone murdered B, and Oneski thinks he murdered A.

(17) Oneski thinks someone murdered A, and Twoski thinks he murdered B.

Sentence (16) is true, and sentence (17) is false. It is hard to see what is going on, however, if exotic objects are involved. How many exotic murderers are there? Three—two for Twoski and one for Oneski? Then one would expect (16) to be false, as they are not thinking about the same thing. Perhaps instead there are two exotic murderers—Twoski thinks each murdered only one of the people, and Oneski thinks only one of them murdered both. But then which one does Oneski think is the murderer? And also why are 'Twoski thinks someone murdered A, and Oneski thinks he murdered B' and (16) both true? Maybe in some difficult-to-see way exotic objects are involved in these sentences, but I think it's very plausible that they are not. The obvious account of these sentences is instead that (16) somehow has a reading where it means the same thing as 'Twoski thinks someone murdered B, and Oneski thinks the murderer of B murdered A'. I discuss such accounts in chapter 4.

Glick (2012) offers a response on behalf of the exoticist. Glick uses Lewis' account of merely possible objects and a Lewisian counterpart relation between objects to give truth conditions for sentences like (16), translating these sentences into counterpart theory talk. His proposal: (16) is true at the original context iff there is a (perhaps merely possible) thing that, at every possible world compatible with Twoski's beliefs, has a counterpart that killed B, and at every possible world compatible with Oneski's beliefs, has a counterpart that killed A.[12] The counterpart relation is context-sensitive: whether something is a counterpart of another thing can change depending on the context. This means that 'thinks' is context-sensitive, and Glick uses this context-sensitivity to account for (16) and (17). If we assume for simplicity that Oneski and Twoski each have a single belief world, here is Glick's picture: there are three exotic objects: one on Oneski's world that killed A and B (call it 'ABKiller'), and then two on Twoski's world, one of whom killed A ('AKiller'), the other of whom killed B ('BKiller'). When the first half of (16) is uttered, Glick posits that this creates a context where the counterpart relation heavily favors the killing-B property when deciding which things are counterparts, and so ABKiller and BKiller are counterparts, whereas AKiller is not a counterpart of anything. Thus, (16) is true, because a B-killer at Twoski's world and an A-Killer at Oneski's world are counterparts. On the other hand, when the first half of (17) is uttered, this

[12]I ignore complications arising from his Lewisian treatment of names in intensional contexts like 'A' and 'B'. See fn. 5 above for more details on Glick's theory.

creates a context where the counterpart relation heavily favors killing A, and so ABKiller and AKiller are counterparts, whereas BKiller ceases to be a counterpart of anything. Now, at this context, (17) is false, since there is nothing that has a B-killing counterpart at Twoski's world and an A-killing counterpart at Oneski's world.

Here is my objection to this response. If the 'counterpart relation' is truly a black box, where we are free to posit it holding of whichever objects we choose, then we can construct semantic accounts that mimic accounts where intentional identity pronouns mean the same thing as particular definite descriptions—in this case, where 'she' in (16) means the same thing as 'the murderer of B'. We can then mimic the predictions of the obvious account I described at the end of the paragraph before the last, and that I explain fully in chapter 4. The obvious account that is being mimicked, however, is compositional, unlike the mimicker, and can be given without appealing to exotica. So we should prefer the mimicked account over the mimicker.

2.5.2 Role sentences

In general, it seems that if we fixate on some description describing a particular role, like 'the wickedest god', if Hob and Nob both believe that the wickedest god is such-and-such, that can be enough for intentional identity sentences to be true, even if they do not believe of the same thing that it is such-and-such.

For example, say there is a community consensus that the wickedest god has been poisoning the well, killing the sow, and in general causing mischief. There is no consensus on who the wickedest god is, however. Some think it is Ares, others that it is Dionysus. Hob believes the consensus that the wickedest god has been poisoning the well, and moreover, he believes that the wickedest god is Ares. Nob believes the consensus, and moreover, believes that the wickedest god is Dionysus. Since they both believe, along with the community, that the wickedest god has been doing all these things, in this scenario the following sentence is true:

(18) Hob believes a god poisoned the well, and Nob believes he killed the sow.[13]

Here it is difficult to see how fictional (or any exotic) objects could be integral to the sentence. True, Hob and Nob have beliefs about Ares and Dionysus, respectively, but (18) seems to be true not because of their particular beliefs about those fictional gods, but because they both believe that the wickedest god, whoever that god is, is such-and-such. Again, we could try to construct an exotic account mimicking one where 'he' at this context just means the same as 'the wickedest god'. For example, we could try and use the Meinongian object that has exactly one property—that of being the wickedest god. This move seems ad hoc and inelegant. In chapter 4 I present a much more straightforward account of these cases where the pronouns here have the same meaning as definite descriptions.

2.5.3 Apartment sentences

Take the following sentences:

(19) Hob believes he will buy an apartment, and Nob thinks it will be near the city centre.

(20) Hob wants to marry a German, and Nob hopes she's from Berlin.

[13] (Friedell Forthcoming) uses a role sentence as his counterexample to the necessity of the truth conditions of an intentional identity account using fictional objects. The counterexample is credited to Donald Martin.

(21) Hob thinks he should apply for a job, and Nob thinks it should involve electrical engineering.

These sentences involve anaphoric links across different intensional contexts quantifying over different worlds, and they do not seem to involve exotica. In (19), for example, Hob may just have a general belief that he will indeed buy some apartment-or-other, and Nob may just have the general belief that, whichever apartment Hob buys, it will be near the city centre. Hob and Nob need have no particular beliefs involving particular exotica. Similarly, in (20), Hob may just have a general desire to marry only Germans. Perhaps Nob would like any German Hob marries to be from Berlin so that Nob can visit them there; if so, then (20) is true, even if there is not some particular thing such that Hob wants to marry *it*.

Exotic theories respond to (G) by positing more objects that Hob and Nob have beliefs about. The analogous move here just does not seem plausible. One could try positing more objects, but this requires saying very strange things. For example, is there a mythical or merely possible object such that Hob wants to marry it, anytime Hob wants to marry only German women? I argue no. This is clearly an incorrect account of Hob's mental state, and of people's mental states in general. For one thing, postulating more objects will entail that Hob, and in general most people, have obviously unsatisfiable wants. If Hob wants to kick a German, for example, though no German in particular, then the postulate-more-objects will have to say there is some exotic object that Hob wants to kick. But one can't kick exotic objects, so Hob has an obviously unsatisfiable want.

(21) in particular reveals the bad consequences resulting from insisting on finding exotic objects to satisfy the sentences. (21) can be true if both (i) Hob just has the general idea that he needs to stop being lazy and apply for

at least one job, and (ii) Nob has the general belief that Hob is particularly suited for electrical engineering jobs, and so should only apply to those. But now what if their beliefs are true! Then if there is indeed an exotic object that they are both thinking of, the following must be true:

> $\exists x$ (Hob should apply for x and x should involve electrical engineering)

But (21) can clearly be true, and Hob and Nob's beliefs can be true, without this existential claim being true. Sometimes it should be the case that you apply for at least one job, even though there is not something such that, it should be the case that you apply for it. In general, because anaphoric linkings across intensional contexts of this kind are possible, this response to the counterexamples would effectively force us to accept that '$\Box\exists xFx$' entails '$\exists x\Box Fx$' for deontological and similar modalities, which is disastrous.

Here is the upshot of this section: either these are counterexamples to exotica theories, or the correct account is a hybrid theory where exotica theories explain some of the discourses in this chapter, and some other account that does not appeal to exotica explains the others. It seems at least plausible that this other account could be extended to cover cases like (G) as well, eliminating the need for exotica, and in chapter 4 I propose such an alternative account and extension. Before moving on, however, I end with an objection to the core exotica idea, applied to any sentences whatsoever, and within any semantic framework whatsoever.

2.6 Objection to the core idea

I see one central problem with the core exotica theory idea, regardless of the semantic framework it is implemented in and the sentences it is implemented on. It does not explain why anaphoric intentional identity pronouns must in general appear in a non-extensional context. Exotica theories seem to predict that discourses like the following should be not only semantically acceptable, but true (since exotic objects do not poison actual wells):

(22) I do not believe there are any witches, but Hob believes a witch poisoned the well. In fact, it didn't poison the well.

It sounds true to say, 'I do not believe there are any witches, but Hob believes a witch poisoned the well. In fact, it didn't poison the well because it doesn't even exist.' So perhaps one could respond to this problem by saying that (22) sounds strange because one should also be saying the stronger claim that it doesn't exist. This is not the reason why it sounds strange, however, because, 'I'm not sure whether there are any witches, but Hob believes a witch poisoned the well. In fact, it didn't poison the well.' sounds similarly strange, though in this case one is not in a position to claim that the witch does not exist.

Of course, if there indeed *are* exotic objects, then we can have wide-scope readings of NPs picking them out, just as we can with ordinary objects. Say you and I are traveling on a bus, and one of the guys sitting next to me happens to be someone hated by our distant friend Luke currently residing off in New Zealand. I can truthfully say:

(23) Luke believes a guy sitting next to me is a jerk, but really he's not a jerk at all.

In (23), 'he' is perfectly fine appearing in an extensional context. With (G) and Geach's scenario, however, it sounds infelicitous to follow (G) with a sentence containing an anaphoric 'she' in an extensional context. So (23)'s antecedent NP at this context is more liberal in allowing anaphoric pronouns than the antecedent NP in (G) and Geach's context. If there indeed are exotic objects, then you could have similar true wide-scope readings of NPs that pick out exotic objects. For example, say that my favorite exotic object is Meinong's golden mountain, and though Luke does not know what my favorite exotic object is, he thinks that Meinong's golden mountain is ugly:

(24) Luke believes my favorite exotic object is ugly, but really it's not ugly at all.

Contexts and sentences like (G), however, have a very different feel to them than the sentences (23)–(24) and their contexts, and the antecedent NPs in sentences and contexts like (G) seem to have very different anaphoric possibilities. The core exotic idea seems to predict that sentences and contexts like (G) are like (24) and its context, with its corresponding anaphoric possibilities, but this is a false prediction, which is evidence that the idea is false.

Chapter 3

Dynamic and Guise Theories

Dynamic semantics—including File Change Semantics (Heim 1982), Discourse Representation Theory (Kamp 1981, Kamp and Reyle 1993), Dynamic Predicate Logic (Groenendijk and Stokhof 1991), and Dynamic Montague Grammar (Groenendijk and Stokhof 1990)—provides new tools for modeling the semantics of natural language. In this chapter, I look at these tools and argue that they are not helpful in giving a correct semantics of intentional identity sentences. Several dynamic semanticists have also proposed accounts of intentional identity using entities that I call 'guises'. In the last part of the chapter, I also argue that the most developed guise account is incorrect, and that guises are also unlikely to be helpful.

Here are two reasons why one might think that the dynamic semantic tools would help. First, the stereotypical intentional identity sentences are combinations of two simpler constructions—propositional attitude ascriptions and cross-sentential anaphora (also called 'discourse anaphora'):

(25) Hob believes a witch has blighted Bob's mare.

(26) A witch has blighted Bob's mare. She killed Cob's sow.

(27) Hob believes a witch has blighted Bob's mare, and Nob believes she

killed Cob's sow.

Sentence (25) is a propositional attitude ascription. Discourse (26) is an instance of cross-sentential anaphora. Stereotypical intentional identity sentences like (27) are combinations of these two simpler constructions. Since we seem to compute the meanings of intentional identity sentences from the meanings of the words and the structure of the sentences, it is natural to think that the correct semantic account of constructions like (25) and (26) would also predict the correct semantics of intentional identity sentences.

Semanticists have not reached a consensus on the correct semantic account of (26), or of cross-sentential anaphora in general (King 2010). As I discuss in the first chapter, there are two traditional ways pronouns have been taken to function: like a free variable, and like a bound variable. The free variable way of functioning is one plausible account of (26). The bound variable account is unacceptable, for the reasons I discuss in the first chapter: it predicts the wrong truth conditions for cross-sentential anaphora involving antecedents with other determiners like 'exactly one', and it violates consensus views in syntax. Because of other data (e.g., donkey sentences) that the traditional semantics has struggled to explain, several new semantics have been proposed. Currently, the two popular views of pronouns are those of (i) dynamic semantics and (ii) description theories. In dynamic semantics, we modify the traditional semantics of the surrounding material in such a way that the pronoun in (26) can indeed function like a bound variable, one that is 'dynamically bound'. Someone who held that dynamic semantics was the correct account of (26) will presumably think that it, combined with an account of propositional attitude ascriptions, will give the correct account of intentional identity.

Modal subordination accounts are a second reason one might think dynamic semantic tools help with intentional identity. Craige Roberts (1989) coined the term ‘modal subordination’ to describe a phenomenon exhibited in the following discourse:

(28) A wolf might come in. It would eat you.

There is a reading of this sentence where ‘might’ takes wider scope than ‘a wolf’, and ‘it’ is anaphorically related to ‘a wolf’, even though there does not have to be a wolf such that it might come in and eat you. Linguists have understandably thought there was a connection between sentences like these and intentional identity sentences (Roberts 1997). Indeed, both (28) and a typical intentional identity discourse involve two intensional contexts, with a pronoun in the latter context anaphorically linked to an antecedent in the former. Several linguists have also thought that dynamic tools provide the correct semantic account of modal subordination (Stone 1999). So it is natural to think that they also provide the correct account of intentional identity.

In our traditional semantics, bound pronouns must be c-commanded by their binders. The central feature underlying all dynamic semantic accounts is that they effectively allow NPs to be semantically bound by binders that do not c-command them. It is very controversial whether or not natural language has this non c-command binding, and the debate will presumably be resolved by looking at a much wider range of data than just intentional identity sentences. In this chapter, therefore, I do not address the broad question of whether dynamic semantic tools are appropriate in giving the entire semantics of English. Rather, I assume that there are English pronouns that are bound dynamically as dynamic semanticists believe, but I argue that the pronouns in intentional identity sentences are not functioning in this way.

Below, I first explain what exactly the tools of dynamic semantics are. I then explain the connection between modal subordination and intentional identity, and how dynamic tools can provide a plausible—though not the only plausible—account of modal subordination. It turns out that modal subordination involves an easier version of the same problem found in intentional identity. While dynamic tools can account for the easier problem in modal subordination, they predict the wrong truth conditions when extended in the obvious way to intentional identity. I then look at less obvious accounts from some dynamic semanticists that involve things I categorize as 'guises', things like Fregean senses. Non-dynamic semantics could also try to give an account using guises, and I evaluate the prospects for such theories as well. The conclusion of this chapter is that dynamic tools and guise theories do not help in providing a semantic account of intentional identity.

3.1 Extensional dynamic semantics

Dynamic semanticists often disagree on substantial goals or claims that they individually incorporate into their particular version of dynamic semantics. For example, Kamp (e.g., Kamp and Reyle 1993) believes that many components of his semantics mirror the way actual human brains compute the semantic values of non-atomic chunks of language. Groenendijk and Stokhof (e.g., 1991 §5.2) do not attempt to model actual human computation as fully as Kamp does. Consequently, Kamp's and Groenendijk and Stokhof's semantics differ from each other in important ways. Despite differences between dynamic semanticists, however, there is one main property shared by all of the accounts. On dynamic semantics, an indefinite description and a pronoun (and in general a definite NP) can behave semantically as if the indefinite semantically binds the pronoun, even if the indefinite does not syntactically c-command the pronoun. As discussed in chapter 1, in traditional semantics, indefinites can bind pronouns and other definites, but only if the indefinites c-command them in the syntactic structure given to semantics. Dynamic semantics introduces new machinery to allow non-c-commanding semantic binding.

Dynamic semantics is much closer to orthodox truth conditional semantics than it might appear in its traditional presentation. If the notations are brought closer to each other, some of the apparent differences are revealed to be superficial notational peculiarities of a particular presentation of one or the other. In this section I will show how to turn the standard, orthodox Heim-Kratzer semantics into a compositional dynamic semantics in a few steps.

Here are traditional, non-dynamic semantic values that one would find in the standard semantics textbook (Heim and Kratzer 1998), which is used as a starting point in much recent semantics literature. For all variable assignments, g:

$[\![\text{`Richard'}]\!]^g = \text{Richard}$

$[\![\text{`he}_1\text{'}]\!]^g = g(1)$

$[\![\text{`Richard paints'}]\!]^g = \text{True iff Richard paints, and False otherwise}$

$[\![\text{`he}_1 \text{ paints'}]\!]^g = \text{True iff } g(1) \text{ paints, and False otherwise}$

$[\![\text{`paints'}]\!]^g = \lambda x_e.x \text{ paints}$

$[\![\text{`physicist'}]\!]^g = \lambda x_e.x \text{ is a physicist}$

$[\![\text{`and'}]\!]^g = \lambda p_t.\lambda q_t.(p = q = \text{True})$

$[\![$'a physicist paints'$]\!]^g$ = True iff $\exists x$(x is a physicist and x paints), and False otherwise

$[\![$'a physicist'$]\!]^g = \lambda f_{et}.\exists x$($x$ is a physicist and $f(x)$ = True)

$[\![$'a'$]\!]^g = \lambda e_{et}.\lambda f_{et}.\exists x(e(x) = f(x) = \text{True})$

Quantifier NPs in predicate position are handled with Quantifier Raising (see section 1.2). There are three semantic composition rules: Functional Application, Predicate Abstraction, and Predicate Modification. For simplicity, I avoid examples where the latter two are necessary, so we only need the first, in the same form as in chapter 1:

Functional Application For all variable assignments g, if α has β and γ as immediate daughters, and $[\![\gamma]\!]^g$ is in the domain of $[\![\beta]\!]^g$ then,

$$[\![\alpha]\!]^g = [\![\beta]\!]^g([\![\gamma]\!]^g)$$

Given the semantic values assigned above to single words, this rule generates the other semantic values above. As in chapter 1, a sentence is true at a context if the sentence's semantic value, relative to the contextually-provided variable assignment, is True.

Now we will turn this traditional semantics into a dynamic one. First, our traditional semantics assigns semantic values relative to variable assignments. A notational variant would be to assign semantic values simpliciter—the new semantic values can be functions from variable assignments to the old semantic values that were had relative to that variable assignment. For step one, we enact this suggestion. We need a new type v, the type of variable assignments.[1] Every expression now will have a new type, vτ, where τ

[1]Since variable assignments are themselves functions, we could incorporate this into

is its old type. On the old semantics, we would define the semantic values by saying that, for some bit of language ϕ, for every variable assignment g, $[\![\phi]\!]^g$ is such-and-such. Now we rephrase all those definitions and instead say that $[\![\phi]\!]$ is $[\lambda g_v.$ such-and-such$]$. So for example, a sentence is of type vt. $[\![$'He$_1$ paints'$]\!]$ is the function that, for every variable assignment, takes it to True if the thing that it assigns to 1 paints (and takes it to False otherwise). A sentence at a context is true iff the semantic value of the sentence takes the variable assignment provided by the context to True.

We also need to rephrase the semantic composition rules. Functional application will now say that, if α has as immediate daughters β and γ, where, for all variable assignments g, $[\![\gamma]\!](g)$ is in the domain of $[\![\beta]\!](g)$, then $[\![\alpha]\!]$ is the function $\lambda g_v.[\![\beta]\!](g)([\![\gamma]\!](g))$. This is just the notational equivalent of the old functional application rule that 'passed down' the variable assignment to the daughter nodes.

This is the most obvious way to incorporate the variable assignments into the semantic values themselves. Instead of using this way, however, I will move to a slightly different way of incorporating variable assignments into semantic values that strikes me as more elegant. We can use a simpler functional application rule, and change only the types of those basic linguistic items that are sensitive to the choice of variable assignments, while letting all the other types be whatever they need to be for functional application to still hold in its simple form.[2]

Functional Application, Simpler Version If α has β and γ as immedi-

the semantics by having a new semantic type i of indices. Instead of having a basic type for variable assignments, they would then be of type ie.

