Semantics & Pragmatics Volume 5, Article 5: 1–57, 2012
http://dx.doi.org/10.3765/sp.5.5

Counterfactual attitudes and multi-centered worlds*

Dilip Ninan
Tufts University

Received 2012-05-22 / Accepted 2012-08-02 / Final version received 2012-08-08 / Published 2012-08-29

Abstract Counterfactual attitudes like imagining, dreaming, and wishing create a problem for the standard formal semantic theory of *de re* attitude ascriptions. I show how the problem can be avoided if we represent an agent's attitudinal possibilities using *multi-centered worlds*, possible worlds with multiple distinguished individuals, each of which represents an individual with whom the agent is acquainted. I then present a compositional semantics for *de re* ascriptions according to which singular terms are *assignment-sensitive* expressions and attitude verbs are *assignment shifters*.

Keywords: *de re* attitudes; *de se* attitudes; counterfactual attitudes; ascription; centered worlds; variables

1 Introduction

I can imagine that Obama lost the 2008 US presidential election even though I know that he won. It might make sense for me to imagine this if, for example, I am trying to work out what would have happened to the auto industry if Obama had lost. *Counterfactual attitudes* are, roughly, attitudes that one

* For providing useful feedback on various incarnations of these ideas, thanks to: Pranav Anand, Derek Ball, Alex Byrne, Mike Caie, Herman Cappelen, Alejandro Pérez Carballo, Nate Charlow, Simon Charlow, Andy Egan, Kai von Fintel, Sally Haslanger, Richard Holton, Ofra Magidor, Sarah Moss, Agustín Rayo, François Recanati, Dave Ripley, Paolo Santorio, Philippe Schlenker, Robert Stalnaker, Eric Swanson, Mike Titelbaum, Jason Turner, Stephen Yablo, Seth Yalcin, and an anonymous referee for *Semantics and Pragmatics*. Thanks also to audiences at MIT, Institut Jean Nicod (Paris), Arché (St Andrews), CSMN (Oslo), Dubrovnik, UC Berkeley, the University of Michigan, Leeds University, and the University of Barcelona. I am especially grateful to Josh Dever and Hans Kamp for providing extensive written comments on earlier drafts of this paper.

can coherently take towards a content p even when one knows that p is false. In addition to imagining, the class of counterfactual attitudes includes dreaming, wishing, and hoping.

This paper is about sentences that report *de re* attitudes, with particular attention to those that report counterfactual *de re* attitudes:

(1) a. Ralph imagined that Ortcutt was flying a kite.
b. Ralph wished that she had not opened the box.

I show that counterfactual attitudes pose a problem for the approach to *de re* attitude ascription standardly assumed in the formal semantics literature (Cresswell & von Stechow 1982), an approach based on Lewis's (1979, 1983) centered worlds account of *de se* and *de re* attitudes.

In response to this problem, I explore an alternative account, one on which attitudinal alternatives are represented by *multi-centered worlds*, possible worlds with multiple distinguished individuals, each of which is used to represent someone from the agent's world. I then examine how one might construct a compositional semantics adequate for predicting multi-centered worlds truth conditions for *de re* ascriptions. The semantics I present is based on the idea that pronouns, indexicals, and proper names are essentially *variables*: their semantic values are sensitive to a variable assignment. This approach synthesizes recent work on indexicals and pronouns (e.g. Heim 2008, Schlenker 2002) and on proper names (e.g. Cumming 2008, Dever 1998, Geurts 1997). Combined with the idea that attitude verbs are 'assignment-shifters' (cf. Cumming 2008, Santorio 2012), this account yields a simple semantics for attitude reports which avoids both the problem raised by counterfactual attitudes, along with a well-known compositional problem facing the standard account.

2 The standard account

2.1 *De se* attitudes

Lewis takes as his starting point *de se* attitudes, thoughts about oneself when one thinks of oneself in a characteristically first-person way. The account of *de se* attitudes that Lewis offers is presented in contradistinction to the classical possible worlds model that is typically presupposed in formal semantic treatments of attitude ascriptions (Hintikka 1962). On the possible worlds approach, an agent's belief state determines a set of possible worlds,

those worlds that are compatible with what the agent believes. A proposition is a set of possible worlds (or the characteristic function thereof), and a proposition p is true at a world w iff $w \in p$ (or $p(w) = 1$). An agent believes a possible worlds proposition p in world w just in case p is true at each world compatible with what the agent believes in w.