[2]In our orthodox semantics with functional application, we have some semantic values—e.g., those of names—that are only input into other functions. We have other semantic values—e.g., those of verb phrases—that can be functions in which other values are inputted. The sort of semantics being given in this paragraph becomes more complicated and unnatural when dealing with variables whose semantic value is of the latter kind. In this chapter we will not be using such variables, so this will not concern us.

ate daughters, and $[\![\gamma]\!]$ is in the domain of $[\![\beta]\!]$ then,

$$[\![\alpha]\!] = [\![\beta]\!]([\![\gamma]\!])$$

To implement this, we again let proper nouns (names and pronouns) be type ve, and again we let sentences be type vt, but instead of also appending a 'v' on the front of every other expression type, we let the other types be dictated by what the simpler functional application rule needs them to be. For example, intransitive verbs are functions from the type of proper nouns to the type of sentences. So now they are of type (ve)vt. Indefinite articles are functions from the type of intransitive verbs to functions from the type of intransitive verbs to the type of sentences. So now they are of type ((ve)vt)((ve)vt)vt. Sentence connectors like 'and' are functions from the type of sentences to functions from the type of sentences to the type of sentences. So now they are of type (vt)(vt)vt. With these type assignments, we can keep our simpler functional application rule. Nodes whose semantic values are computed by this rule are once again just the value of the application of one semantic value to another. Here are sample lexical entries:

$[\![$'Richard'$]\!] = \lambda g_v.$Richard

$[\![$'he$_1$'$]\!] = \lambda g_v.g(1)$

$[\![$'Richard paints'$]\!] = \lambda g_v.$Richard paints

$[\![$'he$_1$ paints'$]\!] = \lambda g_v.g(1)$ paints

$[\![$'paints'$]\!] = \lambda a_{ve}.\lambda g_v.a(g)$ paints

$[\![$'physicist'$]\!] = \lambda a_{ve}.\lambda g_v.a(g)$ is a physicist

$[\![$‘and’$]\!] = \lambda p_{vt}.\lambda q_{vt}.\lambda g_v.(p(g) = q(g) = \text{True})$

$[\![$‘a physicist paints’$]\!] = \lambda g_v.\exists x(x \text{ is a physicist and } x \text{ paints})$

$[\![$‘a physicist’$]\!] = \lambda e_{(ve)vt}.\lambda g_v.\exists x(x \text{ is a physicist and } e(\lambda h_v.x)(g) = \text{True})$

$[\![$‘a’$]\!] = \lambda f_{(ve)vt}.\lambda e_{(ve)vt}.\lambda g_v.\exists x(f(\lambda h_v.x)(g) = \text{True and } e(\lambda h_v.x)(g) = \text{True})$

Note that the semantic values of verbs are now of type ((ve)vt). So they take functions of type (ve), rather than entities. This means that when we quantify over entities in the definitions of the semantic values of ‘a physicist’ and ‘a’, we cannot plug an entity x into the ((ve)vst) functions, so we must turn x into the (ve) function taking every variable assignment to x, and then we plug that (ve) function into the ((ve)vst) function. A similar type-shifting will occur as well on our dynamic semantics.

Now we take the step to a dynamic semantics. Our previous, orthodox semantics has the semantic value of sentences be functions from variable assignments to truth values. The dynamic step is to change the sentence type from vt to vvt, and then to assign special semantic values to certain bits of language. Our traditional semantics in chapter 1 achieved c-command binding by introducing variable assignments. In effect, by complicating the semantic value of sentences further, we can use variable assignments to achieve non c-command binding as well. We subscript indefinite determiners with numerals so that they can bind later pronouns with the same numerals. This will let us say that ‘A_1 physicist paints. He_1 drums.’ is true iff there is something that is a physicist that paints and drums. Effectively, the quantification introduced by ‘A_1’ will bind ‘he_1’.

Here is a rough overview of how we accomplish this, before we look at the details. We can think of our earlier semantic values—functions of type vt—as sets of variable assignments, those that the function takes to True. Context provided a variable assignment, and a sentence at a context was true iff the variable assignment provided by the context was in the sentence's set. On our dynamic semantics, we can think of the semantic values of sentences—functions of type vvt—as relations between variable assignments. Each such function corresponds to a particular relation—the one holding between two variable assignments iff the function takes the first variable assignment to a function that takes the second variable assignment to True. A single sentence is true iff its semantic value relates the contextually-provided variable assignment to at least one variable assignment. So at the level of single sentences, it does not matter which particular variable assignments the contextually-provided assignment relates to. As long as it is related to at least one, it is true. The details of the variable assignments it is related to, and the number, as long as the number is 1 or greater, are irrelevant. These details only become relevant when we look at strings of sentences.

Sentences can be strung together by conjunction-like operators into one large sentence-like thing. Context provides a variable assignment, and this big sentence-like-thing is true at the context if the variable assignment is related by the semantic value of the first sentence to at least one variable assignment that is related by the semantic value of the second sentence to ... at least one variable assignment that is related by the last sentence to at least one variable assignment. In this way, a sentence can influence what variable assignments are received by sentences on down the line, and through this effect we can allow indefinites in earlier sentences to bind pronouns in later sentences.

Now for more precise details. Dynamic semantic values of sentences are functions from variable assignments to functions from variable assignments to truth values. All other type changes follow from this, by following the method of the previous step, adjusting the types of other expressions so that functional application still holds. Intransitive verbs are now of type (ve)vvt. The indefinite article and other determiners are now of type ((ve)vvt)((ve)vvt)vvt. Sentence connectors are now of type (vvt)(vvt)vvt.

As for semantic values, names and pronouns stay the same:

$[\![\text{‘Richard’}]\!] = \lambda g_v.\text{Richard}$

$[\![\text{‘he}_1\text{’}]\!] = \lambda g_v.g(1)$

A simple sentence composed of a name or pronoun and an intransitive verb acts like a ‘filter’. Relationally, it relates variable assignments only to themselves, and then only if they meet certain conditions. A sentence like ‘Richard paints’ relates any variable assignment to itself iff Richard paints—so the sentence either relates all variable assignments to themselves or no variable assignments to themselves, depending on whether it is true or false. A sentence like ‘He$_1$ paints’ relates a variable assignment to itself iff the variable assignment assigns 1 to something that paints:

$[\![\text{‘Richard paints’}]\!] = \lambda g_v.\lambda h_v.(g = h \text{ and Richard paints})$

$[\![\text{‘He}_1\text{ paints’}]\!] = \lambda g_v.\lambda h_v.(g = h \text{ and } g(1) \text{ paints})$

A sentence at a context is true iff the variable assignment given by context is related by the semantic value of the sentence to *at least one* variable assignment. So ‘Richard paints’ is true at a context iff Richard paints. ‘He$_1$ paints’ is true at a context iff the contextually-provided variable assignment assigns 1 to something that paints.

We assign intransitive verbs and common nouns appropriate semantic values that will deliver the above semantic values for simple sentences:

$[\![$'paints'$]\!] = \lambda a_{ve}.\lambda g_v.\lambda h_v.(g = h$ and $a(g)$ paints$)$

$[\![$'physicist'$]\!] = \lambda a_{ve}.\lambda g_v.\lambda h_v.(g = h$ and $a(g)$ is a physicist$)$

Here is how an indefinite in a sentence like 'A_1 physicist paints' manages to semantically bind later pronouns. The semantic value of this sentence is not a filter. Instead, given any variable assignment, it relates that variable assignment to exactly those that both (i) differ at most from it in what they assign to 1 and (ii) assign 1 to a physicist that paints. We write '$g[1]h$' to abbreviate 'g and h differ at most in what each assigns to 1', and we want the following semantic value for a sentence with an indefinite:

$[\![$'A_1 physicist paints'$]\!] = \lambda g_v.\lambda h_v.(g[1]h$ and $h(1)$ is a physicist and paints$)$

Our conjunction-like operator, ';', connects sentences and makes these variable assignments relevant. The semantic value of groups of sentences combined by ';' are of the same type as that of single sentences—type vvt. Groups of sentences relate one variable assignment to another if there is a 'path' through the sentences starting at the former and ending at the latter:

$[\![$'He_1 paints; Richard paints'$]\!] = \lambda g_v.\lambda h_v.\exists i_v([\![$'$\mathrm{He}_1$ paints'$]\!](g)(i) =$ True and $[\![$'Richard paints'$]\!](i)(h) =$ True$)$

So:

$[\![$';'$]\!] = \lambda q_{vvt}.p_{vvt}.\lambda g_v.\lambda h_v.\exists i_v(p(g)(i) =$ True and $q(i)(h) =$ True$)$[3]

Consequently:

$[\![$'A_1 physicist paints; He_1 drums.'$]\!] = \lambda g_v.\lambda h_v.\exists i_v(i[1]g$ and $i(1)$ is a physicist and paints and $i = h$ and $h(1)$ drums$)$

[3]I'm thinking the syntactic trees are right-branching at the level of ';' and the sentences it connects, e.g., like these two trees:

Which reduces to:

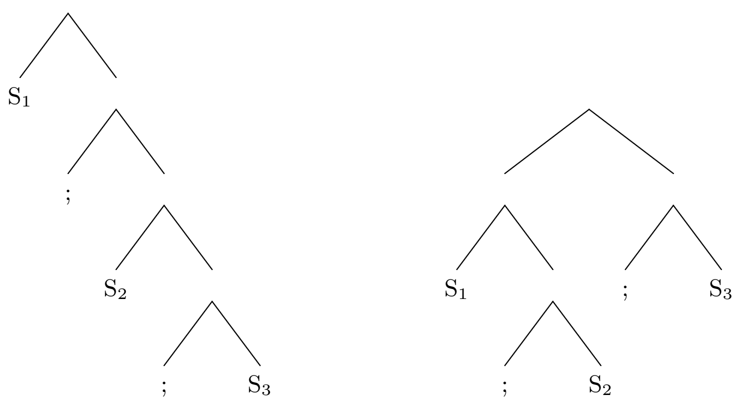
$[\![$‘A_1 physicist paints; He_1 drums.’$]\!] = \lambda g_v.\lambda h_v.(h[1]g$ and $h(1)$ is a physicist and paints and $h(1)$ drums)

The truth of groups of sentences is the same as the truth of a sentence: a group of sentences is true at a context iff the contextually-provided variable assignment is related by the semantic value of the group to at least one variable assignment. So the above group is true iff there is a variable assignment that differs from the contextually-provided one at most in what it assigns to 1, and that assigns to 1 something that is a physicist that paints and drums. Because of our definition of variable assignments, this means that the group is true iff there is something that is a physicist that paints and drums. We have achieved the semantics that allows binding behavior between items neither of which syntactically c-command each other.

Given the semantic value of ‘physicist’, ‘paints’, and ‘A_1 physicist paints’, here are semantic values for ‘A_1 physicist’ and ‘A_1’ that deliver what we want:

$[\![$‘A_1 physicist’$]\!] = \lambda e_{(ve)vvt}.\lambda g_v.\lambda h_v.\exists i_v(g[1]i$ and $i(1)$ is a physicist and

---



Given $[\![$‘;’$]\!]$, both of these trees have the same semantic value. If we want left-branching trees, however, we would want to swap the first instances of ‘p’ and ‘q’ in the description of the semantic value of ‘;’.

$e(\lambda k_v.i(1))(i)(h) = \text{True})$

$$[\![\text{‘A}_1\text{’}]\!] = \lambda f_{(ve)vvt}.\lambda e_{(ve)vvt}.\lambda g_v.\lambda h_v.\exists i_v \exists j_v (g[1]i \wedge f(\lambda k_v.i(1))(i)(j) = \text{True} \wedge e(\lambda k_v.i(1))(j)(h) = \text{True})$$

Note that here we are again type-shifting when we input things into the functions that are the semantic values of common nouns and verb phrases. Also, the quantification over variable assignments in the clause for $[\![$‘a’$]\!]$ (which uses the metavariables ‘i’ and ‘j’) is to allow the $\overline{\text{N}}$ and the VP material attached to the determiner to contain dynamically binding NPs as well.

As in our earlier non-dynamic semantics, indefinite NPs out of subject position would cause a type mismatch on the current semantics once we added transitive verbs, and we could handle this mismatch by introducing either quantifier raising or type-shifting. I will avoid these complications by focusing on simple sentences.

Dynamic semanticists often describe context as being a richer notion than it is on orthodox semantics. To motivate all of this machinery, they often say that the meaning of a sentence is not just its truth-value relative to a variable assignment provided by context, but that sentences have richer meanings—the ways in which they can change the context. There is no need, however, to significantly change our notion of context when we use the dynamic non c-command binding. The traditional Kaplanian (1989) notion of context is of an n-tuple containing an agent, a location, a time, and other variables needed to determine the truth-value of sentences containing context-sensitive words like ‘I’ and ‘this’. In our orthodox semantics we have combined this notion of context with that of a variable assignment. To handle unbound pronouns like ‘he’, we assume that corresponding to an occurrence of a pronoun is an aspect of the context that identifies the object to which the pronoun refers. Formally, we model this by associating a number with each occurrence of a pronoun, and then by postulating a variable assignment that assigns the number to the object to which the pronoun refers. We say that a sentence at a context is true iff it is true relative to the variable assignment given by that context. We also use variable assignments to give the semantics of quantifiers and bound terms. At different points in the computation of semantic values, the variable assignments under consideration will change. This does not mean we need to identify the variable assignments under consideration with the context of the particular part of the sentence whose semantic value is then being computed. We can keep 'context' as a word for the Kaplanian n-tuple, and separate its function from that of the machinery needed for the semantics of binding. If we wanted to, we could make this separation clear by having two devices, a variable assignment that handles binding, and an n-tuple that is the context, each performing distinct functions.

Similarly, dynamic semantics needs a more complicated variable assignment machinery to give its semantics of binding. Again, at different points in the computation of semantic values, the variable assignments under consideration will change. We can still keep 'context' as a word for the Kaplanian n-tuple that is involved in unbound pronouns, and separate its function from that of the variable assignment machinery if we so desire.

We have the dynamic semantics for discourse anaphora, and now we need a semantics for propositional attitude ascriptions, so that we can see what dynamic semantics we can give to stereotypical intentional identity sentences. Focusing on 'believes', there are several different extant accounts in the dynamic semantics literature. Asher (1986, 1987) proposes a complex semantics of propositional attitude verbs, including 'believes', in order to account for a wide range of phenomena. Kamp (1990) proposes to use first-order-logic-sentence-like constructions called DRSs to model the cognitive state of believers. More recently, Cumming (2007) also uses DRSs to mirror private mental states, and Maier (2009) gives an account in which belief *de re* and *de se* are modeled by heavy use of dynamic presuppositions.

None of these accounts are compositional except for Cumming's. Moreover, all of these accounts contain very complicated machinery enabling them to account for, *inter alia*, failures of logical omniscience, Puzzling Pierre cases, Paderewski cases, and belief irreducibly *de se*. None of this complexity is obviously essential for handling just intentional identity, especially given that analogous problems arise with the non-propositional-attitude modal sentences discussed in chapter 1. In traditional formal semantics, there is a simple way of modeling the semantics for 'believes' intensionally—treating it like a quantifier over a restricted set of possible worlds, those in which the subject's beliefs are true. The basic intensional semantics will not necessarily handle belief *de se* and other such puzzles, but it seems plausible that if a basic intensional semantics with dynamic tools cannot provide an account of intentional identity, then neither can the fancier, more complicated semantics. I will supplement our dynamic semantics with such an intensional account of 'believes', so that we can more easily see how intentional identity might be handled with the distinctive dynamic tools. Dynamic accounts of modal subordination tend to use this much simpler model of intensionality, so I turn first to modal subordination before discussing 'believes'. Once we have a simple dynamic account of modal subordination on the table, the analogous intensional account of 'believes' is obvious.

3.2 Modal subordination

3.2.1 Two problems

Linguists have often thought there was a connection between modal subordination and intentional identity. In this section, I describe what the connection is. In brief, the typical modal subordination cases combine two distinct problems—that of continuation and of anaphora. The first problem of continuation is interesting in its own right. What the literature does not seem to have noticed is that the second problem, that of anaphora, is a less difficult instance of the more difficult general problem of intentional identity. Once we see how dynamic semantics has been used to account for the second problem of anaphora, we can see if the same solution can be extended to the more difficult problem of intentional identity.

Roberts (1989) distinguished two problems that modal subordination can involve. Here are some examples of modal subordination:

(29) A wolf might come through the door. Don't worry—I would protect you with my knife.

(30) A wolf might come through the door. It would eat you.

(31) If a hurricane comes, the weather will get ugly. Rain will start pelting down.

(32) If a hurricane comes, it will be a Category 4. The hurricane will destroy all the ocean front property.

(33) According to John's theory, I am a goblin. I dress up like a human to hide my true identity.

(34) According to John's theory, a witch is poisoning the well. She is coming at night and dropping in anthrax bacteria.

What makes these cases of modal subordination is the particular way in which the first sentence of each discourse affects the truth conditions of the second sentence of each discourse. Roughly what is happening is that the first sentence introduces and describes some possibilities, and the second sentence continues to describe those possibilities. For example, the second sentence of (29) means something like, 'Don't worry—if a wolf came through the door, I would protect you with my knife.' Likewise, the second sentence of (32) is not the general prediction that a particular hurricane will destroy some property; instead the prediction is only that *if* a hurricane comes, it will destroy the property. This is the first problem of modal subordination, the continuation problem. We want to find the correct semantic account of how the possibilities introduced by the first sentence affect the truth conditions of the second sentences. Moreover, we want a general description of when such continuation is allowed and when it is not, and we would like an explanation of why this general description is true.

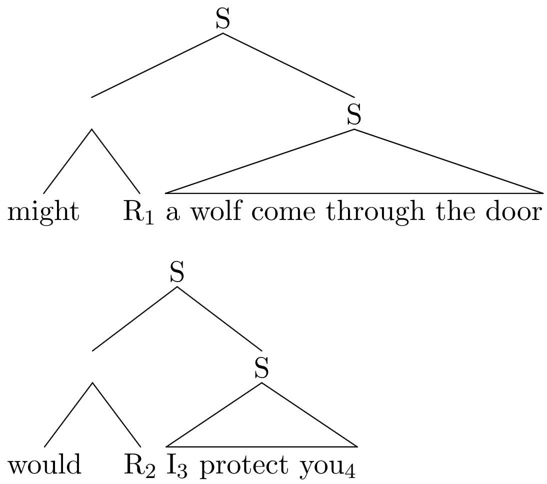
The odd-numbered discourses involve only the first problem of modal subordination. The even-numbered discourses, however, introduce a second problem. In the second sentences of the even-numbered discourses, we have pronouns or definite descriptions whose semantic function somehow depends on an indefinite description in the first sentence. As we will see below, one can handle the first problem, giving truth conditions for all of the odd-numbered sentences, without dealing with the second problem. The two problems are distinct. The first problem is not directly related to intentional identity. The second problem, however, is a specific instance of the more general problem that intentional identity raises. Finding the fully correct account of the even-numbered sentences requires addressing this more difficult general issue.

3.2.2 Continuation: the first problem

By using the standard Kratzer (e.g., 1977) account of modals, we can more precisely describe what is occurring in the first problem of modal subordination. On this account, modal words quantify over a contextually-dependent collection of possible worlds. Words like 'might' existentially quantify over those worlds, and words like 'would' universally quantify over them. As a first pass, 'A wolf might come through the door'—at a context where 'might' is taken to mean something like 'for all we know'—is true iff, at some world compatible with what we actually know, a wolf comes through the door. Context in this case determines that the set of worlds quantified over are those compatible with what we actually know. To allow for the contingency of 'A wolf might come through the door', we say that the context determines a relation between worlds, rather than just a set of worlds. This relation will, in this context, relate each world w to the worlds which are compatible with what we know at w. Thus, 'A wolf might come through the door' is true relative to a world of evaluation w iff, given the contextually determined relation between worlds R, some world that R relates w to is a world where a wolf comes through the door.[4]

We posit the following syntactic structures for the two sentences in (29), shortening the second sentence a bit:

[4]Kratzer gives a complex account of the ways context can determine the worlds quantified over. On her account, context determines two different devices—a modal base, and an ordering source. The modal base is a function from worlds to a set of propositions. In our previous case, this would be the function from worlds to the set of propositions we know at those worlds. The ordering source is a partial ordering of possible worlds that allows Kratzer to handle more complicated scenarios. Like Lewis's account of counterfactuals, Kratzer allows for a partial ordering where, for some world w, there is no world or worlds closer to w than all other worlds. If we reject this possibility, however, we can follow von Fintel and Heim (2011) and simplify her semantics into one where context determines a single accessibility relation between worlds. This relation holds between two worlds iff (i) the former is taken by the modal base to a set of propositions that are all true in the latter world and (ii) according to the ordering source, there is no world closer than the latter world to the former world.



Here each 'R' is a variable of type sst. This will be assigned the accessibility relation determined by context. The first sentence affects which worlds are quantified over in the second sentence. In our model, this means that the first sentence affects the assignment to 'R_2'. Because of the first sentence, 'R_2' appears to be assigned a restriction that relates two worlds iff both (i) the 'R_1' restriction relates the two worlds and (ii) a wolf comes through the door in the second world. If context provides such a relation, then we have an accurate prediction of when the sentence is true at that context, given Kratzer's semantics. Assuming 'R_1' is assigned the relation that relates each world w to the worlds compatible with what we know at w, the sentence is true at that context iff I protect you at every world where (i) every proposition we know is true and (ii) a wolf comes through the door.

A full account of modal subordination has more work to do on the first problem. One might want to more explicitly incorporate into the semantics what I have described as the first sentence affecting the context in which the second is interpreted. Moreover, the exact general rule describing how sentences can affect the worlds quantified over by following sentences is controversial, and the explanation for why this rule exists is controversial as well. Numerous accounts have been given (e.g., Roberts 1989, Stone 1999). Abstracting away from that debate, however, we can use Kratzer's account of modality to describe what in general is going on when the second sentences continue on describing the possibilities raised by the first sentences in modal subordination. With this grip on the first problem, we can look at the second.

3.2.3 Anaphora: the second problem

Without committing ourselves to a more detailed account of the first problem, we can look at how to handle the second problem. Repeated below is our discourse (30):

(35) A wolf might come through the door. It would eat you.

In Kratzer terms, we can see that the contextually-provided accessibility relation paired with the modal in the second sentence must only relate worlds to those where a wolf comes through the door. The pronoun in the second sentence, though, presents a difficulty. Neither of the two ways pronouns function according to traditional semantics will work. Here 'it' is not picking out some particular thing, because I am not speaking of a particular wolf that might come through the door. I am instead saying that the following is possible: some wolf or other comes through the door. Also, 'it' is not functioning as a traditional bound variable. For one thing 'a wolf' needs to take smaller scope than 'might'; 'might' does not have scope over 'It would eat you', so neither does 'a wolf'.

One could try and solve the problem by treating the pronoun semantically like a traditional definite description (Elbourne 2005 §2.6.2), as I discuss in the next chapter. But most proposed accounts use dynamic semantics to solve the second problem of modal subordination (see Stone p. 5).

They treat the pronoun as a variable that is bound dynamically by the indefinite 'a wolf' in the previous sentence. Crucially, an intensional version of dynamic semantics will allow the indefinite to bind the pronoun while still staying within the scope of 'might' in the first sentence. This semantics will effectively say (when combined with a solution to the first problem) that (35) is true iff both (i) it might be the case that a wolf comes through the door and (ii) if a wolf comes through the door, it would be the case that a wolf that comes through the door eats you.

All the dynamic semantic accounts of modal subordination that I know of are either not compositional (Roberts 1989) or very complicated (e.g., Stone 1999, Brasoveanu 2007). In the next section, I present the simplest compositional dynamic semantics that I can see that adequately handles modal subordination. I do not think the more complicated accounts add anything that would help with intentional identity.

3.3 Intensional dynamic semantics

Here is an overview before we dive into the details. To handle modal subordination—and eventually propositional attitudes—we introduce two changes. First, we intensionalize the semantic values. As in traditional intensional semantics, we have a new type s of worlds. In every old type where we had the type t of truth values, we replace this with the type st of functions from worlds to truth values. Since our extensional dynamic semantic values for sentences was vvt, the new intensional type of these sentences is vvst. These can be seen as three-place relations between a variable assignment, another variable assignment, and a world. Informally, $[\![$'John runs'$]\!]$ was a filter—it only related variable assignments to themselves. It related all variable assignments to themselves if John runs, and none to themselves if he does not. Now, it will only relate variable assignments to other variable assignments relative to a world. Relative to worlds where John runs, it will relate a variable assignment to itself (and to itself alone). Relative to worlds where John does not run, it will relate no variable assignments to themselves (or to any other). Similarly, $[\![$'he$_1$ runs'$]\!]$ took a variable assignment to itself iff the thing that it assigned to 1 does indeed run. With our intensional semantics, we will care about what the thing does at a particular world. $[\![$'he$_1$ runs'$]\!]$ will relate a variable assignment to itself, relative to a particular world, iff what the variable assignment assigns to 1 does indeed run at that world.

The second step is that we need to change our variable assignments. Earlier they were assignments of entities. Now, they will need to be assignments of *individual concepts*—functions of type se, functions assigning each world an entity. Here is why we need these more complicated variable assignments. Informally, our dynamic semantics tools let an early sentence with an indefinite in a discourse 'pass on' a set of entities to a later sentence involving a pronoun anaphorically linked to that indefinite. Say the discourse is two sentences: 'A$_1$ man walked in the park. He$_1$ smoked.' Context provides a variable assignment (though it is irrelevant because there are no free-variable-like pronouns). The whole group of sentences is true iff the semantic value of the first sentence relates the contextually-provided variable assignment to a variable assignment that the semantic value of the second sentence relates to at least one variable assignment. Since there is an indefinite description in the first sentence, the semantic value of the first sentence is going to relate the contextually-provided variable assignment to all those that both (i) agree on the non-1 assignments and (ii) assign 1 to some man who walked in the park. In this sense, the sentence is passing on a set of entities to later sentences using pronouns subscripted with '1'—the set of entities that are men who walked in the park. The second sentence will be true iff at least one of these entities smoked.

When we come to intensional languages like those containing 'might', we need to pass along more information than just a set of entities. We need to pass along ordered pairs of worlds and entities. Why do we need this further information? Take (35), repeated below:

(36) A wolf might come through the door. It would eat you.

The second sentence should be true if the domain of 'would' includes just these two worlds: one where wolf A comes in and eats you, one where wolf B comes in and eats you. It should be false if the domain of 'would' instead includes just these two different worlds: one where wolf A comes in and B eats you (perhaps B was already in the room hiding in the corner), and one where wolf B comes in and A eats you. Our semantics needs to discriminate between these two cases, so that the sentence is true in the former case and false in the latter. We need to tie each entity to a world, so that it says in a world where A is the thing that comes in, A is the thing the pronoun must pick out, and in a world where B is the thing that comes in, B is the thing the pronoun must pick out. To achieve this, we move from variable assignments of entities to ones of individual concepts. Since the latter are functions from worlds to entities, they can tie down each entity, and in effect say, 'this pronoun must pick out one of these things at this world, and one of those things at that world', rather than only communicating a set of entities, one of which the pronoun must pick out. The move from entities to individual concepts will only become relevant when we are looking at pronouns within an intensional context that are bound from outside that context. The rest of our semantics will function as if assignments just assigned entities rather than individual concepts.

History shows that when individual concepts are used one needs to take care to avoid making '$\exists x \Box Fx$' equivalent to '$\Box \exists x Fx$' (see Aloni 2001 §2.3). Indeed, the semantics I give below does not provide accurate truth conditions for discourses with an indefinite outside of the scope of any modal word, that then binds a pronoun inside a modal context. I believe that one could adjust the account below to handle these sentences as well, but as it does not concern intentional identity sentences I have not worked it out.

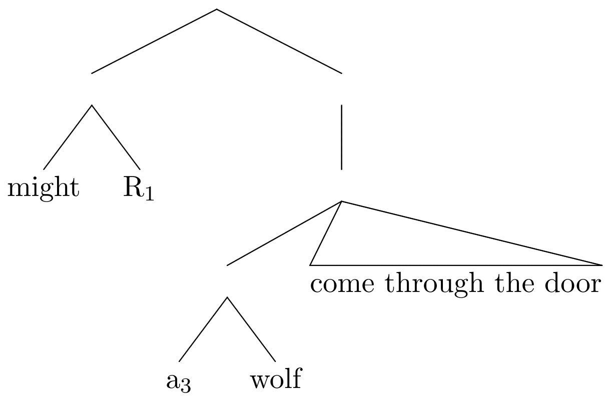
Informally, here is what our semantics will achieve.

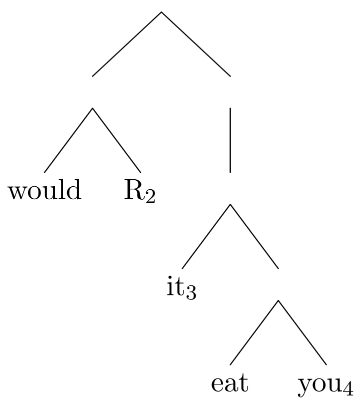
(37) A_3 wolf might-R_1 come through the door. It_3 would-R_2 eat you.

The two modals are associated with accessibility relations R_1 and R_2, which each associate each world with a set of worlds that are quantified over. We will care about which sets are associated with the actual world, since that is the world in which (37) is being evaluated. R_2, because of the continuation, will take the actual world to a subset of those worlds that R_1 does—the R_1 worlds where a wolf does come through the door. Then the modals act like quantifiers—the force of the first being existential, the second being universal. The first sentence says that, at some R_1-accessible world, a wolf comes through the door. The second sentence says that, at every R_2-accessible world, a wolf that comes in eats you. We use individual concepts to compositionally achieve this non c-commanding semantic binding. At the level of individual concepts, the whole discourse will be true iff there is an individual concept c such that: (i) at some R_1-accessible world w, a wolf comes through the door at w, and (ii) at every R_1-accessible world w where a wolf comes through the door, $c(w)$ is a wolf that comes through the door, and (iii) at every R_2-accessible world w, $c(w)$ eats you. Given the plenitude of individual concepts and the relation between R_1 and R_2, these truth conditions reduce to our earlier ones that did not mention individual concepts. We achieve these truth conditions compositionally by using variable assignments. Context will provide a variable assignment (now assigning individual concepts, not entities), and—as long as at least one R_1-accessible world has a wolf coming through the door—the first sentence's semantic value will relate the contextually-provided variable assignment to all those variable assignments that both (i) differ at most in what they assign to 3 (the index of (37)'s indefinite article) and then at most in what entities they assign to 3 relative to R_1-accessible worlds where a wolf comes through the door and (ii) assign to 3 an individual concept assigning, to every R_1-accessible world where a wolf comes through the door, such a wolf. Crucially, the individual concept can assign different things to those different worlds, but at each world, the thing assigned will be a wolf coming through the door at that world. The semantic value of the second sentence will act as a filter, since there are no indefinites in it, and it will let through those variable assignments that, for each R_2-accessible-world, assign to that world an entity that, at that world, eats you. As long as at least one variable assignment gets through, the discourse is true. These truth conditions reduce to the ones involving just individual concepts, which again reduce to the ones that do not mention individual concepts.[5]

Here are the details of our intensional dynamic semantics. We use the following syntax:

[5] I assume that the contextually-provided variable assignment does not assign anything to 3 that might help make the sentence true. If we are going to rely on context to provide an appropriate individual concept, the dynamic semantics binding tools will presumably not be doing any work.





Our only rule is the simpler version of functional application:

Functional Application If α has β and γ as immediate daughters, and $[\![\gamma]\!]$ is in the domain of $[\![\beta]\!]$ then,

$$[\![\alpha]\!] = [\![\beta]\!]([\![\gamma]\!])$$

A sentence with a semantic value p_{vvst} is true at a context providing a variable assignment g iff there is some variable assignment h such that $p(g)(h)$(the actual world) is True. Variable assignments are functions from the set of indices into the set of functions from the set of worlds into the set of entities (except for the assignments for accessibility relation 'R' variables). In other words, if indices were of type i, variable assignments would be of type (ise). '$g[1(w)]h$' means that g and h differ at most in what they assign to 1 relative to world w. In other words, for all numbers n and worlds w', $g(n)(w') \neq h(n)(w')$ only if both $n = 1$ and $w' = w$.

First, here are the semantic values that do not change or are just intensionalized in the obvious way:

$[\![$'it$_3$'$]\!]$ is $\lambda g_v.g(3)$

$[\![$'you$_4$'$]\!]$ is $\lambda g_v.g(4)$

$[\![$'R$_1$'$]\!]$ is $\lambda g_v.g(1)$

$[\![$'R$_2$'$]\!]$ is $\lambda g_v.g(2)$

$[\![$'eat'$]\!]$ is $\lambda b_{vse}.\lambda a_{vse}.\lambda g_v.\lambda h_v.\lambda w_s.\big(g = h$ and $a(g)(w)$ eats $b(g)(w)$ at $w\big)$

$[\![$'come through the door'$]\!]$ is $\lambda a_{vse}.\lambda g_v.\lambda h_v.\lambda w_s.\big(g = h \ \wedge \ a(g)(w)$ comes through the door in $w\big)$

$[\![$'wolf'$]\!]$ is $\lambda a_{vse}.\lambda g_v.\lambda h_v.\lambda w_s.\big(g = h \ \wedge \ a(g)(w)$ is a wolf in $w\big)$

$[\![$';'$]\!]$ is $\lambda q_{vvst}.\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s.\exists i_v\big(p(g)(i)(w) = \text{True} \wedge q(i)(h)(w) = \text{True}\big)$

As for indefinite descriptions, here is the semantic value we want for a whole sentence:

$[\![$'a$_3$ wolf come through the door'$]\!]$ is $\lambda g_v.\lambda h_v.\lambda w_s.\big(g[3(w)]h$ and $h(3)(w)$ is a wolf that comes through the door in $w\big)$

So

$[\![$'a$_3$ wolf'$]\!]$ is $\lambda f_{(vse)vvst}\lambda g_v.\lambda h_v.\lambda w_s.\exists i_v\big(g[3(w)]i$ and $i(3)(w)$ is a wolf in w and $f([\lambda k_v.\lambda w'_s.i(3)(w)])(i)(h)(w)$ is True$\big)$

$[\![$'a$_3$'$]\!]$ is $\lambda e_{(vse)vvst}.\lambda f_{(vse)vvst}\lambda g_v.\lambda h_v.\lambda w_s.\exists i_v\exists j_v\big(g[3(w)]i$ and $e([\lambda k_v.\lambda w'_s.i(3)(w)])(i)(j)(w)$ is True and $f([\lambda k_v.\lambda w'_s.i(3)(w)])(j)(h)(w)$ is True$\big)$

Note that we are not feeding the individual concept assigned to 3 into the semantic value of the predicate combining with 'a$_3$ wolf'. Instead, we are looking at the entity the individual concept assigns to the world of evaluation. A (possibly different) individual concept that takes every world to this entity is then fed into the predicate. This more subtle procedure does not have an effect on this fragment of our semantics, since we have no modal words that change the world of evaluation. Once we introduce modal words, however, this procedure will ensure that indefinites in intensional contexts leave the right sort of information for future pronouns. We now turn to the semantic values of modal words.

Say we have a 'might-R_1' that combines with a sentence containing an indefinite. If there is at least one R_1-accessible world where the sentence is true (more precisely, where the sentence's semantic value takes the variable assignment to at least one variable assignment), then the semantic value of the whole combination is going to take variable assignments to those that differ at most in what they assign to the indefinite at R_1-accessible worlds where the sentence is true, and that, at each of these worlds, assign the indefinite to something satisfying the sentence. For example, $[\![$'might-R_1 a$_3$ wolf is blue'$]\!]$ is going to take variable assignments to those (i) that differ at most in what they assign to 3, and whose individual concepts assigned to 3 differ at most in the entities assigned to the R_1-accessible worlds where a wolf is blue and (ii) that assign to 3 an individual concept such that each entity assigned to an R_1-accessible, blue-wolf-containing world is a blue wolf at that world and (iii) are such that at least one R_1-accessible world has a blue wolf.

To try to make the above explanations of the semantic values more comprehendible, I have not relativized them to the world of evaluation, but that would be done by sticking in 'at the world of evaluation' at various points in the previous sentences. Here is the official semantic value:

⟦'[might-R_1] [a$_3$ wolf is blue]'⟧ is

$\lambda g_v.\lambda h_v.\lambda w_s.\Big($

$(i)g$ and h differ at most in the following: what they assign to 3 at the R_1-accessible-from-w worlds where there is a blue wolf and

(ii) h assigns to 3 an individual concept that assigns to each R_1-accessible-from-w, blue-wolf containing world w' something that is a blue wolf in w' and

(iii) at least one R_1-accessible-from-w world has a blue wolf$\Big)$

Here is that same semantic value in more symbolic terminology:

⟦'[might-R_1] [a$_3$ wolf is blue]'⟧ is

$\lambda g_v.\lambda h_v.\lambda w_s.\Bigg($

(i) $\forall$ indices ι, $\forall {w'}_s\Big($if $g(\iota)(w') \neq h(\iota)(w')$ then $\big(\iota$ is 3, and $R_1(g)(w)(w')$ = True, and there is a blue wolf at $w'\big)\Big)$ and

(ii) $\forall {w'}_s\Big($if $\big(R_1(g)(w)(w')$ = True, and there is a blue wolf at $w'\big)$ then $h(3)(w')$ is a blue wolf at $w'\Big)$ and

(iii) $\exists {w'}_s\Big(R_1(g)(w)(w')$ = True, and there is a blue wolf at $w'\Big)\Bigg)$

Next, I give the general semantic value for 'might-R_1' that allows it to combine with any sentence. The first change from the more specific case is that the sentence might have an indefinite with a different index than

'a_3'. To see which indexed indefinites the sentence contains, the semantic value for 'might' looks at the semantic value p of the sentence. If we put a variable assignment g into p, and we receive some function taking some variable assignment h to some world to True, where h disagrees with g at a particular index, then we know the sentence contains an indefinite with that index. The second change is in how we get the '$h(3)(w')$ is a blue wolf' requirement. If we have some semantic value p of a sentence containing an indefinite, here is how we see if h assigns to 3 an entity that satisfies the sentence's requirements of the indefinite. We check to see if p relates h back to h itself. For example, given some world w, if h assigns to 3 at w an entity that is indeed a blue wolf at w, then $[\![$'A_3 wolf is blue'$]\!](h)(h)(w) = $ True. With these two changes, we have:

$$
\begin{aligned}
&[\![\text{'might-}R_1\text{'}]\!] = \lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s.\Big(\\
&\text{(i) } \forall \text{ indices } \iota, \forall w'_s\big(g(\iota)(w') \neq h(\iota)(w') \rightarrow (R_1(g)(w)(w') = \text{ True } \wedge\\
&\qquad\qquad \exists i_v(p(g)(i)(w') = \text{ True } \wedge g(\iota)(w') \neq i(\iota)(w')))\big) \wedge\\
&\text{(ii) } \forall w'_s\big((R_1(g)(w)(w') = \text{ True } \wedge \exists i_v(p(g)(i)(w') = \text{ True })) \rightarrow\\
&\qquad\qquad p(h)(h)(w') = \text{ True }\big) \wedge\\
&\text{(iii) } \exists w'_s\big(R_1(g)(w)(w') = \text{ True } \wedge \exists i_v(p(g)(i)(w') = \text{ True })\big)\Big)
\end{aligned}
$$

By generalizing to any accessibility relation, we get the semantic value for 'might' (with similar semantic values for other modals with existential force):

$$[\![\text{'might'}]\!] = \lambda R_{vsst}.\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s.\Big($$

$$\text{(i) } \forall \text{ indices } \iota, \forall w'_s\big(g(\iota)(w') \neq h(\iota)(w') \rightarrow (R(g)(w)(w') = \text{ True } \wedge$$

$$\exists i_v(p(g)(i)(w') = \text{ True } \wedge g(\iota)(w') \neq i(\iota)(w')))\big) \wedge$$

$$\text{(ii) } \forall w'_s\big((R(g)(w)(w') = \text{ True } \wedge \exists i_v(p(g)(i)(w') = \text{ True })) \rightarrow$$

$$p(h)(h)(w') = \text{ True }\big) \wedge$$

$$\text{(iii) } \exists w'_s\big(R(g)(w)(w') = \text{ True } \wedge \exists i_v(p(g)(i)(w') = \text{ True })\big)\Big)$$

The semantic value for 'would' (and for other modals with universal force) just changes the last big conjunct. Instead of requiring that at least one R-accessible world is such-and-such, we require that every R-accessible world is such-and-such:

$$[\![\text{'would'}]\!] = \lambda R_{vsst}.\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s.\Big($$

$$\text{(i) } \forall \text{ indices } \iota, \forall w'_s\big(g(\iota)(w') \neq h(\iota)(w') \rightarrow (R(g)(w)(w') = \text{ True } \wedge$$

$$\exists i_v(p(g)(i)(w') = \text{ True } \wedge g(\iota)(w') \neq i(\iota)(w')))\big) \wedge$$

$$\text{(ii) } \forall w'_s\big((R(g)(w)(w') = \text{ True } \wedge \exists i_v(p(g)(i)(w') = \text{ True })) \rightarrow$$

$$p(h)(h)(w') = \text{ True }\big) \wedge$$

$$\text{(iii) } \forall w'_s\big(R(g)(w)(w') = \text{ True } \rightarrow \exists i_v(p(g)(i)(w') = \text{ True })\big)\Big)$$

With these intricate semantic values, we can derive the welcome result that 'A_3 wolf might come through the door; it_3 would eat you_4' is true (given the accessibility relations that our account of continuation described) iff both (i) there is an accessible world where a wolf comes through the door and (ii) at every accessible world where a wolf comes through the door, a wolf that comes through the door eats you.[6] See Appendix A for the derivation of the truth conditions, and see Appendix B for a novel but useful notation of the semantic values and calculations, one that I find significantly improves the speed in which I can grasp and calculate dynamic semantic values. This shows that dynamic semantics can solve the second problem of modal subordination, though again it is not obviously the only way to solve it.