Although the previous paragraph only talks about beliefs, similar clauses can be given for mental states other than belief (desire, knowledge, imagination, etc.). I shall for the moment continue to focus on belief, but I assume that both the possible worlds account and Lewis's alternative to it are intended as general accounts of attitude content.

Lewis thought that the possible worlds account of attitudes could not accommodate certain features of *de se* attitudes, while his alternative account could. He offers a number of arguments and examples intended to establish this point (cf. Lewis 1979, 1983a). Here, I trace out one argumentative thread that can be found in Lewis 1983a.

We begin by considering Perry's character, Rudolf Lingens, an amnesiac lost in the Stanford Library.[1] In some intuitive sense, Lingens does not know who he is: he doesn't remember his name, where he went to school, who his parents are, what he does for a living, etc. Suppose Lingens, having touched his face, comes to believe *de se* that he has a mustache. What, on the possible worlds theory, is the content of Lingens's belief?

The most natural answer to our question is to say that the content of Lingens's belief is the set of possible worlds in which Rudolf Lingens has a mustache. To say that this is the content of Lingens's belief is to say that all of his belief worlds are worlds in which Rudolf Lingens has a mustache. Call this the *singular proposition view* of the *de se*. On this view, an agent x believes *de se* that she is F just in case every possible world compatible with what she believes is contained in the singular proposition $\{w\colon x \text{ is } F \text{ in } w\}$. Lewis would of course reject such a proposal since he holds that a given individual exists in at most one world (Lewis 1986). But those of us not bound by his metaphysics are free to consider this proposal.

One might object to this view as follows: Suppose Lingens has come across a biography of himself in the library. Since he doesn't know that he is Rudolf Lingens, he doesn't realize that the biography is about himself. The biography contains many picture of Lingens; in all of them, Lingens is

1 This is actually Frege's character (Frege 1918/1956), though Perry (1977) turned him into an amnesiac. Other classic papers on *de se* attitudes include Castañeda 1966, 1967, Perry 1979, and Stalnaker 1981.

clean-shaven. So Lingens doesn't believe that Rudolf Lingens has a mustache. So, the objection goes, he does not believe the set of possible worlds in which Rudolf Lingens has a mustache, after all. But that contradicts our earlier claim that Lingens does believe that singular proposition. The singular proposition view lands us in a contradiction, and so ought to be rejected.

But the objection takes it for granted that the proposition Lingens rejects when he says, "Rudolf Lingens doesn't have a mustache" is a singular proposition, the set of worlds in which Rudolf Lingens has a mustache. Admittedly, this is a natural assumption, but the possible worlds theorist *qua* possible worlds theorist is not forced to accept it. Instead, she might hold that the proposition Lingens rejects is a *descriptive* proposition, e.g. the set of worlds w such that the individual named "Rudolf Lingens" in w has a mustache in w. Since this proposition is independent of the singular proposition Lingens accepts (neither proposition entails/includes the other), Lingens can accept the singular proposition while rejecting this descriptive proposition.

But there is another objection to the singular proposition proposal that is more difficult to evade.[2] To see the problem, note that Lingens's origins presumably involve a certain sperm and egg. Moreover, the property of originating from a sperm and an egg is presumably a necessary property of Lingens's, one he has in every possible world in which he exists. Now, since Lingens believes *de se* that he has a mustache, he believes *de se* that he exists. So every possible world compatible with what he believes is, on the singular proposition view, a world in which Lingens exists. Since every world in which Lingens exists is a world in which Lingens comes from a sperm and an egg, every world compatible with what Lingens believes is a world in which Lingens comes from a sperm and an egg. But on the present proposal, this means that Lingens *believes de se* that he comes from a sperm and an egg. But surely Lingens can believe *de se* that he has a mustache even if he is convinced, for example, that "the myth of sperm and eggs is scientistic tommyrot and that a rival hypothesis is correct" (Lewis 1983a: p. 385).