This intensional dynamic semantics can be readily extended to 'believes' by treating it like a universal-force quantifier over worlds given by a different accessibility relation—the one taking each world w to those worlds compatible with the subject's beliefs in w. So 'John believes' will receive the same semantic value as 'would-R$_1$', but where R_1 is replaced by the belief accessibility relation I just mentioned in the last sentence. The semantic value of 'believes' will generalize this to any believer:

$[\![$'believes'$]\!]$ $= \lambda a_{vse}\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s.\Big($

(i) $\forall$ indices $\iota, \forall w'_s\big(g(\iota)(w') \neq h(\iota)(w') \rightarrow (w'$ is compatible with $a(g)(w)$'s beliefs at w $\wedge \exists i_v(p(g)(i)(w') = \text{True} \wedge g(\iota)(w') \neq i(\iota)(w'))))\big) \wedge$

(ii) $\forall w'_s\big((w'$ is compatible with $a(g)(w)'s$ beliefs at w $\wedge \exists i_v(p(g)(i)(w') =$ True$) \rightarrow p(h)(h)(w') =$ True$\big) \wedge$

(iii) $\forall w'_s\big(w'$ is compatible with $a(g)(w)$'s beliefs at w $\rightarrow \exists i_v(p(g)(i)(w') =$ True$))\Big)$

This will predict that 'Hob believes a$_1$ witch came in; *Hob* believes she$_1$ ate you' is true iff at every world compatible with Hob's beliefs, there is a witch that came in and that ate you. With the right sort of example, we can even construct modal subordination continuation examples where

[6]If the dynamic semanticist wanted the stronger truth condition requiring that *every* wolf that comes through the door eats you, this would require a more complicated semantics, and I have not worked out the details.

the 'Hob believes'-part of the second sentence is implicit, and so we need to use a solution to the first problem of modal subordination to introduce quantification over the right worlds into the second sentence. An example is the above (34), repeated below:

(38) According to John's theory, a witch is poisoning the well. She is coming at night and dropping in anthrax bacteria.

For this Hob-*Hob* sentence, we could use an account of modal subordination to derive the truth conditions.

The important thing to note is that the second problem of modal subordination is an easier version of the difficult problem of intentional identity involved in Hob-Nob, intentional identity sentences. The second problem of modal subordination is the same problem we face when we look at Hob-*Hob* sentences like in the previous paragraph—sentences where the subject of each propositional attitude attribution is the same.[7] Since modal subordination involves continuation, the accessibility relation in the second sentence relates the world of evaluation to a (typically proper) subset of worlds related to the world of evaluation by the first sentence's accessibility relation.

When we face the more difficult Hob-Nob cases, however, these accounts will not work. Take 'Hob believes a_1 witch poisoned the well. Nob believes she_1 ate you.' On the extended modal subordination semantics, the first sentence places requirements on what 'she_1' can pick out. At Hob's belief

[7] More accurately, sentences where the possible worlds quantified over by the second intensional context are a subset of the possible worlds quantified over by the first intensional context. We could have cases where the subject stays the same but where this more accurate criterion is not met . For example, one could say the sentences so slowly that the subject has undergone significant mental changes between the first and second sentence, or one could use a propositional attitude verb in the second sentence that is quantifying over new worlds not quantified over in the first sentence (e.g., 'wonders whether' in the second, and 'believes' in the first). These latter cases are as difficult as (G), and are not accurately described by this account of modal subordination straightforwardly extended to 'believes'.

worlds, 'she$_1$' must pick out a witch that poisoned the well. Assuming Nob's belief worlds are a subset of Hob's, this means the discourse will be true iff Hob believes a witch poisoned the well, and Nob believes a witch that poisoned the well ate you. But the assumption is only rarely true. For most cases, Nob's belief worlds will not be a subset of Hob's; in many cases, they will not even overlap. At worlds that are Nob belief worlds but are not Hob belief worlds, the pronoun's assignment is just going to be given by contextually-provided variable assignment, so the dynamic tools are doing no work. The pronoun is, for the most part, functioning like a free-variable-like deictic pronoun (though now it is picking out an individual concept), and the dynamic tools have not helped us give its semantics.

By slightly modifying the values of the modal and propositional attitude verbs, we could predict instead that the discourse is true iff Hob believes a witch poisoned the well, and Nob believes a witch that poisoned the well ate you, regardless of whether Nob's belief worlds are a subset of Hob's.[8] As we have seen, and as Geach (1967) immediately pointed out after describing the intentional identity sentences, these truth conditions are not accurate, because not necessary. Or we could retreat from individual concepts and say that the discourse is true iff Hob believes a witch poisoned the well, and at each of Nob's belief worlds, you were eaten by one of the things that at one of Hob's belief worlds poisoned the well.[9] This would, however, give an inaccurate account of modal subordination (for the reason that forced us to move to individual concepts). Also, it looks similar to an exotica

[8]Where p is the semantic value of a sentence, we would say that $[\![$'might-R_1'$]\!](p)$ takes a variable assignment g to a variable assignment h to a world w to True iff (i) h disagrees with g only on new variables in p and (ii) for every world where there is a variable assignment that p relates g to, p takes h to itself, and (iii) there is at least one R_1-accessible-from-w world for which p takes g to some variable assignment.

[9]To do this we would move back to our old variable assignments assigning entities to indices, and then make the small adjustments that this would require to the semantic values of the modal and propositional attitude verbs.

account, and it would have some of the same problems that I discuss in chapter 2, for example, Oneski-Twoski sentences. So this would be giving up an accurate account of one phenomenon to gain an inaccurate account of another phenomenon.

In general, here is why the intentional identity pronouns do not seem to be functioning like dynamically-bound pronouns. A dynamic semantics can make the pronoun bound by material that does not c-command it, which allows us to require that, at Nob worlds, the thing(s) the pronoun picks out satisfies certain predicates that previously appear in the linguistic material around the binder. To do this, we need to write up more complicated semantic values for the previous linguistic material, that lay down in advance precise formulas for constructing the predicates that the pronoun appearing later will need to satisfy. Oneski-Twoski sentences, however, seem to need a different formula than the apartment sentences, and the predicates that role sentences seem to need do not seem to appear in the previous linguistic material at all. As for the original Geach sentence (G) and the causal connection cases, the predicates not only do not seem to appear in the previous linguistic material, but we do not even know what the predicates would be. Consequently, it does not look like these pronouns are functioning simply as ones dynamically bound by previous material.

The problem is not just with intentional identity sentences involving propositional attitude or intensional transitive verbs. Just like modal subordination phenomena can occur in propositional-attitude-like contexts such as (38), the more-difficult-than-modal-subordination-problem-two intentional identity phenomena can occur in modal contexts, and we discussed some examples in chapter 1. For example, imagine an elderly woman wondering how her life could have been different:

(39) If I had married Bobby, I could have had a daughter with red hair. Or if I had married Presley, she could have had blonde hair.

In this case, the worlds being quantified over in the second sentence are not a subset of the worlds being quantified over in the first sentence. Our dynamic semantics would not handle these cases either. And the second sentence does not mean that it could have been that a daughter with red hair has blonde hair. It also does not mean that something that was—at one of the marrying-Bobby-worlds—a daughter with red hair is also—at one of the marrying-Presley-worlds—a daughter with blonde hair. Therefore, neither of our two contemplated revisions in the last paragraph will deliver acceptable truth conditions for these sentences.

Whatever the solution is to intentional identity, it looks like it will solve the easier second problem of modal subordination. One consequence of this is that, though dynamic semantics solves the second problem of modal subordination, this is no intrinsic point in its favor, since it alone cannot solve the more general intentional identity problem, and since the correct solution looks like it will solve the second problem of modal subordination. The main conclusion of this section, however, is that dynamic tools extended in the obvious way, though they handle modal subordination, cannot handle the harder cases of intentional identity.

3.4 Guise theories

In the remaining sections of this chapter, I look at theories that use entities I call 'guises'. I argue that the currently proposed theories are incorrect, and that this approach to intentional identity will probably not prove helpful.

As briefly mentioned earlier, most work in dynamic semantics can be roughly sorted into two different groups. The first group does not take itself to be heavily modeling actual human mental states or linguistic processing. Instead, this work is mostly trying to model truth conditions and semantic properties of bits of language, like most non-dynamic formal semantics, and the tools it would typically need are just the more complicated semantic values in the previous section that allow non c-command binding. Much of this work is done by Groenendijk, Stokhof, and people using their Dynamic Predicate Logic and Dynamic Montague Grammar (Groenendijk and Stokhof 1990, 1991).

On the other hand, there is a second body of work that involves trying to construct accurate models of mental activity. Much of this work is done by Kamp and people using his Discourse Representation Theory (Kamp 1981, Kamp and Reyle 1993). Just because work is done with DRT does not mean it is in this second group—a lot of the semantic models described using DRT tools are merely meant to predict truth conditions and other semantic properties of bits of language, rather than model mental states or linguistic processing. The second body of work goes beyond that, using DRT tools not just to model semantic properties, but also to model mental states. This work tends to use notions from DRT—e.g., DRT's discourse referents and DRSs—that I have not discussed in the earlier sections, because they are not necessary to achieve non c-command binding. Semanticists working in this tradition have tried to use these models of mental states to give an account of intentional identity sentences (Asher 1987, Kamp 1990, Cumming 2007, 2013), and I discuss accounts of this kind here. On these accounts, people have something like mental files or mentalese names, and intentional identity requires Hob and Nob, or whoever the two subjects are, to have mental files or tokens of mentalese names that correspond in some way, both being of a certain type or bearing a certain property.

I group these into a more general class of intentional identity accounts, guise theories. These theories use what I will call *guises*, a vague category meant to include things like senses, modes of presentation, and other entities that are more fine-grained than Russellian semantic values yet perform similar semantic work. Guise theories of intentional identity are those that use guises in a central way to explain the semantics of intentional identity sentences. The most obvious way would be to posit some guises, and then say that typical intentional identity sentences require that there is a single guise that is involved in both of the subject's beliefs in an appropriate way.

Guise theories will typically involve two parts. First, there will be an account of guises. Second, there will be the semantics that employs those guises. These two parts are relatively independent of each other—different guise accounts could be fitted with different semantics—and there is no reason why the semantics needs to be of a dynamic sort. I include guise theories in this chapter, however, because every currently-endorsed dynamic account of intentional identity that I have seen uses guises, because every guise account I have seen has been proposed by a dynamic semanticist, and because their conceptions of guises and of the accompanying semantics tend to appropriate ideas from the dynamic tradition.[10]

Here is how I am thinking any plausible guise theory will go. It is not as if the typical Hob-Nob sentence is true iff there is some guise such that Hob believes it is such-and-such and Nob believes it is what-not. Hob can believe a witch is cackling without believing that a mode of presentation is cackling. (Or, if this is how the guise theory goes, then really it is an

[10](Burge 1983) is a possible exception. I have trouble understanding the central technical concepts of his account ('Applications' and 'notions'), however, so I am unsure what exactly his account is.

exotica theory, and my thoughts on those are in chapter 2.) Instead, the quantification over guises will be more subtle. One would say something like, the Hob-Nob sentence is true iff Hob believes a witch is such-and-such, and Nob believes something is what-not, and there is some guise that plays the right role in both Hob's belief and Nob's belief. In many guise stories, guises will be involved in beliefs in such a way that one can believe that x under guise g is F, and the guise story will spell out what it means to believe that something under a certain guise is such-and-such. Then, one will probably say that the typical Hob-Nob sentence is true iff there is a guise g such that Hob believes that a witch under guise g is such-and-such, and Nob believes that something under guise g is what-not. This will be the role the guise will need to play in Hob and Nob's beliefs. Then one will try to extend this account to all intentional identity discourses.

The general property uniting guise accounts is that they are things that are related to Russellian semantic values, but are more fine-grained. The Russellian semantic values of names are objects, and those of verbs are properties. Of course, in at least one good sense of 'object', everything is an object. Guises are more fine-grained in the sense that each object may have multiple guises of that object, but each guise is the guise of only one object. There are guise stories where guises might be the guise of no object at all (and presumably such guises will be the ones involved in Geach's (G)), but even on such stories they are typically the exception, not the rule. So given the sentence 'Bono sings', the Russellian semantic value of 'Bono' is an object, namely Bono. The semantic value of 'sings' is the property of singing. On a guise account, there are multiple guises of Bono. What are these guises? Different guise accounts will give different answers. The guises are meant to fill some role in semantic theorizing, and so entities will be chosen that seem to fit the role. For example, they might be properties corresponding to descriptions—that of being the singer of U2 might be one, that of being the guy I just saw on the street might be another.

In general, the motivation for introducing guises is often one that would lead you to postulate guises not just for the Russellian semantic values of names, but for the Russellian semantic values of most, if not all, kinds of words or morphemes. For example, guises are often introduced to allow two names bearing the same Russellian semantic value—e.g., ‘Hesperus’ and ‘Phosphorous’—to have different semantic properties. This would, among other things, be useful in allowing ‘John believes Hesperus is magical’ to be true without ‘John believes Phosphorous is magical’ being true. Similar sentence pairs can be constructed for just about any kind of word, however. Perhaps the English ‘if... then’ construction really is a sentential connector with the same Russellian semantic value as the logician's English ‘$\rightarrow$’. Nevertheless, if one thought John can believe that Hesperus is magical without believing that Phosphorous is magical, one should think that a mistaken philosopher can believe that grass is blue $\rightarrow$ grass is not blue, without believing that if grass is blue then grass is not blue. Consequently, most guise accounts should have multiple guises not just of Bono, but of singing, of the Russellian semantic value of ‘every’, and so on for perhaps every kind of word.

Below, I evaluate guise theories. First, I present the most detailed guise theory in the literature, that of (Cumming 2013). I compare it to the other guise theories (Asher 1987, Kamp 1990), and discuss why I think it is incorrect. Then I evaluate guise theories in general.

3.5 Cumming's (2013) account

Cumming's (2013) account can be divided into two parts: the guise part, and the semantic part. I focus on (Cumming 2013) because it is the most worked-out guise account I have seen published.

The guises are called 'thought objects' (TOs). TOs are 'of' or 'denote' ordinary objects (though of course there can be thought objects that are of other thought objects too). For example, there is the Parthenon, but there are also thought objects that are of the Parthenon. These thought objects are very different than the Parthenon. The Parthenon itself is in Greece, is made out of stone, has a large mass, etc., and none of these properties are true of the thought objects of the Parthenon. There are also thought objects that do not denote anything, and it is these thought objects that will feature in typical intentional identity sentences.

Cumming also posits 'descriptive mental files' (MFs). MFs are private—each MF is the MF of some particular thinker. My MFs are distinct from your MFs. MFs always 'denote' a TO, and TOs are public, in the sense that you and I can have distinct MFs that denote the same TO. MFs also 'refer' to ordinary objects. So for example I might have an MF that refers to the Parthenon, and that denotes a thought object that is of the Parthenon. An MF is basically a collection of properties or 'conditions', which I will pick out by using lambda notation. One of these properties is the special 'rubric condition', and the others are 'non-rubric conditions'. MFs only refer to things that satisfy their rubric condition.[11] For example, I might have an MF with the rubric condition [$\lambda x.x$ is an ancient temple on the Acropolis]

[11] Cumming leaves it open whether it is always the case that rubric conditions uniquely pick out the thing (if such there be) that the MF is of. Also, one might expect me to have a descriptive MF for each description that I believe is uniquely satisfied, but Cumming denies this.

and the non-rubric conditions [$\lambda x.x$ is made out of stone] and [$\lambda x.x$ is heavy]. This MF may refer to the Parthenon. You may have another MF with the same rubric condition that also refers to the Parthenon. Since MFs are private, we cannot have the same MF referring to the Parthenon, but we can have MFs that denote the same TO. While some MFs do not refer to anything, they always denote a TO. So, for example, Le Verrier had an MF with the rubric condition [$\lambda x.x$ is a planet between Mercury and the Sun]. This MF did not refer to anything, but it still must have denoted a TO. Note that TOs are not Salmon's fictional objects. I do not believe that my Parthenon-denoting thought idea is made out of stone; I only believe the Parthenon is made out of stone. And Le Verrier did not have the false belief that the TO denoted by his MF was in between Mercury and the Sun.

I find the following diagram helpful in keeping track of the basic ontology and technical language:

$$\begin{tikzcd}
\text{Mental Files} \arrow[r, "\text{denote}"] \arrow[d, "\text{refer to}"'] & \text{Thought Objects} \arrow[dl, "\text{denote}"'] \\
\text{Ordinary Objects} &
\end{tikzcd}$$

In sum, we have three kinds of things: MFs, TOs, and ordinary objects (though of course these kinds are not mutually exclusive, for example, we could have MFs that refer to MFs, and so 'ordinary object' is a bit of a misnomer). MFs 'refer' to ordinary objects and 'denote' TOs. TOs are 'of' or 'denote' ordinary objects. The referring relation is a *partial* function from (the set of) MFs to (the set of) ordinary objects; it is not one-one (i.e., not injective), and not onto (i.e., not surjective), assuming there are some ordinary objects that no thinker has an MF referring to.[12] The MF-denoting

[12] All this talk of functions and sets is merely for helpful explication; there is probably no set of all objects capable of being referred to by an MF.

relation is a *total* function from MFs to TOs. It is not one-one; Cumming leaves it open whether it is onto. The of-relation is a partial function from TOs to ordinary objects. Cumming leaves it open whether it is one-one (though if TOs were like typical guises it would not be) or onto.

When do MFs denote the same TO? One person's distinct MFs can pick out different TOs even if the person believes that the two MFs refer to the same ordinary object, i.e., even if each MF has the condition that it is identical to the referent of the other one. Also, MFs of two different people can pick out different TOs even if the MFs have the same rubric condition. Cumming does not give complete necessary and sufficient conditions for co-denotation, but he does give a sufficient condition for MF co-denotation: two distinct people's MFs pick out the same TO if the two people agree that the MFs co-refer (Cumming 2013 p.3). Importantly, this sufficient condition does not hold for one person within herself. As noted above, one person can believe that two of her MFs co-refer without the two MFs co-denoting.[13]

This sufficient condition entails that the MFs of people connected by chains of agreement also co-refer. For example, if A has MF1, B has MF2, and C has MF3, then, if A and B agree that their two MFs co-refer, and B and C agree that their two MFs co-refer, then by the sufficiency condition, MF1 and MF2 co-denote, and MF2 and MF3 co-denote, and so MF1 and MF3 co-denote, even if A and C have never met, and so are not familiar with each other's MFs and so do not agree that they co-refer. As you can see, TOs are effectively dispensable, and they could be replaced by partitions of MFs—dividing MFs into groups that all co-denote.

[13]Cumming cites his paper (Cumming Forthcoming) for a more detailed theory of content coordination. I do not discuss (Cumming Forthcoming) because I am not sure how to fit it with the account of (Cumming 2013), and because what (Cumming 2013) says about the subject is enough for my argument in the incorrectness of his current account, and in the difficulty of fixing it.

Cumming's account focuses on correctly describing the Oneski-Twoski case from chapter 2, repeated below:

> *Oneski-Twoski:* Person A has killed himself, but when the two detectives arrive on the scene, they think he has been murdered. Later, Person B kills himself, and when the two detectives arrive there, Detective Oneski thinks that the same person that murdered A murdered B. Detective Twoski disagrees, and thinks a different person murdered B.

Here is Cumming's analysis of the mental states of Oneski and Twoski. Each of their MFs is displayed below as a box, with the rubric condition appearing at the top, and the non-rubric condition at the bottom.

Twoski:

$\lambda x.x$ murdered A	$\lambda x.x$ murdered B
$\lambda x.x$ did not murder B	$\lambda x.x$ did not murder A

Oneski:

$\lambda x.x$ murdered A	$\lambda x.x$ murdered B
$\lambda x.x$ murdered B	$\lambda x.x$ murdered A

Call Twoski's first MF, the one on the upper left, 'TA' (since it is Twoski's and its rubric condition concerns murdering A). Similarly, call Twoski's upper-right, 'TB', Oneski's lower-left 'OA', Oneski's lower-right 'OB'. Note that Oneski has two MFs, even though he thinks the two MFs co-refer.

The pattern of co-denotation is as follows: TA and OA, the two boxes on the left, co-denote; and TB and OB, the two boxes on the right, co-denote. There are no other co-denotations.[14] None of the MFs refer to anything,

[14] Here is what seems to be the general idea of why Cumming thinks these co-denotation facts arise from the agreement sufficiency condition. Say Twoski says to himself 'I'm going to call the murderer of A 'TA', and I'm going to call the murderer of B 'TB'. Hmm. TA and TB are not the same person.' Say also that Oneski says to himself, 'I'm going to

and none of the TOs are of anything, since the two dead people were not actually murdered.

Here is the semantic part of Cumming's theory. The semantic story is not fully compositional; Cumming focuses on giving truth conditions for intentional identity sentences, though he does occasionally discuss what the semantic values of some of the sub-sentential bits are. The key idea is that intentional identity requires the two subjects to have MFs that denote the same TO, i.e., that are co-denoting.