This line of reasoning is quite general. Suppose the singular proposition view of the *de se* is correct, and let F be an arbitrary essential property of Lingens's. Then if Lingens believes *de se* that he exists, he will, on this proposal, count as believing *de se* that he has F. Since this holds for an arbitrary essential property of Lingens's, it holds for them all. But surely Lingens can believe *de se* that he exists without possessing a comprehensive

2 Lewis 1983a: pp. 384-385 discusses the counterpart-theoretic version of this objection. See also Lewis 1981 and Lewis 1986: p. 32.

set of true beliefs about his essence. Call this the *problem of essential properties*. The problem appears to sink the singular proposition view of the *de se*.[3]

If Lingens's *de se* belief that he has a mustache cannot be represented by a singular possible worlds proposition, can it be represented by a *descriptive* possible worlds proposition? On the descriptive proposal, there is a property F such that (i) Lingens is the unique F in the actual world; and (ii) Lingens believes *de se* that he has property G just in case every possible world compatible with what he believes contains a unique individual who is F, and who also has G. So for Lingens to believe *de se* that he has a mustache, for example, is for all his belief worlds w to be such that the F in w has a mustache in w. Call this the *descriptive proposition view* of the *de se*.

Note that in order for this proposal to avoid the problem of essential properties, it must be the case that there are some possible worlds w such that the unique F in w is someone other than Lingens. This essentially means that, in some of Lingens's belief worlds, someone other than Lingens represents Lingens as existing there. In some of his belief worlds, Lingens is represented not by himself, but by a *counterpart*, something that is similar to Lingens in some respect. Suppose w is one of Lingens's belief worlds. His *de se* counterpart c in w is the unique F in w; Lingens and c are similar insofar as Lingens is the unique F in the actual world and c is the unique F in w.

A standard way of arguing against this sort of view is to try out some candidates for property F and then argue that Lingens could believe *de se* that he has a mustache without believing that the bearer of the candidate property has a mustache, since Lingens might fail to realize that *he* is the bearer of the candidate property (cf. Perry 1977, 1979). For example: suppose Lingens is the unique man in the Stanford Library wearing a red sweater. Even so, the property of being a man in the Stanford Library wearing a red sweater is a bad candidate for being property F, since Lingens might fail to realize that he is in the Stanford Library. Thus, he might believe *de se* that he has a mustache without believing that the only man in the Stanford Library wearing a red sweater has a mustache (he may not even believe that there is a unique man in the Stanford Library wearing a red sweater). A similar fate awaits other properties that Lingens alone possesses: the property of being the only bearded philosopher west of the Mississippi, the property of being

3 This applies, of course, only to the singular *possible worlds* proposition view. The idea that the content of a *de se* attitude might be a singular *structured* proposition is, of course, a live option, but not one relevant for present purposes.

the only person with a scar shaped like *this*, the property of having exactly $2.34 in change in one's left trouser pocket.

But there are more plausible candidates for property F. Stalnaker (2008b: p. 70) suggests that Lingens will always be in a position to know the proposition he could express by saying, "I am the thinker of this (token) thought." Let T be the thought-token that Lingens has when he thinks to himself that he has a mustache. Stalnaker's suggestion is that our elusive property F might be the property of being the thinker of T. On this proposal, the content of Lingens's *de se* belief that he has a mustache is the set of worlds w in which the thinker of T in w has a mustache in w.

In order for this to avoid our original problem of essential properties, it must be that individuals other than Lingens can be the thinker of T. One might wonder if this is really possible, but even if we overlook this, another problem awaits, a problem that arises in connection with T's essential properties. For the proposition that Stalnaker claims that Lingens believes when he believes *de se* that he has mustache is, I take it, a singular proposition about T. (I will return to consider an alternative interpretation shortly.) Whatever a thought-token is, it presumably has some non-trivial essential properties. If T is a brain state, for example, then it is presumably essentially a brain state. Now suppose again that Lingens believes *de se* that he has a mustache. On this proposal, this means that all his belief worlds are worlds in which the thinker of T has a mustache; thus, they are all worlds in which T exists. Since T is a brain state in each world in which it exists, it is a brain state in each of Lingens's belief worlds. It follows that Lingens believes that T is brain state. But why should Lingens's believing *de se* that he has a mustache commit him to believing that T is a state of the brain? Maybe the latter claim conflicts with his general anti-scientistic outlook; maybe Lingens believes that T is a state of his soul and that souls are distinct from brains. His metaphysics may be exotic, but that shouldn't prevent him from holding the mundane *de se* belief that he has a mustache.