Take:

(40) Twoski believes someone murdered A

Cumming argues that (40) sometimes has two readings. First, there is the 'non-suspect' reading that means the same as 'Twoski thinks A was murdered'. The non-suspect reading roughly corresponds to what others call the de dicto, unspecific, and notional readings of an indefinite. Second, there is the 'suspect' reading, that requires Twoski to have a somewhat worked-out conception of 'the murderer', rather than merely thinking that A was murdered. This roughly corresponds to what others have called the de re, specific, and relational readings of an indefinite, though unlike some of those other conceptions, it does not require there to genuinely be someone that is suspected by Twoski. Both readings require Twoski to have certain MFs, but the differences in the readings are captured by the differences between rubric and non-rubric conditions. Also, a caveat: Cumming believes that sometimes only one of these readings is available (cf. his fn. 29), though he

call the murderer of A 'OA' and the murderer of B 'OB'. Aha! OA and OB are the same person.' Now say that when Oneski and Twoski are discussing the deaths, they share these privately coined names with each other. Oneski and Twoski will agree with the sentence 'TA and OA are the same person, and TB and OB are the same person'. However, though Oneski will agree with the sentence 'TA and OB are the same person', Twoski will not. So the sufficiency condition is not met for TA and OB, and likewise it is not met for TB and OA.

does not give an account of when this happens.

The non-suspect reading of (40) is true iff there is an x such that (i) x is an MF of Twoski's and (ii) x has [$\lambda y.y$ murdered A] as a *rubric* condition. The suspect reading of (40) is true iff there is an x such that (i) x is an MF of Twoski's and (ii) x has [$\lambda y.y$ murdered A] as a *non-rubric* condition. These analyses are meant to generalize to all sentences of the form ⌜H believes someone is G⌝.[15]

Given the underspecification in the account of when MFs are created, the account of non-suspect readings seems less a prediction about new cases, and more just a definition of when MFs are created. As long as you believe that A was murdered, you automatically have an MF with the rubric condition [$\lambda y.y$ murdered A]. The prediction about new cases comes when we add the semantic account of later sentences with pronouns anaphorically linked to the previous indefinites. A new sentence 'Oneski believes he is G', where the pronoun 'he' is anaphorically linked to the indefinite in (40), is true iff (i) Oneski has an MF with [$\lambda y.y$ is G] as a condition, and (ii) this MF co-denotes with at least one of Twoski's MFs that satisfy the existentially quantified statement in the truth conditions of (40). Condition (ii) is the hallmark of intentional identity—Oneski's MF must denote the same TO as at least one of Twoski's MFs that made (40) true.

So, here are the general truth conditions that we arrive at:

[15]When Cumming discusses sentences with a non-suspect reading of ⌜an F⌝ in place of 'someone', he adds a complication to the analysis so that not all of the predication in the sentence has to appear in the rubric condition. The new truth condition: ⌜H believes an F is G⌝ on the non-suspect reading is true iff there is an x such that (i) x is an MF of H's, (ii) x has [$\lambda y.y$ is F and G] as a condition, and (iii) x has a rubric condition that either matches or is entailed by [$\lambda y.y$ is F and G]. (Of course, if every condition entails itself, the 'matches' part is redundant.) I ignore this complication because it does not affect my discussion of Cumming, and because he does not offer an analysis of the sentence when the indefinite has a suspect reading, and so since I do not know how he would propose to treat the $\overline{\mathrm{N}}$ material of suspect readings of indefinite NPs, I restrict my attention to 'someone'. Also note that Cumming treats 'someone' as one would imagine 'something' would be, presumably to avoid unnecessary complexity.

(41) H believes someone_1 is F. N believes she_1 is G.

Non-suspect reading of 'someone': (41) is true iff there is an x and a y such that (i) x is one of H's MFs, (ii) x has $[\lambda z.z$ is F$]$ as a *rubric* condition, (iii) y is one of N's MFs, (iv) y has $[\lambda z.z$ is G$]$ as a condition, and (v) x and y denote the same TO.

Suspect reading of 'someone': (41) is true iff there is an x and a y such that (i) x is one of H's MFs, (ii) x has $[\lambda z.z$ is F$]$ as a *non-rubric* condition, (iii) y is one of N's MFs, (iv) y has $[\lambda z.z$ is G$]$ as a condition, and (v) x and y denote the same TO.

Note that the only difference in the two readings is whether the Hob-sentence condition must be a rubric or a non-rubric condition.

Here again are the Oneski and Twoski sentences:

(42) Twoski thinks someone murdered B, and Oneski thinks he murdered A.

(43) Oneski thinks someone murdered A, and Twoski thinks he murdered B.

Intuitively, (42) is true and (43) is false. Given the MFs of Oneski and Twoski, and the co-denotations between them, the account does indeed deliver these truth-values to the *non*-suspect readings of (42) and (43). (42) is true because of the MFs TB and OB—they have the right conditions and they both co-denote. (43) is false because, to satisfy the existential quantified statement in the truth conditions, 'x' would need to be assigned to OA, since that is Oneski's MF that has $[\lambda y.y$ murdered A$]$ as a rubric condition. But TA is the only MF of Twoski's that OA co-denotes with, and TA does not have the condition [$\lambda y.y$ murdered B].[16]

This in sum is the most detailed guise account of intentional identity sentences. The crucial features are shared with the other accounts of Asher (1987) and Kamp (1990, see also Bende-Farkas and Kamp 2001): we have private mental objects in the minds of the subjects, and distinct private mental objects in two different minds can be of the same public type. Intentional identity requires that the private objects in each of the two subjects are of the same type.

3.6 Objections to the account

In this section, I step through each of the difficult kinds of sentences raised in chapter 2, and argue that the (Cumming 2013) account does not correctly handle them. First, the Oneski-Twoski sentences. As the case is currently spelled out, the account delivers the right verdict on (43), but if we change the scenario slightly, it no longer does.

> *Oneski Jr.*: Everything is exactly like in the Oneski-Twoski scenario, except that Oneski has a partner Oneski Jr., who consults with and agrees with Oneski that a single person murdered both A and B.

Oneski Jr. will have similar MFs as Oneski, with the same conditions. So Oneski Jr. has the following two MFs: one ('JA') with the rubric condition [$\lambda x.x$ murdered A] and the non-rubric condition [$\lambda x.x$ murdered B]; one ('JB') with the rubric condition [$\lambda x.x$ murdered B] and the non-rubric condition [$\lambda x.x$ murdered A]:

[16]On the *suspect* readings, Cumming's account predicts that (42) and (43) are both true.

Oneski Jr.:		
	$\lambda x.x$ murdered A	$\lambda x.x$ murdered B
	$\lambda x.x$ murdered B	$\lambda x.x$ murdered A

Since two MFs of two different people co-denote if the two people agree that they co-refer, and since Oneski and Oneski Jr. consult and agree that a single person murdered both A and B, JA and OA co-denote, JA and OB co-denote, JB and OA co-denote, and JB and OB co-denote. But then, because co-denoting is an equivalence relation, since OA co-denotes with TA and OB co-denotes with TB, all of Oneski, Twoski, and Oneski Jr.'s MFs co-denote with every other MF. So, (43) on both readings is predicted to be true. The non-suspect reading is true, for the existentially quantified statement of the truth condition is satisfied if 'x' is assigned to OA and 'y' is assigned to TB. And the suspect reading is true, if 'x' is assigned to OB and 'y' is assigned to TB. Yet there is clearly a reading (the most prominent reading), where (43) on this scenario is false. So Cumming's account is incorrect.

More generally, Cumming's account of Oneski-Twoski sentences only works because of the following aspect of it: if two distinct people agree that their MFs co-refer, then they co-denote, and intentional identity can be true, but if a single person thinks that two of her MFs co-refer, they do not necessarily co-denote. It is this distinction between the two-person and the one-person case that allows the account to deliver the right verdict on the Oneski-Twoski sentences, but this distinction can be gotten around by pairing the single person with a partner who agrees, which now turns their situation into a two-person one. I do not see how Cumming's account can be fixed, especially given that here it needs to be more restrictive in allowing co-denoting, whereas in the cases below it will need to be more accommodating in allowing co-denoting. This pull in different directions seems hard to accommodate at the same time.[17]

What about Geach's original intentional identity sentence (G)? In one way of spelling out the scenario, there need be no chains of agreement from Hob to Nob. For example, say they each read the local newspaper's witch article alone right after they wake. No third party knows anything about their mental files created in reading the article, and so no third party agrees that those mental files co-denote with any of their own mental files. So the sufficiency account given for co-denotation is not met. So the account is not specific enough to predict the truth-value of (G).

The role sentences and the causal connection cases from chapter 2 have the same result. One can spell them out with no chains of agreement between the subjects. Something interesting can be seen if we consider Scenario 2 of the causal connection case—the scenario where the intentional identity sentence is false because there is no connection. We can specify in this scenario that Hob and Nob would indeed come to believe that their mental files co-referred if they were to meet. Despite this, the intentional identity sentence remains false, because there is no actual causal connection. This shows that we cannot improve Cumming's account by adding as a sufficient

[17]What about an account where two MFs having the same rubric condition is both necessary and sufficient for co-denotation? This effectively turns the account into one where an intentional identity sentence ⌜Hob believes someone is G, and Nob believes she is H⌝ is true on the suspect reading when there is some description ⌜the D⌝ such that ⌜Hob believes the D is G and Nob believes the D is H⌝ is true. For one seems to have an MF with the rubric condition $[\lambda x.x \text{ is a } D]$ iff one believes that someone is a D, and then if one believes that the D is G, this MF will have $[\lambda x.x \text{ is } G]$ as a condition. I believe these truth conditions are much too permissive—many of our previous cases would be counterexamples. For example, take chapter 2's Causal Connection Scenario 2. Hob and Nob will each have MFs with at least the following conditions:

$\lambda x.x$ is a witch wreaking havoc
$\lambda x.x$ is a witch in town

$\lambda x.x$ is a witch in town
$\lambda x.x$ is wreaking havoc

But then the account will incorrectly predict that there is no false reading of 'Hob believes someone is a witch in town, and Nob believes she is wreaking havoc'. For on the suspect reading it is true, since Hob and Nob's MFs on the left co-denote; and on the non-suspect reading it is true, since Hob and Nob's MFs on the right co-denote.

condition for co-denoting that, if the two subjects were to meet, they would agree that their mental files co-referred. Such an account would wrongly predict that the intentional identity sentence in Scenario 2 is true. This also shows that any account of counterfactual coordination (i.e., that the subjects would be able to coordinate their mental activities in some way if certain things were to happen) is probably going to fail. Also, note again that to deal with these problem cases we need to spell out how to allow more co-denotation than we currently have, whereas with Oneski-Twoski we need to figure out how to limit the co-denotation we already have.

I think that some of the apartment sentences from chapter 2 are counterexamples to this theory of anaphoric indefinites, for e.g., if Hob wants to marry a German, it does not seem that he has to have a descriptive mental file in his head with a rubric condition like [λy. I want to marry y]. However, these counterexamples involve intensional transitive verbs and modals, and Cumming specifically mentions that his theory is not meant to apply to sentences where indefinites are within the scope of those, so I merely note that the account will not cover those sentences.

3.7 Evaluating all guise theories

Of course, the much more general task is not to see whether the particular account in (Cumming 2013) is correct, but whether any guise theory is correct. The main question a guise account needs to tackle is, what are the right conditions for co-denotation or guise-sharing that will deliver the right verdict for the intentional identity sentences? Getting just one of the above kinds of sentences correct seems difficult, and getting the right verdict on all of them at the same time seems much harder.

In general, whatever the correct account of intentional identity sentences is, there will presumably be some way that Hob and Nob's minds and the world outside their minds has to be for the intentional identity sentence to be true. Whatever this way is, it is plausible we could postulate guise-like entities, and then rephrase the account in guise-talk, by rephrasing the account in terms of requirements to be thinking under the same guise. So I think it is quite plausible that there is a correct account using guise-language. The above objections seem to show, however, that no currently proposed guise account is correct, and that modifying them in order to make them correct looks very difficult.

Starting out by trying to find a guise account of intentional identity by asking, 'what are the conditions under which two people share the same guise?' seems to be a very similar project to just asking, 'what are the conditions under which two people are 'thinking about the same thing', even if technically there is no such thing?'. Because of this, I worry that thinking in terms of guises does not help much in trying to find the correct account of intentional identity sentences.

Chapter 4

Description Theories

In this chapter I set out what I believe is the correct account of intentional identity. First, I shift focus from intentional identity sentences with pronouns to sentences with definite descriptions. I propose a semantics for such sentences, and then extend it to intentional identity sentences with pronouns and other NPs. Finally, I compare it with other accounts and discuss areas for further work.

4.1 Intentional identity with definite descriptions

There is a striking fact that apparently has never been stated in the intentional identity literature. All intentional identity sentences, including Geach's (G), have analogues with definite descriptions in place of pronouns, and these analogues seem practically identical in meaning and truth conditions:

(G) Hob thinks a witch has blighted Bob's mare, and Nob wonders whether she killed Cob's sow.

(G′) Hob thinks a witch has blighted Bob's mare, and Nob wonders whether the witch killed Cob's sow.

I cannot find a context of utterance and a circumstance of evaluation where these two sentences clearly diverge in truth-value.[1] This fact suggests an obvious research strategy. We need a semantic account of definite description sentences like (G′) as much as of pronoun sentences like (G). Since the semantics of definite descriptions is a topic that has benefitted from much work, it seems promising to try and first apply the standard semantics of definite descriptions to sentences like (G′), and then to postulate the same truth conditions for the pronoun sentences like (G). I call this kind of theory a 'description theory' of intentional identity—one where the pronouns in (G) and corresponding sentences function semantically like definite descriptions have traditionally been taken to function.[2,3]

This postulation receives further motivation from other areas of semantics. Many people have proposed that certain pronouns in non-intentional-identity sentences function semantically like traditional descriptions.[4] When we combine this postulation with an account of how sentences like (G′) function, we have an account of how sentences like (G) function.

[1] The only possible exception I can see is where Nob does not believe that the thing is a witch. My intuitions are quite weak in these contexts, which is why I include the qualifier 'clearly' in the footnoted sentence above. If one's intuitions are stronger, one could replace (G′)'s 'the witch' with 'the thing', while putting the sentence in a context where it's clear that 'the thing' is anaphorically linked to 'a witch'.

[2] Elbourne's (2005) account seems to entail that (G′) and (G) have the same truth conditions, though he does not discuss intentional identity, and his account does not entail what these truth conditions might be.

[3] I believe that there are intentional identity sentences with names as well—'Hob thinks a witch named 'Gertie' has blighted Bob's mare, and Nob wonders whether Gertie killed Cob's sow'—though it is controversial whether these sentences have anaphoric links as in (G). These sentences arguably show that there is a link between the problem of intentional identity and that of empty names. Hawthorne and Manley (2012) suggest applying semantic accounts of empty names to intentional identity sentences, but I believe that significantly more progress has been made on the semantics of definite descriptions than that of empty names, and so I think that we should first see if definite description semantics suggest a correct account of (G′) and thus of (G).

[4] For example: Quine (1956), Evans (1977, 1980)—though Evans' pronouns in some ways do not function like descriptions, Parsons (1978), Cooper (1979), Davies (1981), Neale (1990), Heim (1990), Cherchia (1995), and Elbourne (2005).

We need a semantic account of sentences like (G′), and clearly the anaphoric NP in (G′) functions semantically like a definite description, since it *is* a definite description. So there must be an account of intentional identity sentences like (G′) where the anaphoric NPs function semantically like definite descriptions. Given that these sentences do not seem to have any significant semantic differences from pronoun-involving sentences like (G), and given the plausibility that at least some pronouns function semantically like definite descriptions, an account of intentional identity sentences like (G) seems to follow almost for free. Thus, if we conservatively assume that descriptions function as traditionally thought, a correct description theory seems to me almost inevitable. In this chapter, I propose an account of how sentences like (G′) work, and then extend it to intentional identity sentences where definite descriptions are replaced by pronouns and other NPs.

4.2 The semantics of definite descriptions

In this section, I first classify the definite description in (G′) as an 'incomplete' description and explain what that is. Then I discuss the general semantics of definite descriptions, including incomplete ones.

I will call an occurrence of a definite description ⌜the N'⌝ in a certain sentence at a context *incomplete* iff (a) the sentence at that context is felicitous and yet (b) the definite description does not trigger a presupposition that there is (unrestrictedly speaking) exactly one N'. Here is an example of a *complete* definite description: 'The best philosopher in the world just published another book' seems to take for granted that there is exactly one best philosopher in the world. In general, many definite descriptions seem to give rise to a presupposition that there is exactly one thing satisfying the semantic value of their N′. Other definite descriptions, however, clearly seem not to give rise to that particular presupposition. Take (while in a lecture), 'The lecturer is drunk'. This carries no suggestion that there is (unrestrictedly speaking) exactly one lecturer. Here 'the lecturer' is incomplete.

In clause (b) of the definition of 'incomplete description', I use the word 'trigger', because some phrases in simple sentences can trigger certain presuppositions on their own, even if those presuppositions change or disappear when the simple sentences are embedded in larger passages. For example, 'wonders whether' is typically taken to be a presuppositional 'transformer' or 'filter'. If you take a sentence S that on its own presupposes that it is the case that P, and then you append S to $\ulcorner X$ wonders whether$\urcorner$, the whole larger sentence does not presuppose that it is the case that P, but rather that X believes that it is the case that P (Karttunen 1974). (G$'$) does not presuppose that Nob believes there is (unrestrictedly speaking) exactly one witch. So, the definite description 'the witch' in (G$'$) did not trigger a presupposition that there is (unrestrictedly speaking) exactly one N'; for if it had, the 'Nob wonders whether' would have transformed it into a presupposition that Nob did indeed believe there is exactly one witch. So 'the witch' in (G$'$) is at this context an *incomplete* description. So the correct semantic account of (G$'$) will depend on that of incomplete definite descriptions.

There are two traditional semantic theories of definite descriptions: Fregean and Russellian.[5,6] I wish to stay neutral between these theories. On the Fregean account of definite descriptions, their contribution to the truth conditions of simple, non-intensional sentences is a thing. The N$'$ material

[5]On the semantics of definite descriptions, see (Heim 1991), (Heim and Kratzer 1998), (Elbourne 2005). The main salient rival to these two theories would be a theory that uses dynamic tools in giving the semantics of definite descriptions, but because of the arguments in chapter 3 I do not think those tools would be of help here.

[6]I focus on singular definite descriptions, for I assume that plural definite descriptions raise no unique puzzles of intentional identity, and that a correct account of singular descriptions will readily extend to intentional identity with plural descriptions when a correct account of plural descriptions is found.

plays some role in picking out the thing that is the contribution—the semantic value—of the definite description. Incomplete definite descriptions show that the N′ material does not solely determine the semantic value, since different occurrences of ‘the lecturer’ can have different semantic values, even if the things that are lecturers remain the same. Some aspect of the context helps to determine which lecturer is the semantic value of such an occurrence. When we complicate our semantics to handle intensional contexts, the contribution of definite descriptions is instead a (possibly partial) function from possible worlds to things. Incomplete definite descriptions show that context must then play some role in which such function is the semantic value.

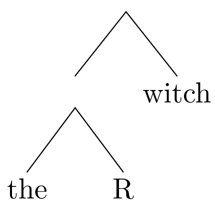
On the Russellian account of definite descriptions, ‘the’ functions semantically like a quantifier, explicitly restricted by the explicit material in the N′. Most semanticists agree that, in particular contexts, quantification can also be implicitly restricted by restrictions not explicitly found in N′, and the most popular Russellian account of definite descriptions says that the quantification involved in ‘the’ can be similarly restricted by restrictions not explicitly found in N′. This is the popular Russellian account of incomplete definite descriptions (e.g., Neale 1990; Stanley and Szabo 2000).[7]

Looking at the second attitude ascription of (G′) on its own—‘Nob wonders whether the witch killed Cob’s sow’—we are already in a position to apply these theories of incomplete descriptions to ‘the witch’. On either theory, context affects the semantic value of the incomplete description. The challenge is then to identify the semantic value of the incomplete description

[7] There are also theories of incomplete definite descriptions where the semantic value of an incomplete definite description is not contextually-dependent, but where the context instead only affects the situations described by the sentence. Barwise and Perry (1983) give an influential theory of this kind. Kratzer (2011) provides an overview of these types of theories, which I henceforth do not discuss.

at a context where the previous utterance is a sentence like, 'Hob thinks a witch has blighted Bob's mare'. There are different views on the details of how context affects incomplete descriptions, and thus on how to model such interactions.[8] I will abstract away from these details by using a simple toy model of incomplete definite descriptions. This is not a proposal for the actual semantics of incomplete descriptions; I use it merely to make my intentional identity account more explicit. Whatever the actually correct account of incomplete descriptions is, it should then be easy to translate my proposal into the framework that is ultimately found to be correct.