Note that the objection doesn't depend on T's being a brain state *per se*, but only on T's having some non-trivial essential properties which Lingens might fail to believe that T possesses. This problem points to a more general moral about candidates for property F: they probably should not be *haecceitist properties*, where a property G is *haecceitist*$_{df}$ just in case there is an individual x such that G is instantiated in a possible world w only if x exists in w. The property of being identical to Barack Obama is haecceitist in this sense, as is the property of being a child of Barack Obama — both are

instantiated at a world w only if Obama exists in w. The property of being a thinker of T is also haecceitist in this sense, since it will be instantiated at a world w only if T exists in w. The general problem with using a haecceitist property G to play the role of F is that the resulting theory will tell us that if Lingens believes *de se* that he exists, then all of his belief worlds are worlds in which G is instantiated; this in turn will mean that there is an x such that all of his belief worlds will be worlds in which x — along with any essential properties x may have — exists. And this will mean that Lingens can't believe anything about himself without believing that x is H, where H is any one of x's essential properties.

If we ought to steer clear of haecceitist properties in our search for property F, then it seems like F will have to be a *qualitative property*, a property which, intuitively speaking, is not 'about' an individual. But there is an old and well-known problem with thinking that F is a qualitative property (Strawson 1959: p. 20). For all we've said, Lingens might be in a 'reduplication universe', a universe in which the same set of qualitative properties are distributed in the same pattern across two distinct regions of space. Then there is no qualitative property G that Lingens possesses uniquely; *a fortiori*, there is no qualitative property G that Lingens possesses uniquely and which is such that Lingens believes that the unique bearer of G has a mustache. But living in a reduplication universe does not prevent Lingens from believing *de se* that he has a mustache. Note that, for all I've said, Lingens may believe *de se* that he is the unique G and that the unique G has a mustache. The point is that the content of his *de se* belief cannot be identified with the set of possible worlds w in which the unique G in w has a mustache in w. For that proposition is not true, whereas Lingens's belief is true — he really does have a mustache.

Before moving on, let me close one loophole in the preceding argument. My argument above against Stalnaker's 'thought-token' proposal assumed that he was proposing to represent Lingens's *de se* belief with a singular proposition about T. But at one point, Stalnaker suggests an alternative: T might be represented at other possible worlds by a *counterpart* (Stalnaker 2008b: 70, footnote 1). But this just postpones the pain. For T's counterpart at another possible world is presumably something that is similar to it in some respect. If the respect of similarity is haecceitistic — if there is an x such that any world in which T has counterpart is a world in which x exists — then the problem of essential properties reappears. If the relevant type of similarity is qualitative, then we run into the problem raised by the reduplication universe.

Stalnaker 2008a suggests yet another alternative: we can represent attitudinal alternatives using *pairs* of a possible world and a 'counterpart function'. Depending on how it is fleshed out, I think a proposal along these lines might avoid the foregoing difficulties. But note that to adopt it is to concede the present point, that the content of a *de se* attitude cannot be represented by a set of possible worlds. I will later discuss at length one way of working out Stalnaker's suggestion, but for the moment, I wish to put it aside and focus on Lewis's account of *de se* attitudes.

Lewis's idea is that we should represent attitudinal alternatives — the possibilities over which attitude verbs quantify — not with possible worlds, but with *centered worlds*, where a centered world is a triple consisting of a possible world, a time, and an individual (the *center*), who exists at the time and world in question. A centered world (w', t', x') will be compatible with what Lingens believes (at time t in world w) if and only if x' has, at t' in w', all the properties that Lingens believes *de se* that he himself has at t in w. The content of a belief, on the centered worlds proposal, is a *centered proposition*, a set of centered worlds (or the characteristic function thereof). A centered proposition p is true at a centered world (w, t, x) iff $(w, t, x) \in p$ (or $p(w, t, x) = 1$). Lingens believes a centered proposition p at t in w iff p is true at all the centered worlds compatible with what Lingens believes at t in w. If Lingens believes a centered proposition p at t in w, his belief is true *simpliciter* iff p is true at $(w, t,$ Lingens$)$.[4]

What is it, on this proposal, for Lingens to believe *de se* (at t in w) that he has a mustache? It is for all the centered worlds compatible with what he believes at t in w to be contained in the following centered proposition:

(2) $\{(w', t', x') : x'$ has a mustache at t' in $w'\}$

Unlike the singular proposition proposal, this account avoids the problem of essential properties. Let us suppose that y is some genetic property of Lingens — having such-and-such a DNA sequence, for example. Then it's plausible that Lingens has y essentially, and also plausible to suppose that Lingens fails to believe (at t in w) that he has y. On the centered worlds proposal, this means he doesn't believe (3):

(3) $\{(w', t', x') : x'$ has y at t' in $w'\}$

4 As should be clear, I take psychological attitudes to be four-place relations between an individual, a time, a world, and a content. But I will sometimes suppress reference to time and world to improve readability. So I will sometimes say, "So-and-so believes p" where it would be more precise to say, "So-and-so believes p at time t in world w".