In the toy model, I will suppose that the English 'the' is the Fregean 'the'. I will also suppose that, in the syntactic structure given to the semantics, each 'the' is adjoined with a variable of the same semantic type as that of common nouns like 'witch'. This variable will be assigned different semantic values at different contexts, which is how we will model the effect of context on the semantic value of instances of incomplete definite descriptions. So, with 'R' as one of our variables, 'the witch' is, in the syntactic structure given to semantics:



As in our semantics from chapter 1, context provides a variable assignment. Variable assignments will take variables like 'R' to semantic values of the same type as that of 'witch'. The semantic values of definite descriptions are (possibly partial) functions from worlds to entities. We use the same composition rules as chapter 1, and here are some semantic values. For all

[8] Stanley and Szabo (2000) and Elbourne (2008) discuss a wide range of these views.

variable assignments g,

$$
\begin{aligned}
[\![\text{‘witch’}]\!]^g &= \lambda x_{\text{se}}.\lambda w_{\text{s}}.x(w) \text{ is a witch in } w \\
[\![\text{‘R’}]\!]^g &= g(\text{‘R’}) \\
[\![\text{‘the’}]\!]^g &= \lambda f_{(\text{se})\text{st}}.\lambda g_{(\text{se})\text{st}}.\text{the unique thing } x \text{ such that} \\
&\quad\; f([\lambda w'.x])(w) = g([\lambda w'.x])(w) = \text{True}
\end{aligned}
$$

For example, if at a particular context, an incomplete definite description ‘the witch’ has the same semantic value as the complete definite description ‘the witch that crushed Dorothy’, then the contextually provided variable assignment can assign ‘R’ the same semantic value as that of ‘crushed Dorothy’, i.e., $[\lambda x_{\text{se}}.\lambda w_{\text{s}}.x(w)$ crushed Dorothy in $w]$. With our values and rules, the semantic value of ‘the R witch’ at this context will be the same as that of ‘the witch that crushed Dorothy’.

So, to give an account of (G′), we want to know the semantic value of the incomplete ‘the witch’ at certain contexts, and we can work this out by trying different semantic values, and testing them against our knowledge of the truth-values of the sentence in particular scenarios.

4.3 Easier cases

Before we get to (G′), I would first like to look at cases from earlier chapters that are easier to handle. First, here are definite-description analogues of our apartment sentences from chapter 2:

(44) Hob wants to buy an apartment. Nob wants the apartment to be near the city centre.

(45) Hob believes he will buy an apartment. Nob believes the apartment will be near the city centre.

Here the full value of 'the apartment' is plausibly just that of 'the apartment that Hob buys'. So 'R' can be assigned $[\lambda x_{\mathrm{se}}.\lambda w_{\mathrm{s}}.\text{Hob buys } x(w) \text{ in } w]$. It looks promising to suggest that, in general, many contexts assign to 'R' semantic values like those above—ones for the most part that are not identical to any from the first sentence, but that could be constructed from values in the first sentence.

Similarly, here are definite-description analogues of the Oneski-Twoski sentences and the role sentences from chapter 2:

(46) Twoski thinks a guy killed B, and Oneski thinks the guy killed A.

(47) Hob believes a god poisoned the well, and Nob believes the god killed the sow.

At the contexts in chapter 2, the full semantic value of 'the guy' can be that of 'the guy that killed B', and the full semantic value of 'a god' can be 'the wickedest god'. Note how context-dependent the full semantic values of the incomplete definite descriptions are. In different contexts, these sentences could have their incomplete definite descriptions pick out completely different semantic values.[9]

Here are definite-description analogues of our earlier modal sentences:

(48) A wolf might come through the door. The wolf would eat you.

(49) If I use the MegaLure, I'll catch a fish big enough to eat. If I use the

[9]Edelberg (1986) endorses his Example 7 as a counterexample to specific description theories of intentional identity sentences with pronouns, but Edelberg's Example 7 has a definite description analogue ('Detective A thinks some person murdered the mayor, and Detective B thinks the person murdered the commissioner'), just as (G) has as an analogue (G′), so there must be some account of what the incomplete definite description in the analogue of Edelberg's counterexample means. I think his Example 7 and similar ones are like role sentences and their contexts. 'the person' in the definite description analogue of Edelberg's (26) can have the same semantic value as that of 'the person who shot Smith', and 'the person' in the analogue of his (27) can have that of 'the person who shot Jones'. So I disagree with (Glick 2012)'s assertion that Edelberg's example 'convincingly demonstrate[s]' that all definite description theories are incorrect.

MiniLure, the fish won't be big enough.

Because of the first sentence in (48), 'The wolf would eat you' is plausibly somehow restricted to quantifying over situations where a wolf comes through the door, meaning something like 'If a wolf came through the door, it would be the case that the wolf eats you'. If so, then the full value of 'the wolf' is plausibly just that of 'the wolf that comes through the door'. Similarly, the full value of 'the fish' in (49) seems to be just the value of 'the fish I catch'. Again, this is a semantic value mostly constructed from values in the first sentence.

4.4 Harder cases

So far, assigning semantic values to particular incomplete definite descriptions has been easy. When we turn to (G′) and the causal connection cases from chapter 2, however, it becomes more difficult. Geach pointed out that the sorts of complete descriptions we have used so far will not work for (G). They also will not work for (G′). (G′) does not mean the same thing as 'Hob thinks a witch has blighted Bob's mare, and Nob wonders whether the witch that blighted Bob's mare killed Cob's sow', for in our scenario Nob does not necessarily have any beliefs about Bob's mare. Similarly, (G′)'s incomplete 'the witch' does not have the same value as something like 'the witch that Hob thinks blighted Bob's mare', since Nob need not have any beliefs about Hob. In general, there is no obvious semantic value for the incomplete 'the witch' in (G′) that delivers the right truth conditions.

With the causal connection scenarios from chapter 2 in mind, one might propose that the full value of 'the witch' in (G′) is that of 'the witch described by the newspaper', since the newspaper is where, in the context, Hob and Nob both heard about the witch. This is better, but it does not get the intension of the sentence right. To illustrate this, take a slightly more complicated case where the speaker is unsure of the connection between the subjects. Say that a speaker sees Hob looking behind the television, and Nob looking in the cabinets, and she speculates that they are both searching for the same fairy. The speaker does not know exactly how their beliefs might be connected (did they read the same newspaper article? did they hear about a fairy from Hob's mother?), but the speaker speculates that they are indeed connected (their searching appears to be coordinated, they both keep muttering on about 'catching the fairy'), and says:

(50) Hob believes a fairy is behind the television, and Nob believes the fairy is in the cabinet.

It seems that this sentence at this context could be true in many different ways. Perhaps Hob and Nob have both read in the same newspaper about a fairy, and they each have the required beliefs about where it is. Or perhaps Hob heard about a fairy from his mother, who also told Nob about it, and now they are performing a coordinated search, starting in the places where they believe it is. There are many other scenarios in which (50) seems to be true. But now exactly which fuller descriptions correspond to 'the fairy' in this sentence at this context?

In general, the causal connection scenarios suggest that there are salient readings of intentional identity sentences that require some sort of communication chain between the subjects. In Causal Connection Scenario 1 (§2.3), there are instances of communication that connect Hob and Nob to each other.[10] Al and Bob communicate with each other, then Al communicates

[10]Burge (1983) notes that causal links between Hob and Nob are important, though I am unsure exactly how important he thought them (see the above footnote 10 in chapter 3). Also, Devitt (1981) suggests that intentional identity requires that the names the subjects

with Hob, and Nob communicates with Bob. Al and Bob are links in the communication chain connecting Hob to Nob. As with a causal story of the passing-down of the meanings of names, the requirement of a communication chain will need many qualifications. In general, however, it seems that we want to give (50) an intension where, at worlds where the communication link from Nob to Hob is a newspaper, Nob must believe that the fairy described by the newspaper is in the cabinet. At worlds where the communication link is Hob's mother, Nob must believe that the fairy described by Hob's mother is in the cabinet. There is indeed a technical tool that can deliver these truth conditions—the philosopher's 'actual'. With this tool, 'the fairy' could correspond to 'the fairy described by the actual previous link in the communication chain from Nob to Hob'.

The word 'actual' in the technical sense can best be explained with the possible worlds semantics we have been using. In this semantics, 'John believes that the length of my yacht is thirty feet' is true iff, at every possible world compatible with John's beliefs, the length of my yacht is thirty feet. When we use the word 'actual' in a belief attribution, it can shift the world of evaluation from the possible worlds we are quantifying over back to the actual world. For example, 'John believes that the length of my yacht is longer than the actual length of my yacht' has a reading that is true iff, at every possible world compatible with John's beliefs, the length of my

would use to express their propositional attitudes must be related causally, though the details of this proposal are not worked out. Edelberg (1992) fleshes out Devitt's proposal before offering two counterexamples. I agree that the contexts of Edelberg's counterexamples do not require communication chains between the subjects, but as with (Edelberg 1986)'s examples, there are incomplete-definite-description analogues to Edelberg's counterexample sentences, and there must be *some* semantic value that we can assign to the incomplete descriptions (cf. fn. 9 above). I think Edelberg's counterexamples to Devitt are like our role sentences above, and the incomplete descriptions have full values similar to some of the phrases Edelberg uses in describing the mental states of the people in the contexts of the scenarios, e.g., that of 'the driver of the car' and 'the overdensity drawing the supercluster'.

yacht is longer than whatever length my yacht is at the actual world. In other words, it is true iff the actual length of my yacht L is such that, John believes that the length of my yacht is longer than L.

In (50), 'the fairy' can correspond to 'the fairy described by the actual previous link in the communication chain from Nob to Hob'. Here I am intending the reading of 'actual' where (50) is true at a context in world w_1 iff (i) for each of Hob's belief worlds w_2, something that is a fairy at w_2 is behind the television at w_2 and, (ii) for each of Nob's belief worlds w_3, the fairy in w_3—the one that is described in w_3 by what is, in w_1 the previous link in the communication chain from Nob to Hob—is in the cabinet in w_3.[11] So, at a context in world w_1, 'R' can be assigned the function $[\lambda x_{\mathrm{se}}.\lambda w_{\mathrm{s}}.x(w)$ is described in w by what is, in w_1, the previous link in the communication chain from Nob to Hob$]$. This is equivalent to saying that (50) is true at this context iff (i) Hob believes a fairy is behind the television, and (ii) the previous link L in the communication chain from Nob to Hob is such that Nob believes the fairy described by L is in the cabinet.

In many contexts 'the previous link in the communication chain from Nob to Hob' will admittedly be an incomplete definite description as well. There might be, for example, several previous links from Nob to Hob, even if we restrict our attention to chains about fairies. Perhaps they read one story about fairies in the newspaper, hear a different story about fairies from Hob's mother, and the latter prompts them to search for what Hob's mother was talking about. In either case, (50) should be true, and yet 'the previous

[11] Soames (1998 13–17) and Williamson (2000 293–4) argue that many propositional attitude ascriptions containing 'actual' require that the subject of the ascription has *de re* attitudes about the actual world—attitudes that would have been very difficult to have if the actual world had not obtained. For those who accept their picture of 'actual', my complete description containing 'actual' does not accurately communicate the truth conditions I am intending for these sentences, and they should focus instead on this more detailed gloss not containing 'actual'.

link in the communication chain between Hob and Nob' will fail to uniquely pick out something. Here one would need a more detailed description to more fully describe the truth conditions. The rough shape of the semantic account, however, will be the same, and giving this more detailed description seems to involve general issues about incomplete definite descriptions, rather than issues specific to intentional identity sentences.

With (50), the world where there must be a communication chain is the world in which the intentional identity sentence is uttered. By embedding intentional identity sentences in modal operators, however, we can see that the important world is the evaluation world of the belief ascriptions, not necessarily the evaluation world of the whole sentence. For example, say that the speaker thinks about various possible scenarios of how Hob's and Nob's beliefs could be connected—a newspaper, Hob's mother—and is unsure which is true, but knows that at least one of them is true, and so, given what the speaker knows:

(51) It must be the case that: Hob believes a fairy is behind the television, and Nob believes the fairy is in the cabinet.

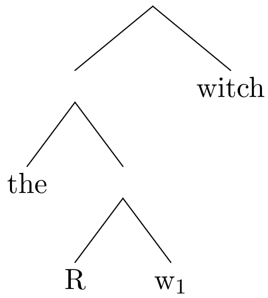
Here, on the standard semantics for 'it must be the case that', the matrix clause—the whole sentence—involves quantification over possible worlds, some where Hob and Nob both read the same newspaper article, and some where they both hear Hob's mother talking about a fairy. In a newspaper world, Hob and Nob are connected by a different chain than they are in a Hob's mother world. There is no single communication chain that can be appealed to in giving the description. The full semantic value of 'the fairy' at this context cannot be that of 'the fairy described by the actual previous link'. The 'actual' would take us back to the world of evaluation of the whole sentence, when instead we just want to go back to the world of evaluation of the belief ascriptions.

Here is a way to get the truth conditions we want. Some semanticists posit slightly complicated syntactic structures to handle sentences like 'Every cop took a photo with a certain crook'. It seems like such sentences can, at the right contexts, mean something like 'Every cop took a photo with a certain crook—the favorite crook of the cop', where each cop might be taking a photo with a different crook, but each cop takes a photo with that cop's favorite crook.[12] On this account, at this context, 'a certain crook' is, at the level of semantic analysis, 'a certain [R x] crook', where 'R' and 'x' are both variables. 'Every cop' effectively binds the variable 'x', and then the value of 'R' is that of 'favorite crook of'. When combined with the variable 'x', the value of '[R x]' is that of 'favorite crook of x'. The whole sentence then means, at this context: every cop x is such that, x took a photo with the favorite crook of x.

We can do something similar with these intentional identity sentences. The quantification introduced by 'it must be the case that' might bind some element in the implicit material in 'the fairy' in such a way that the sentence is true iff at every world w_1 compatible with what the speaker knows, (i) for every Hob belief world w_2, there is something that is a fairy in w_2 and is behind the television in w_2, and (ii) for every Nob belief world w_3, the fairy at w_3—the one described at w_3 by what is, at w_1, the previous link in the communication chain from Nob to Hob—is in the cabinet at w_3.

To achieve this, we can postulate that the 'R' variable is adjoined by a variable 'w_1', so we have the following structure:

[12] Stanley and Szabo (2000) propose an influential theory of such cases with quantifier phrases and provide citations of earlier discussions. Schwarzschild (2003) discusses cases involving indefinite descriptions.



'[R w_1]' is assigned the value of 'described by what is, at world w_1, the previous link in the communication chain from Nob to Hob', and 'w_1' is bound by the quantification across worlds involved in 'it must be the case that'. In detail: $[\![$'w_1'$]\!]^g$ is $\mathrm{g}(\mathrm{w}_1)$, which will let the quantification introduced by 'it must be the case that' bind it, and at this context 'R' is assigned the semantic value $[\lambda w_{\mathrm{s}}.\lambda x_{\mathrm{se}}.\lambda w'_s.x(w')$ is described at w' by what is, at w, the previous link in the communication chain from Nob to Hob].[13] Consequently, (51) is, at this context, true iff, for every world w_1 compatible with what the speaker knows, (i) for every Hob-at-w_1 belief world w_2, something is a fairy at w_2 and is behind the television at w_2, and (ii) the previous link L in the communication chain from Nob to Hob at w_1 is such that, for every Nob-at-w_1 belief world w_3, the fairy in w_3 described in w_3 by L is in the cabinet in w_3.

The previous link in newspaper worlds is the newspaper, and in Hob's-mother worlds it is the mother. So, (51) is true. In general, this semantic account seems to predict the correct truth judgments of sentences like (51) in the contexts in which we were evaluating them.[14] This proposal also cor-

[13]See appendix C for details of the proposal in a world-variable framework like that in (von Fintel and Heim 2011, c. 8).

[14]For uniformity considerations, I propose that these are also the sorts of syntactic structures involved in sentences and contexts like the non-modally-embedded Geach ones. In cases like (G′) where there is no modal operator under which the intentional identity sentences are embedded, the 'w' variable in the implicit restrictor is assigned to the matrix clause world of evaluation. This is what would naturally happen in, e.g., von Fintel and

rectly predicts that the sentence in Scenario 2 of the causal connection case is not true. If there is no communication chain in a particular world of evaluation of the propositional attitude clauses, then the incomplete description 'the witch' will not be true of anything at any of the belief worlds, and the matrix sentence will not be true.[15]

There is, however, still a remaining causal connection difficulty. The true intentional identity sentence in chapter 2's Scenario 1 stays true even if Nob loses his belief of where he heard about the witch. Our account seems to wrongly predict that the sentence is now no longer true, because Nob will have some belief worlds where there is no witch described by Bob—the actual previous communication link. I believe that here we should be liberal about what it takes for some *de re* ascriptions of belief to be true, and say that Bob can be such that Nob still believes the witch described by him is wreaking havoc, for example, because Nob always retains the trivial belief that the witch described by whoever he actually heard about the witch from is wreaking havoc.[16]

Heim's framework (2011, c. 8). In effect, this gives the description the semantic value of 'the witch described by the *actual* common communication link', as we proposed earlier.

[15]Whereas on the Russellian 'the', the sentence would be false, on the Fregean 'the', the most we can say right now is that the sentence is not true, though one can flesh out the Fregean story so that sentences with definite descriptions that fail to pick something out are indeed false (see Heim and Kratzer 1998 §4.4.2). There is nothing unique about intentional identity here; Fregean accounts of anaphoric definite descriptions anaphorically linked to non-intensional-context indefinites would have the same issue. Take 'A guy came to our house last night; John believes the guy was staking the place out' and say that the first sentence is true, but the second sentence is apparently false because John has no beliefs about any guy coming to the house, or any guy staking (or about the actual guy who, unbeknownst to John, came to the house); in this case, the Fregean story must be that the incomplete definite description 'the guy' fails to pick out anything at some of John's belief worlds, which is why the sentence is not true. The Fregean will also need an explanation of why the sentence is not only not true, but also apparently false. The same account should be applied by the Fregean to these intentional identity sentences.

[16]Hawthorne and Manley (2012) give arguments that support being liberal about *de re* thought in general. The trivial belief Nob has might be different in different scenarios. For example, if Nob believes he was really the first to describe the witch to others, he will always retain the trivial belief that the witch described by any speakers he actually informed is wreaking havoc.

Here is one final problem. What if Nob believes there are multiple things described by the actual previous link? For example, if Hob and Nob both read about a fairy in the *Times*, and then have the right beliefs, (50) will be true. If Nob also believes that there is another fairy that the *Times* at one point described, however, then 'the fairy described by the actual previous link' no longer uniquely picks out something in each of Nob's belief worlds, and so our account fails to accurately identify (50) as true. Here I think it is important to pay attention to which particular objects are the links in the chain. Potential links can be 'broad' or 'narrow'; for example, here are some potential links in increasing order of narrowness: the media, the print media, the *Times*, this particular daily issue of the *Times*, this particular article in this particular daily issue of the *Times*. I believe that these cases show that we want to use narrow links in our account. If the particular article in the particular daily that Hob and Nob both read is such that Nob truly believes that there are multiple fairies described by it, then (50) strikes me as not clearly true, with perhaps an unmet presupposition, as our account would predict.

In sum, there appear to be at least two kinds of contextually-provided semantic values for the incomplete definite descriptions: first, those that are for the most part made up from semantic values appearing in the first sentence or in the surrounding context; second, those that pick out, at each world, the target (if such there be) of the previous link in the communication chain from the propositional attitude clause world of evaluation.

4.5 Other NPs

As discussed earlier, it has frequently been proposed that at least some pronouns function like definite descriptions. If we accept that (G′) and (G) have the same meaning, then we can assume that 'she' has the same semantic value as 'the witch' (as it would in the account of (Elbourne 2005)), with the same possibilities for context supplementing the incomplete value.[17] This proposal will extend our account to cover pronoun NPs.

As discussed in chapter 1, intentional identity can also occur with other NPs besides pronouns and definite descriptions. For example, take chapter 1's sentence (12) and its scenario, repeated below:

> Hob and Nob both read the local newspaper saying fifty witches from the west came to town, and they believe what they read. Furthermore, Nob thinks that every witch that the newspaper wrote about flew in on a broomstick, though he thinks a few witches from the east also came to town without using broomsticks. In actuality, the newspaper's article is false, and there are no witches.

(52) Hob thinks fifty witches came to town, and Nob thinks that every witch flew in on a broomstick.

There is a large consensus that incomplete definite descriptions and implicitly restricted quantification phrases function very similarly to each other, though whether they function identically to each other is more controversial (Larson and Segal 1995, 329ff.). Mechanisms similar to our toy model of

[17]In Elbourne's account, the pronoun has the same value as the determiner 'the' adjoined with the topmost N′ of the antecedent NP (I ignore his indexes that allow for binding). In some cases of intentional identity, it might be more accurate just to adjoin 'the' with the lowest N′ of the antecedent NP. That should be acceptable even if the key parts of Elbourne's account are correct, since his mechanism for the construction, N′ deletion (which he calls 'NP deletion' since he accepts the Abney DP hypothesis) can function similarly: 'The department's been downsized, but not drastically. All philosophers that were working on aesthetics were fired, but few were working on aesthetics'. What about cases like (G) but where Nob does not believe the thing is a witch? Again, N′ deletion exhibits similar behavior: 'Hob believes many jumpers looked angry as they fell. Nob believes some didn't jump, but were pushed.' So Elbourne's mechanism would work here as well, and this is a question about N′ deletion, rather than intentional identity.

incomplete definite descriptions are often used to model implicit restrictions of quantifier restrictions (Stanley and Szabo 2000), and if these are on the right track, then our model can easily be extended to these NPs as well. In (52), 'every witch' will be supplemented with a variable like our 'R' above, and it will receive similar values as it did earlier in the chapter.

4.6 Conclusion

With this description theory of intentional identity, we can give what look like correct truth conditions to all of our examples of anaphora without appealing to mythical objects, guises, or dynamic semantics. Nevertheless, there are still questions remaining. Perhaps most prominent is the metasemantic question: why do some contexts provide these sorts of semantic values rather than others? This might seem more pressing given that, by positing bindable world variables in the syntax, there are more kinds of values context could assign. A general account of bindable world variables should provide some constraints on how they can function (von Fintel and Heim 2011, c. 8). Richer knowledge of the relation between context and the semantic value of incomplete definite descriptions in non-intentional-identity sentences should also constrain the semantic values. I think one can find cases of anaphoric incomplete definite descriptions in non-intentional-identity sentences that also have values as complicated as our intentional identity ones, and that ultimately the main reason behind these values lies in the meaning of the general anaphoric link itself, but there is clearly more work to be done here.

Given chapter 3, dynamic tools simply cannot wholly accommodate all of the instances of intentional identity. The best currently developed guise theories likewise do not accurately handle the data, and the prospects for using the concept of guises looks unpromising. As chapter 2 shows, using exotic objects at first looks promising, but exotica theories struggle with accounting for Oneski-Twoski, role, and apartment sentences, while our description theory without exotica handles these sentences readily. One could try a description theory that mimics ours in the easier cases of §4.3, but that uses exotica instead of communication chains in the harder cases of §4.4. Such a theory, however, still faces the objection to the core exotica idea in §2.6. It falsely predicts that intentional identity antecedent indefinites should be able to link anaphorically with extensional pronouns, when they clearly cannot. In contrast, the communication chain restrictions offer a clear reason why this cannot happen: if one follows (G) with a pronoun 'she' in an extensional context, it will try to pick out the thing described by the actual previous link in the communication chain, but there is no such thing, so we have a failed definite description. For these reasons, I believe that the account in this chapter is superior to the other accounts that have been proposed in the literature.

But of course, the goal is not just to find an account more accurate than any previously proposed, but to find a *correct* account. If there is in fact good evidence against this chapter's account that I have not been able to find, the conclusions of §4.1 above remain. There must be some correct description theory of (G), and if this chapter has failed to find it, we should continue to search.

Appendix A

Modal subordination truth condition calculation

By the definition of 'is true', where g is the variable assignment provided by the context and where $w_@$ is the actual world, 'A_3 wolf-R_1 might come through the door; it_3 would-R_2 eat you_4' is true iff:

$$\exists h\big(⟦\text{'A}_3\text{ wolf might-}R_1\text{ come through the door; it}_3\text{ would-}R_2\text{ eat you}_4\text{'}⟧ \\ (g)(h)(w_@) = \mathrm{T}\big)$$

By functional application, our syntax, and our semantic values:

$$\begin{aligned} & ⟦\text{'A}_3\text{ wolf might-}R_1\text{ come through the door; it}_3\text{ would-}R_2\text{ eat you}_4\text{'}⟧ \\ = \quad & ⟦\text{';'}⟧(⟦\text{'it}_3\text{ would-}R_2\text{ eat you}_4\text{'}⟧)(⟦\text{'A}_3\text{ wolf might-}R_1\text{ come through the door'}⟧) \end{aligned}$$

First I calculate the semantic value for the 'it_3' sentence, and then for the

'A_3' sentence. By our semantic values, functional application, and our syntax:

$$
\begin{aligned}
& [\![\text{'it}_3 \text{ would-}R_2 \text{ eat you}_4\text{'}]\!] \\
= \quad & [\![\text{'would'}]\!]([\![\text{'R}_2\text{'}]\!])([\![\text{'eat'}]\!]([\![\text{'you}_4\text{'}]\!])([\![\text{'it}_3\text{'}]\!]))
\end{aligned}
$$

$[\![\text{'eat'}]\!]([\![\text{'you}_4\text{'}]\!])([\![\text{'it}_3\text{'}]\!])$ is:

$$
\begin{aligned}
& [\lambda b_{vse}.\lambda a_{vse}.\lambda g_v.\lambda h_v.\lambda w_s.(g = h \wedge a(g)(w) \text{ eats } b(g)(w) \text{ at } w)] \\
& ([\lambda g_v.\lambda w_s.g(4)(w)])([\lambda g_v.\lambda w_s.g(3)(w)]) \\
= \quad & [\lambda g_v.\lambda h_v.\lambda w_s.(g = h \wedge g(3)(w) \text{ eats } g(4)(w) \text{ at } w)]
\end{aligned}
$$

And ⟦'would'⟧(⟦'R$_2$'⟧) is:

$$
\begin{aligned}
& [\lambda R_{vvst}.\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s. \\
& \text{(i) } \forall \text{ indices } \iota\ \forall w'_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(R(g)(w)(w') = \mathrm{T}\ \wedge \\
& \qquad\qquad\qquad \exists i_v(p(g)(i)(w') = \mathrm{T} \wedge g(\iota)(w') \neq i(\iota)(w'))\big)\Big) \wedge \\
& \text{(ii) } \forall w'_s\Big(\big(R(g)(w)(w') = \mathrm{T} \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\big) \rightarrow p(h)(h)(w') = \mathrm{T}\Big) \wedge \\
& \text{(iii) } \forall w'_s\Big(R(g)(w)(w') = \mathrm{T} \rightarrow \exists i_v(p(g)(i)(w') = \mathrm{T})\Big)] \\
& ([\lambda g_v.\lambda w_s.\lambda w'_s.w \text{ is related to } w' \text{ by } g\text{'s } R_2]) \\
= \ & [\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s. \\
& \text{(i) } \forall \text{ indices } \iota\ \forall w'_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \wedge \\
& \qquad\qquad\qquad \exists i_v(p(g)(i)(w') = \mathrm{T} \wedge g(\iota)(w') \neq i(\iota)(w'))\big)\Big) \wedge \\
& \text{(ii) } \forall w'_s\Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\big) \rightarrow \\
& \qquad\qquad\qquad p(h)(h)(w') = \mathrm{T}\Big) \wedge \\
& \text{(iii) } \forall w'_s\Big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \rightarrow \exists i_v(p(g)(i)(w') = \mathrm{T})\Big)]
\end{aligned}
$$

So $[\![$'it$_3$ would-R_2 eat you$_4$'$]\!]$ is:

$$
\begin{aligned}
& [\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s. \\
& \text{(i) } \forall \text{ indices } \iota\ \forall {w'}_s \Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \wedge \\
& \qquad\qquad\qquad \exists i_v (p(g)(i)(w') = \mathrm{T} \wedge g(\iota)(w') \neq i(\iota)(w'))) \Big) \wedge \\
& \text{(ii) } \forall {w'}_s \Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \wedge \exists i_v (p(g)(i)(w') = \mathrm{T}) \big) \rightarrow \\
& \qquad\qquad\qquad p(h)(h)(w') = \mathrm{T} \Big) \wedge \\
& \text{(iii) } \forall {w'}_s \Big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \rightarrow \exists i_v (p(g)(i)(w') = \mathrm{T}) \Big)] \\
& ([\lambda g_v.\lambda h_v.\lambda w_s.(g = h \wedge g(3)(w) \text{ eats } g(4)(w) \text{ at } w)]) \\
= \ & [\lambda g_v.\lambda h_v.\lambda w_s. \\
& \text{(i) } \forall \text{ indices } \iota\ \forall {w'}_s \Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \wedge \\
& \qquad\qquad \exists i_v (g = i \wedge g(3)(w') \text{ eats } g(4)(w') \text{ at } w' \wedge g(\iota)(w') \neq i(\iota)(w'))) \Big) \wedge \\
& \text{(ii) } \forall {w'}_s \Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \wedge g(3)(w') \text{ eats } g(4)(w') \text{ at } w' \big) \rightarrow \\
& \qquad\qquad h(3)(w') \text{ eats } h(4)(w') \text{ at } w' \Big) \wedge \\
& \text{(iii) } \forall {w'}_s \Big(w \text{ is related to } w' \text{ by } g\text{'s } R_2 \rightarrow g(3)(w') \text{ eats } g(4)(w') \text{ at } w' \Big)]
\end{aligned}
$$

Here we have the three universally-quantified conjuncts: (i), (ii), and (iii). The consequent of (i) requires both that $g = i$ and that g and i differ, so the consequent is never satisfied. So the antecedent is false, and $g = h$. The conjunct (ii) is vacuously true since $g = h$. So this reduces to:

$$
\begin{aligned}
& [\lambda g_v.\lambda h_v.\lambda w_s.(g = h \wedge \forall {w'}_s (w \text{ is related to } w' \text{ by } g\text{'s } R_2 \rightarrow \\
& \qquad\qquad\qquad g(3)(w') \text{ eats } g(4)(w') \text{ at } w'))]
\end{aligned}
$$

Now we calculate the semantic value of the 'A_3' sentence.

$$\begin{aligned}
& [\![\text{'A}_3 \text{ wolf might-}R_1 \text{ come through the door'}]\!] \\
= \; & [\![\text{'might'}]\!]([\![\text{'}R_1\text{'}]\!])([\![\text{'a'}]\!]([\![\text{'wolf'}]\!])([\![\text{'come through the door'}]\!]))
\end{aligned}$$

$[\![\text{'a'}]\!]([\![\text{'wolf'}]\!])([\![\text{'come through the door'}]\!])$ is:

$$\begin{aligned}
& [\lambda e_{(vse)vvst}.\lambda f_{(vse)vvst}.\lambda g_v.\lambda h_v.\lambda w_s.\exists i_v \exists j_v (g[3(w)]i \wedge \\
& e([\lambda k_v.\lambda w'_s.i(3)(w)])(i)(j)(w) = \text{T} \wedge f([\lambda k_v.\lambda w'_s.i(3)(w)])(j)(h)(w) = \text{T})] \\
& (\lambda a_{vse}.\lambda g_v.\lambda h_v.\lambda w_s.(g = h \wedge a(g)(w) \text{ is a wolf at } w)) \\
& (\lambda a_{vse}.\lambda g_v.\lambda h_v.\lambda w_s.(g = h \wedge a(g)(w) \text{ comes through the door at } w)) \\
= \; & [\lambda g_v.\lambda h_v.\lambda w_s.\exists i_v \exists j_v (g[3(w)]i \; \wedge i = j \wedge i(3)(w) \text{ is a wolf at } w \wedge \\
& j = h \wedge i(3)(w) \text{ comes through the door at } w)] \\
= \; & [\lambda g_v.\lambda h_v.\lambda w_s.(g[3(w)]h \wedge \\
& h(3)(w) \text{is a wolf at } w \wedge h(3)(w) \text{ comes through the door at } w)]
\end{aligned}$$

And ⟦'might'⟧(⟦'R_1'⟧) is:

$$[\lambda R_{vvst}.\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s.$$

$$\text{(i) } \forall \text{ indices } \iota\ \forall {w'}_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(R(g)(w)(w') = \mathrm{T} \wedge$$

$$\exists i_v(p(g)(i)(w') = \mathrm{T} \wedge g(\iota)(w') \neq i(\iota)(w'))\big)\Big)\wedge$$

$$\text{(ii) } \forall {w'}_s\Big(\big(R(g)(w)(w') = \mathrm{T} \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\big) \rightarrow p(h)(h)(w') = \mathrm{T}\Big)\wedge$$

$$\text{(iii) } \exists {w'}_s\Big(R(g)(w)(w') = \mathrm{T} \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\Big)]$$

$$([\lambda g_v.\lambda w_s.\lambda {w'}_s.w \text{ is related to } w' \text{ by } g\text{'s } R_1])$$

$$= [\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s.$$

$$\text{(i) } \forall \text{ indices } \iota\ \forall {w'}_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge$$

$$\exists i_v(p(g)(i)(w') = \mathrm{T} \wedge g(\iota)(w') \neq i(\iota)(w'))\big)\Big)\wedge$$

$$\text{(ii) } \forall {w'}_s\Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\big) \rightarrow p(h)(h)(w') = \mathrm{T}\Big)\wedge$$

$$\text{(iii) } \exists {w'}_s\Big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\Big)]$$

So $[\![$'A$_3$ wolf might-R_1 come through the door'$]\!]$ is:

$$
\begin{aligned}
& [\lambda p_{vvst}.\lambda g_v.\lambda h_v.\lambda w_s. \\
& \text{(i) } \forall \text{ indices } \iota\ \forall w'_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \\
& \qquad\qquad\qquad \exists i_v(p(g)(i)(w') = \mathrm{T} \wedge g(\iota)(w') \neq i(\iota)(w'))\big)\Big) \wedge \\
& \text{(ii) } \forall w'_s\Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\big) \rightarrow \\
& \qquad\qquad\qquad p(h)(h)(w') = \mathrm{T}\Big) \wedge \\
& \text{(iii) } \exists w'_s\Big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \exists i_v(p(g)(i)(w') = \mathrm{T})\Big)] \\
& ([\lambda g_v.\lambda h_v.\lambda w_s.(g[3(w)]h \wedge h(3)(w) \text{is a wolf at } w \wedge \\
& \qquad h(3)(w) \text{ comes through the door at } w)]) \\
= \quad & [\lambda g_v.\lambda h_v.\lambda w_s. \\
& \text{(i) } \forall \text{ indices } \iota\ \forall w'_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \\
& \qquad\qquad\qquad \iota = 3 \wedge \text{there is a blue wolf at } w'\big)\Big) \wedge \\
& \text{(ii) } \forall w'_s\Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \text{a wolf comes through the door at } w'\big) \rightarrow \\
& \qquad h(3)(w') \text{ is a wolf that comes through the door at } w'\Big) \wedge \\
& \text{(iii) } \exists w'_s\Big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \text{a wolf comes through the door at } w'\Big)]
\end{aligned}
$$

So ⟦'A$_3$ wolf might-R_1 come through the door; it$_3$ would-R_2 eat you$_4$'⟧ is:

$$
\begin{aligned}
& [\lambda g_v.\lambda h_v.\lambda w_s.\exists i_v(\\
& \text{(i) } \forall \text{ indices } \iota\ \forall w'_s\Big(g(\iota)(w') \neq i(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \\
& \qquad\qquad\qquad \iota = 3 \wedge \text{there is a blue wolf at } w'\big)\Big)\wedge \\
& \text{(ii) } \forall w'_s\Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \text{a wolf comes through the door at } w'\big) \rightarrow \\
& \quad i(3)(w') \text{ is a wolf that comes through the door at } w'\Big)\wedge \\
& \text{(iii) } \exists w'_s\Big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \text{a wolf comes through the door at } w'\Big)\wedge \\
& \text{(iv) } i = h \wedge \forall w'_s(w \text{ is related to } w' \text{ by } i\text{'s } R_2 \rightarrow i(3)(w') \text{ eats } i(4)(w') \text{ at } w'))] \\
= \ & [\lambda g_v.\lambda h_v.\lambda w_s. \\
& \text{(i) } \forall \text{ indices } \iota\ \forall w'_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \\
& \qquad\qquad\qquad \iota = 3 \wedge \text{there is a blue wolf at } w'\big)\Big)\wedge \\
& \text{(ii) } \forall w'_s\Big(\big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \text{a wolf comes through the door at } w'\big) \rightarrow \\
& \quad h(3)(w') \text{ is a wolf that comes through the door at } w'\Big)\wedge \\
& \text{(iii) } \exists w'_s\Big(w \text{ is related to } w' \text{ by } g\text{'s } R_1 \wedge \text{a wolf comes through the door at } w'\Big)\wedge \\
& \text{(iv) } \forall w'_s(w \text{ is related to } w' \text{ by } h\text{'s } R_2 \rightarrow h(3)(w') \text{ eats } h(4)(w') \text{ at } w'))]
\end{aligned}
$$

So 'A$_3$ wolf-R_1 might come through the door; it$_3$ would-R$_2$ eat you$_4$' is true at the context providing the variable assignment g iff

$\exists h_v([\lambda g_v.\lambda h_v.\lambda w_s.$

(i) $\forall$ indices ι $\forall w'_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w$ is related to w' by g's $R_1 \wedge$

$\iota = 3 \wedge$ there is a blue wolf at $w'\big)\Big)\wedge$

(ii) $\forall w'_s\Big(\big(w$ is related to w' by g's $R_1 \wedge$ a wolf comes through the door at $w'\big) \rightarrow$

$h(3)(w')$ is a wolf that comes through the door at $w'\Big)\wedge$

(iii) $\exists w'_s\Big(w$ is related to w' by g's $R_1 \wedge$ a wolf comes through the door at $w'\Big)\wedge$

(iv) $\forall w'_s(w$ is related to w' by h's $R_2 \rightarrow h(3)(w')$ eats $h(4)(w')$ at $w'))]$

$(g)(h)(w_@) = \mathrm{T})$

iff

$\exists h_v\Big($

(i) $\forall$ indices ι $\forall w'_s\Big(g(\iota)(w') \neq h(\iota)(w') \rightarrow \big(w_@$ is related to w' by g's $R_1 \wedge$

$\iota = 3 \wedge$ there is a blue wolf at $w'\big)\Big)\wedge$

(ii) $\forall w'_s\Big(\big(w_@$ is related to w' by g's $R_1 \wedge$ a wolf comes through the door at $w'\big) \rightarrow$

$h(3)(w')$ is a wolf that comes through the door at $w'\Big)\wedge$

(iii) $\exists w'_s\Big(w_@$ is related to w' by g's $R_1 \wedge$ a wolf comes through the door at $w'\Big)\wedge$

(iv) $\forall w'_s(w_@$ is related to w' by h's $R_2 \rightarrow h(3)(w')$ eats $h(4)(w')$ at $w')\Big)$

Conjunct (iv) becomes (since R_2 relates the actual world to the R_1-actual-world-related worlds where a wolf comes through the door, because of the continuation at the context):

$$\forall w'_s((w_@ \text{ is related to } w' \text{ by } h\text{'s } R_1 \wedge \text{ a wolf comes through the door at } w') \rightarrow$$
$$h(3)(w') \text{ eats } h(4)(w') \text{ at } w')$$

The contextually-provided variable assignment will assign to 4 the individual concept taking every world to you. Let the accessible worlds be those that R_1 relates to the actual world. Then the whole thing holds iff both (i) at some accessible world a wolf comes through the door and (ii) at every accessible world where a wolf comes through the door, a wolf that comes through the door eats you.

Appendix B

Arrow Notation

Some dynamic semantic values are hard to parse. I find them easier to understand and quicker to calculate if they're displayed in a different notation using arrows. On extensional dynamic semantics, the semantic value of a sentence is a function from a variable assignment to a function from a variable assignment to a truth value. This can be thought of as a relation between variable assignments—the relation holding between one variable assignment and another iff the semantic value takes the first variable assignment to a function taking the second variable assignment to True. The semantic value of 'Richard paints', thought of relationally, relates two variable assignments iff (a) they are identical and (b) the semantic value of 'Richard' applied to one of them is a thing that runs. In arrow notation, for all variable assignments g and h:

$$g \xrightarrow{\text{Richard paints}} h \text{ iff } g = h \wedge [\![\text{`Richard'}]\!](\text{g}) \text{ paints}$$

The arrow notation is just a more visually obvious way of thinking about the dynamic semantic values of sentences. To translate from arrow notation into the orthodox set-theory language, given variable assignments γ, δ, and sentence Φ, the following arrow:

$$\gamma \xrightarrow{\Phi} \delta$$

is equivalent to the following language using set theory:

$$[\![\phi]\!](\gamma)(\delta) = \text{True}$$

With intensional semantics, semantic values now can be thought of as relating one variable assignment to another variable assignment relative to a world. We relativize the arrows to a world by subscripting a world-name under the arrow. So the following arrow:

$$\gamma \xrightarrow[\sigma]{\Phi} \delta$$

is equivalent to:

$$[\![\Phi]\!](\gamma)(\delta)(\sigma) = \text{True}$$

So, for all variable assignments g, h, and all worlds w:

$$g \xrightarrow[w]{\text{Richard paints}} h \text{ iff } g = h \wedge [\![\text{Richard}]\!](\text{g}) \text{ paints in } w$$

Speaking relationally, a discourse is true relative to a variable assignment iff the semantic value of the discourse takes that variable assignment to at least one variable assignment relative to the actual world. So ‘Richard paints’ is true relative to a variable assignment g and the actual world $w_@$ iff there is some h such that:

$$g \xrightarrow[w_@]{\text{Richard paints}} h$$

We derive the truth conditions by getting rid of arrows according to the arrow notation equivalents of the semantic rules. ‘paints’ takes the semantic value of a name or pronoun and delivers an arrow. Given some semantic value α of a name or pronoun, ‘paints’ gives the following arrow:

$$g \xrightarrow[w]{\text{paints}(\alpha)} h \text{ iff } g = h \wedge \alpha(g)(w) \text{ paints in } w$$

The dynamic conjunctive ‘;’ has the following clause:

$$g \xrightarrow[w]{\Phi;\Psi} h \text{ iff } \exists i\big(\, g \xrightarrow[w]{\Phi} i \ \wedge\ i \xrightarrow[w]{\Psi} h \,\big)$$

In other words, for $[\![\phi;\psi]\!]$ to relate g and h, ϕ must relate g with some variable assignment that ψ then relates to h.