Note that there are centered worlds (w', t', x') such that: (i) x' has a mustache at t' in w', and (ii) x' lacks y at t' in w'. Right now, I have a mustache here in the actual world. But I presumably don't have the same DNA sequence as Lingens does, and so I lack y right now in the actual world. This means that (2) is not a subset of (3), since (actual world, now, DN) is contained in the former but not in the latter. So even if all of Lingens's centered belief worlds are elements of (2), some of them may not be elements of (3), which means Lingens can believe (2) without believing (3).

Nor does the centered worlds approach have any trouble dealing with the possibility that Lingens might fail to possess uniquely any qualitative properties. Suppose he doesn't possess any at time t in world w, perhaps because w is a reduplication universe. Still, for him to believe (at t in w) that he has a mustache is for him to believe (2). His belief is true *simpliciter* just in case (2) is true at $(w, t,$ Lingens). And (2) is true at $(w, t,$ Lingens) iff Lingens has a mustache at t in w. So we predict the correct truth conditions for Lingens's *de se* beliefs even when Lingens finds himself in a reduplication universe.

2.2 *De re* attitudes

Quine (1956) observed an ambiguity in attitude reports like (4):

(4) Ralph believes that someone is a spy.

On the first reading, the sentence says that Ralph, like most of us, believes that there are spies — he believes that *someone or other* is a spy. This is the *de dicto* reading. The second reading is more interesting: it says that there is someone — Frank, perhaps — such that Ralph believes that *that person* is a spy. This is the *de re* reading, which appears to be strictly stronger than the *de dicto* reading.

In the first instance we are applying the term "*de re*" to a certain reading of an ascription like (4), but we can also apply that term to the attitude being ascribed. Let us suppose that (4) is true on its *de re* reading; then, intuitively, Ralph has someone in particular in mind — Frank, say — who he believes to be a spy. Then we can say that Ralph's belief about Frank is a *de re belief*. More could be said about this distinction, but I assume the reader is familiar with the contrast, and so I won't delve further into the issue here.

A natural way to try to capture Quine's ambiguity in the centered worlds framework is as follows. The *de dicto* reading of (4) is true at time t in world

w iff Ralph believes the following *general centered proposition* at t in w:

(5) $\{(w', t', x')$: there is a spy at t' in $w'\}$

Note that in order for Ralph to believe this it must be that every (w', t', x') compatible with what Ralph believes is such that there is a y' such that y' is a spy at t' in w' ($\forall\ \exists$). But it need not be the case that there is a y' such that every (w', t', x') compatible with what Ralph believes is such that y' is a spy at t' in w' ($\exists\ \forall$). But one might think that this is precisely the condition that must hold if (4) is to be true on its *de re* reading. That is, one might think that the *de re* reading is true at t in w iff there is a y such that Ralph believes, at t in w, the following *singular centered proposition* concerning y:

(6) $\{(w', t', x')$: y is a spy at t' in $w'\}$

In order for Ralph to count as believing this centered proposition, there must be a particular individual — Frank, say — who is a spy in all of Ralph's centered belief worlds. Note that this approach makes the *de re* reading strictly stronger than the *de dicto* reading, as desired.

Unfortunately, there are at least two problems with the *singular proposition view* of the *de re*. The first problem is one we have seen before: the problem of essential properties. Since we have discussed this at length in the *de se* case, I won't re-visit it here. The second problem — sometimes called the *problem of double-vision* — is well-known. To illustrate it, we adapt a much-discussed example from Quine 1956. Suppose Ralph sees a man in a brown hat behaving suspiciously one evening on the waterfront and comes to believe that the man he is observing is a spy. On the singular proposition account, this means that in every centered world compatible with what Ralph believes, this man is a spy. But suppose that Ralph also believes that Bernard Ortcutt, who is the mayor of the local town and who Ralph has seen on several occasions, is not a spy. According to the singular proposition account, this means that every centered world compatible with what Ralph believes is a world in which Ortcutt is not a spy. Finally, let us suppose that the man Ralph observes on the waterfront just *is* Ortcutt — Ralph is acquainted with Ortcutt in two different ways but fails to realize this. Given what we've said so far, this means that every centered world compatible with what Ralph believes is such that Ortcutt/that man is a spy *and* such that Ortcutt/that man is not a spy. Since no centered world meets that condition, it follows that no centered world is compatible with what Ralph believes. This is a bad result. For one, it means we represent Ralph as having contradictory