The formidable ‘a_1’ receives the following clause, where ϕ and ψ are the semantic values of the two things ‘a’ combines with, in the simplest case these will be a common noun and intransitive verb:

$$g \xrightarrow[w]{a_1(\phi)(\psi)} h \text{ iff}$$

$$\exists i \exists j\big(g[1(w)]i \wedge\ i \xrightarrow[w]{\phi(\lambda k_v.\lambda w'_s.i(1)(w))} j \ \wedge\ j \xrightarrow[w]{\psi(\lambda k_v.\lambda w'_s.i(1)(w))} h \,\big)$$

Here’s how we calculate the semantic effect of our SVs on a sample sentence, ‘A physicist paints.’

$$g \xrightarrow[w_@]{\text{A}_1 \text{ physicist paints}} h \ \text{ iff}$$

$$\exists i \exists j\big(g[1(w_@)]i \wedge\ i \xrightarrow[w_@]{\text{physicist}(\lambda k_v.\lambda w'_s.i(1)(w_@))} j \ \wedge\ j \xrightarrow[w_@]{\text{paints}(\lambda k_v.\lambda w'_s.i(1)(w_@))} h \,\big) \text{ iff}$$

$$\exists i \exists j \big(g[1(w_@)]i \wedge i = j \wedge [\lambda k_v.\lambda w_s.i(1)(w_@)](i)(w_@) \text{ is a physicist in } w_@ \wedge$$

$$j = h \wedge [\lambda k_v.\lambda w_s.i(1)(w_@)](i)(w_@) \text{ paints in } w_@ \big) \text{ iff}$$

$$\exists i \exists j \big(g[1(w_@)]i \wedge i = j \wedge i(1)(w_@) \text{ is a physicist in } w_@ \wedge$$

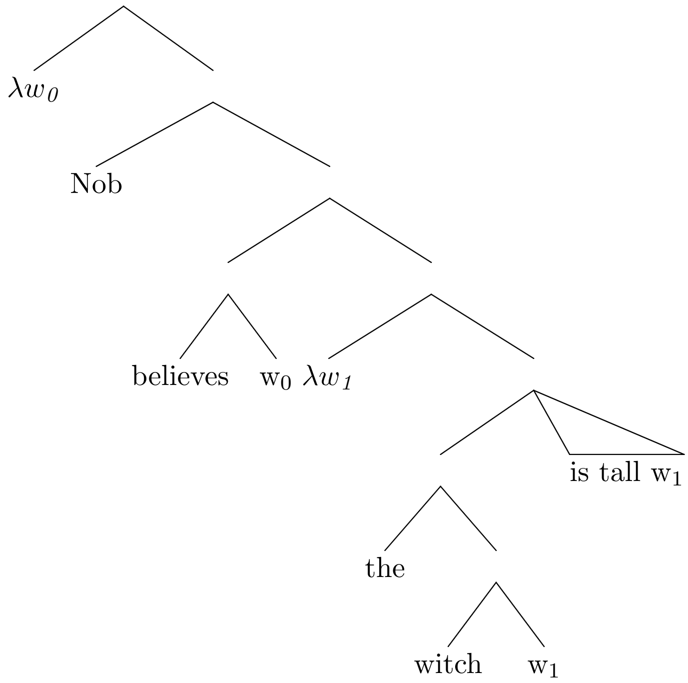
$$j = h \wedge i(1)(w_@) \text{ paints in } w_@ \big) \text{ iff}$$

$$g[1(w_@)]h \wedge h(1)(w_@) \text{ is a physicist in } w_@ \wedge h(1)(w_@) \text{ paints in } w_@$$

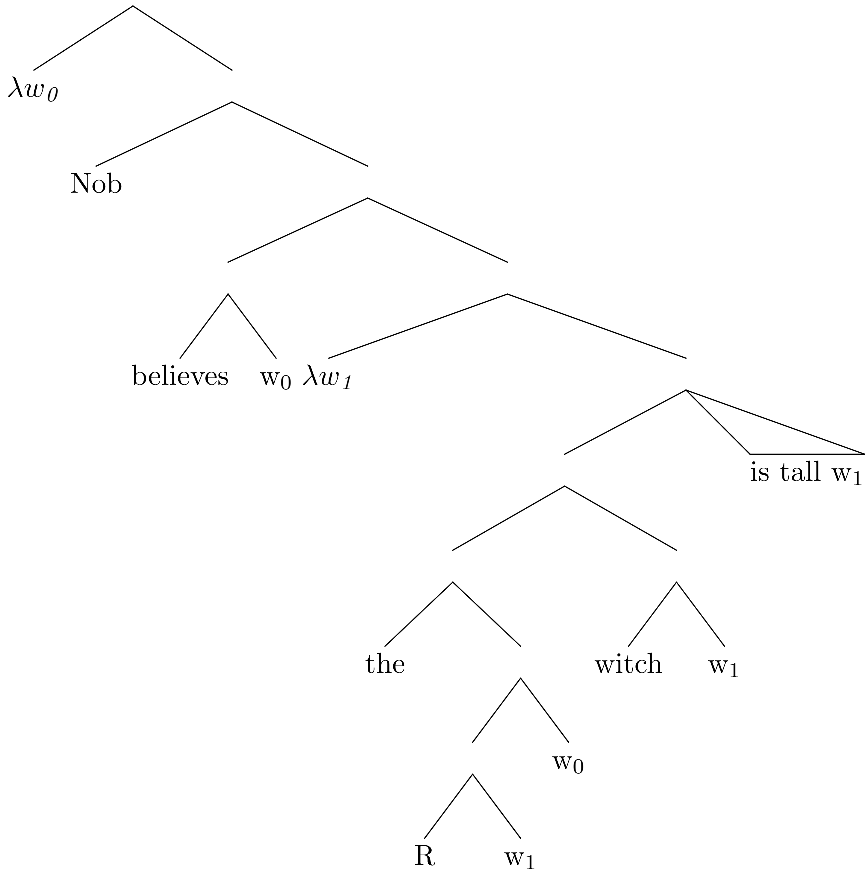
Appendix C

Implicit descriptions with world variables

In the world-variable semantics of (von Fintel and Heim 2011, c. 8), each verb and common noun is adjoined with a world variable, each sentence is adjoined with a world-variable binder, and different '*de dicto*' or '*de re*' readings are associated with different binding relationships between the binders and the variables. The sentence 'Nob believes the witch is tall' at our context would have the following syntactic structure given to semantics:



When we include our toy model of the contextual effect on the incomplete definite description, we add a variable 'R' (which, since it is the same semantic type as a common noun, will need to be conjoined with a world variable 'w_1', as 'witch' is) along with a world variable bound by 'λw_0'. This gives us:


λw_0
Nob
believes
w_0
λw_1
is tall w_1
the
witch
w_1
w_0
R
w_1

We have the old rules of Functional Application and Lambda Abstraction, and we have the following semantic values:

$$
\begin{array}{lcl}
[\![\text{`Nob'}]\!]^g & = & \text{Nob} \\
[\![\text{`witch'}]\!]^g & = & \lambda w_{\mathrm{s}}.\lambda x_{\mathrm{e}}.x \text{ is a witch in } w \\
[\![\text{`[is tall w}_1\text{]'}]\!]^g & = & \lambda x_{\mathrm{e}}.x \text{ is tall in } g(1) \\
[\![\text{`R'}]\!]^g & = & g(\text{`R'}) \\
[\![\text{`w}_0\text{'}]\!]^g & = & g(0) \\
[\![\text{`w}_1\text{'}]\!]^g & = & g(1) \\
[\![\text{`the'}]\!]^g & = & \lambda f_{\mathrm{et}}.\lambda g_{\mathrm{et}}.\text{the unique thing } x \text{ such that} \\
 & & f(x) = g(x) = \text{True} \\
[\![\text{`believes'}]\!]^g & = & \lambda w_{\mathrm{s}}.\lambda p_{\mathrm{st}}.\lambda x_{\mathrm{e}}.\forall w'(w' \text{ is one of } x\text{'s belief worlds at } w \rightarrow \\
 & & p(w') = \text{ True})
\end{array}
$$

At this context, 'R' is assigned the following function:

> $[\lambda w_{\mathrm{s}}.\lambda w'_{\mathrm{s}}.\lambda x_{\mathrm{e}}.x$ is described in w by what is, in w', the previous link in the communication chain between Nob and Hob$]$

Sentences at contexts are true if their semantic value relative to the variable assignment provided by the context takes the actual world to True. With these values, rules, and syntactic structure, the truth conditions described in chapter 4 are predicted. If we embed these sentences under further intensional phrases like 'necessarily', we again get the right results.

Bibliography

Abney, Steven. 1987. *The English Noun Phrase in its Sentential Aspect.* PhD diss., MIT.

Aloni, Maria. 2001. *Quantification under Conceptual Covers.* PhD diss., University of Amsterdam.

Barwise, Jon and John Perry. 1983. *Situations and Attitudes.* Cambridge, MA: The MIT Press.

Bende-Farkas Agnes and Hans Kamp. 2001. *Indefinites and Binding: From Specificity to Incorporation.* Lecture Notes. ESSLLI, Helsinki.

Brasoveanu, Adrian. 2007. *Structured Nominal and Modal Reference.* PhD diss., Rutgers.

Breckenridge, Wylie and Ofra Magidor. 2012. 'Arbitrary Reference'. *Philosophical Studies* 158: 377–400.

Burge, Tyler. 1983. 'Russell's Problem and Intentional Identity'. In *Agent, Language, and the Structure of the World*, edited by James Tomberlin, 79–110. Indianapolis, IN: Hackett.

Büring, Daniel. 2005. *Binding Theory.* Cambridge: Cambridge University Press.

Carnie, Andrew. 2013. *Syntax: A Generative Introduction*, Third Edition. Oxford: Wiley-Blackwell.

Chierchia, Gennaro. 1995. *Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar.* Chicago: University of Chicago.

——. 2005 'Definites, Locality, and Intentional Identity'. In *Reference and Quantification: The Partee Effect*, edited by Gregory Carlson and Jeffrey Pelletier, 143–77. Stanford: CSLI.

Cooper, Robin. 1979. 'The Interpretation of Pronouns'. In *Syntax and Semantics, Vol. 10: Selections from the Third Gröningen Round Table*, edited by Frank Heny and Helmut Schnelle, 61–92. New York: Academic Press.

Cumming, Sam. 2007. *Proper Nouns.* PhD diss., Rutgers.

——. 2013. 'Indefinites and Intentional Identity'. *Philosophical Studies.* Forthcoming in print. Online at: http://link.springer.com/article/10.1007%2Fs11098-013-0131-9

——. Forthcoming. 'Creatures of Darkness'. *Analytic Philosophy.*

Davies, Martin. 1981. *Meaning, Quantification, Necessity.* London: Routledge and Kegan Paul.

Devitt, Michael. 1981. *Designation.* New York: Columbia University Press.

Edelberg, Walter. 1986. 'A New Puzzle About Intentional Identity'. *Journal of Philosophical Logic* 15: 1–25.

——. 1992. 'Intentional Identity and the Attitudes'. *Linguistics and Philosophy* 15 (6): 561–96.

——. 1995. 'A Perspectivalist Semantics for the Attitudes'. *Noûs* 29 (3): 316–42.

Elbourne, Paul 2005. *Situations and Individuals.* Cambridge, MA: MIT Press.

——. 2008. 'The Interpretation of Pronouns'. *Language and Linguistics Compass* 2 (1): 119–50.

Evans, Gareth. 1977. 'Pronouns, Quantifiers, and Relative Clauses (I)'. *Canadian Journal of Philosophy* 7: 467–536.

——. 1980. 'Pronouns'. *Linguistic Inquiry* 11: 337–62.

——. 1982. *The Varieties of Reference.* Oxford: Clarendon Press.

Fine, Kit. 1983. 'A Defence of Arbitrary Objects'. *Proceedings of the Aristotelian Society, Supplementary Volumes* 57: 55–77.

——. 1984. 'Critical Review of Parsons' "Non-Existent Objects"'. *Philosophical Studies* 45 (1): 95–142.

——. 1985. *Reasoning with Arbitrary Objects.* Oxford: Blackwell.

von Fintel, Kai and Irene Heim. 2011. 'Intensional Semantics: MIT Spring 2011 Edition'. Online at: http://web.mit.edu/fintel/fintel-heim-intensional.pdf

Friedell, David. Forthcoming. 'Salmon on Hob and Nob'. *Philosophical Studies.*

Geach, Peter. 1967. 'Intentional Identity'. *The Journal of Philosophy* 64 (20): 627–32.

Glick, Ephraim. 2012. 'A Modal Approach to Intentional Identity'. *Noûs* 46 (3): 386–99.

Groenendijk, Jeroen and Martin Stokhof. 1990. 'Dynamic Montague Grammer'. In *Papers from the Second Symposium on Logic and Language*, edited by László Kálmán and László Polos, 3–48. Budapest: Akadémiai Kiadó.

——. 1991. 'Dynamic Predicate Logic'. *Linguistics and Philosophy* 14: 39-100.

Hawthorne, John and David Manley. 2012. *The Reference Book*. Oxford: OUP.

Heim, Irene. 1982. *The Semantics of Definite and Indefinite Noun Phrases.* PhD diss., University of Massachusetts Amherst.

——. 1990. 'E-type Pronouns and Donkey Anaphora'. *Linguistics and Philosophy* 13: 137–77.

——. 1991. 'Articles and Definiteness'. Published in German as 'Artikel und Definitheit' in *Sematics. An International Handbook of Contemporary Research*, edited by Arnim von Stechow and Dieter Wunderlich, 487–535. Berlin: de Gruyter.

Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford: Blackwell.

Jacobson, Pauline. 1999. 'Towards a Variable-Free Semantics'. *Linguistics and Philosophy* 22 (2): 117–84.

Kamp, Hans. 1981. 'A Theory of Truth and Semantic Representation'. In *Formal Methods in the Study of Language, Part I*, edited by Jeroen Groenendijk, T.M.V. Janssen, and Martin Stokhof, 277–322. Amsterdam: Mathematisch Centrum.

——. 1990. 'Prolegomena to a Structural Theory of Belief and Other Attitudes'. In *Propositional Attitudes: The Role of Content in Logic,* *Language, and Mind*, edited by C. Anthony Anderson, and Joseph Owens, 27–90. Menlo Park, CA: CSLI.

—— and Uwe Reyle. 1993. *From Discourse to Logic.* Dordrecht: Kluwer.

Kaplan, David. 1989. 'Demonstratives'. In *Themes from Kaplan*, edited by Joseph Almog, John Perry, and Howard Wettstein, 481–564. Oxford: OUP.

Karttunen, Lauri. 1974. 'Presupposition and Linguistic Context'. *Theoretical Linguistics* 1: 181–94.

——. 1976. 'Discourse Referents'. *Syntax and Semantics, Vol. 7: Notes from the Linguistic Underground.* London: Academic Press.

King, Jeffrey. 1993. 'Intentional Identity Generalized'. *Journal of Philosophical Logic* 22 (1): 61–93.

——. 1994. 'Anaphora and Operators'. *Philosophical Perspectives, 8, Logic and Language*: 221–50.

——. 2010. 'Anaphora'. *The Stanford Encyclopedia of Philosophy (Spring 2010 Edition)*, edited by Edward Zalta. Online at: http://plato.stanford.edu/archives/spr2010/entries/anaphora/

Kratzer, Angelika. 1977. 'What 'Must' and 'Can' Must and Can Mean'. *Linguistics and Philosophy* 1: 337–355.

——. 2011. 'Situations in Natural Language Semantics'. *The Stanford Encyclopedia of Philosophy (Fall 2011 Edition)*, edited by Edward Zalta. Online at: http://plato.stanford.edu/archives/fall2011/entries /situations-semantics/

Kripke, Saul. (1973) 2011. 'Vacuous Names and Fictional Entities'. *Philosophical Troubles: Collected Papers Volume 1.* Oxford: Oxford University Press. This article was first presented in 1973, and is a precursor of his Locke Lectures.

Larson, Richard and Gabriel Segal. 1995. *Knowledge of Meaning: An Introduction to Semantic Theory.* Cambridge, MA: MIT Press.

Lewis, David. 1968. 'Counterpart Theory and Quantified Modal Logic'. *Journal of Philosophy* 65: 113–26.

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

——. 1979. 'Attitudes *De Dicto* and *De Se*'. *The Philosophical Review* 88: 513–43.

——. 1986. *On the Plurality of Worlds.* Oxford: Blackwell.

Maier, Emar. 2009. 'Presupposing Acquaintance: a Unified Semantics for *De Dicto*, *De Re* and *De Se* Belief Reports'. *Linguistics and Philosophy* 32: 429–74.

Moltmann, Friederike. 2006. 'Unbound Anaphoric Pronouns: E-Type, Dynamic, and Structured-Propositions Approaches'. *Synthese* 153: 199–260.

Muskens, Reinhard. 1996. 'Combining Montague Semantics and Discourse Representation'. *Linguistics and Philosophy* 19: 143–86.

Neale, Stephen. 1990. *Descriptions.* Cambridge, MA: MIT Press.

Parsons, Terence. 1978. 'Pronouns as Paraphrases'. Unpublished manuscript.

——. 1980. *Nonexistent Objects.* New Haven, CT: Yale University Press.

Pietarinen, Ahti. 2001. 'Intentional Identity Revisited'. *Nordic Journal of Philosophical Logic* 6 (2): 147–88.

Quine, W.V.O. 1940. *Mathematical Logic.* Cambridge, MA: Harvard University Press.

——. 1956. 'Quantifiers and Propositional Attitudes'. *Journal of Philosophy* 53: 177–87.

Roberts, Craige. 1989. 'Modal Subordination and Pronominal Anaphora in Discourse'. *Linguistics and Philosophy* 12 (6): 683–721.

——. 1997. 'Anaphora in Intensional Contexts'. In *Handbook of Contemporary Semantic Theory*, edited by Shalom Lappin, 215–46. Oxford: Blackwell.

Saarinen, Esa. 1978. 'Intentional Identity Interpreted: A Case Study of the Relations among Quantifiers, Pronouns, and Propositional Attitudes'. *Linguistics and Philosophy* 2 (2): 151–223.

Salmon, Nathan. 1998. 'Nonexistence'. *Noûs* 32: 277–319.

——. 2002. 'Mythical Objects'. In *Meaning and Truth: Investigations in Philosophical Semantics*, edited by Joseph Campbell, Michael O'Rourke, and David Shier, 105–23. New York: Seven Bridges Press.

——. 2008. 'That *F*'. *Philosophical Studies* 141: 263–80.

——. Forthcoming. 'The Philosopher's Stone and Other Mythical Objects'. In a volume edited by Stuart Brock and Anthony Everett.

Schwarz, Bernhard. 2011. 'Long Distance Indefinites and Choice Functions'. *Language and Linguistics Compass* 5 (12): 880–97.

Schwarzschild, Roger. 2002. 'Singleton Indefinites'. *Journal of Semantics* 19: 289–314.

Soames, Scott. 1998. 'The Modal Argument: Wide Scope and Rigidified Descriptions'. *Noûs* 32: 1–22.

Stanley, Jason and Zoltán Gendler Szabó. 2000. 'On Quantifier Domain Restriction'. *Mind & Language* 15 (2–3): 219–261.

Stone, Matthew. 1999. 'Reference to Possible Worlds'. Rutgers Center for Cognitive Science Technical Report 49.

van Inwagen, Peter. 1977. 'Creatures of Fiction'. *American Philosophical Quarterly* 14 (4): 299–308.

van Rooy, Robert. 2000. 'Anaphoric Relations Across Attitude Contexts'. In *Reference and Anaphoric Relations*, edited by Klaus von Heusinger and Urs Egli, 157–82. Dordrecht: Kluwer.

Williamson, Timothy. 1998. 'Bare Possibilia'. *Erkenntnis* 48 (2/3): 257–73.

——. 2000. *Knowledge and its Limits*. Oxford: OUP.

——. 2002. 'Necessary Existents'. In *Logic, Thought and Language*, O'Hear (ed.), pp. 233–51. Cambridge: Cambridge University Press.

——. 2013. *Modal Logic as Metaphysics*. Oxford: OUP.