beliefs, even though it seems that, for all we've said, Ralph's beliefs might be perfectly consistent. A second problem is that, given a standard semantics for attitude verbs, every sentence of the form *Ralph believes such-and-such* will be vacuously true if no centered worlds are compatible with what Ralph believes.[5]

Following Quine and Kaplan, Lewis offers a *descriptivist* solution to these problems.[6] On Lewis's version of the descriptivist story, the fundamental notion of *de re* belief is the idea of believing something about someone *relative to an acquaintance relation*.[7] An acquaintance relation is a relation "of a sort apt for the reliable transmission of information" (Lewis 1979: p. 155). As I shall use the term, an acquaintance relation is simply any relation that underwrites an agent's ability to have a thought about an object.[8] The notion of believing something about someone relative to an acquaintance relation is a notion with reasonably clear intuitive content. This can be seen by reflecting on examples. Ralph is acquainted with Ortcutt in two different ways, but fails to realize this. He bears relation Q to Ortcutt, where Q is the relation x bears to y just in case y is the unique individual that x sees sneaking around on the waterfront. Relative to this relation, Ralph believes that Ortcutt is a spy. Ralph also bears relation S to Ortcutt, where S is the relation x bears to y just in case y is the unique individual that x has heard of under the name "Bernard J. Ortcutt". Relative to this relation, Ralph believes that Ortcutt is *not* a spy. In some cases it will be useful to put the "relative to" qualification next to the referring expression at issue; for example, I will often say things like: "Ralph believes that Ortcutt (relative to Q) is a spy, but not that Ortcutt (relative to S) is a spy".

5 If one thinks that belief reports presuppose that the agent's belief state is non-empty, then all sentences of that form will be infelicitous and/or neither-true-nor-false — an equally bad result.

6 See Quine 1956, Kaplan 1968, and Lewis (1979: §XIII).

7 Lewis (1979: §XIII) talks about "ascribing a property to an individual under a description". But he takes a "description" to be a relation, and in the case of *de re* belief, he requires the relation to be a relation of acquaintance. So this comes to the same thing as my 'believing something about someone relative to an acquaintance relation'.

But this is, in any case, something of a simplification; something more general is needed for the 'multiply *de re*' case, as when Ralph believes that he is standing in between Jones and Ortcutt. The general notion is something like 'believing something about a plurality of individuals-relative-to-acquaintance relations'.

8 The idea that an agent x can have a *de re* attitude about an object y only if x stands to y in a special causal or epistemic relation is controversial; see Hawthorne & Manley 2012 and the papers in Jeshion 2010 for some recent discussion of the issues involved.

Lewis offers the following analysis of this notion:[9]

(7) An agent x believes, at t in w, that y is F, relative to acquaintance relation R iff:

(i) x bears R uniquely to y at t in w, and

(ii) x believes *de se* (at t in w) that the thing to which he bears R is F.

Given Lewis's account of *de se* belief, (ii) says that all the of centered worlds (w', t', x') compatible with what x believes at t in w are such that the thing to which x' bears R at t' in w' is F at t' in w'.

So when Ralph believes, relative to Q, that Ortcutt is a spy, Ralph believes that the man to whom he bears Q is a spy, i.e. he believes that the man he saw sneaking around on the waterfront is a spy. The content of his belief is the following centered proposition:[10]

(8) $\{(w', t', x')$: the individual that x' sees sneaking around on the waterfront at t' in w' is a spy at t' in $w'\}$

And when Ralph believes, relative to S, that Ortcutt is not a spy, he believes that the man to whom he bears S is not a spy, i.e. he believes that the man he knows under the name "Bernard Ortcutt" is not a spy. The content of his belief is the following centered proposition:

(9) $\{(w', t', x')$: the individual that x' has heard of under the name "Bernard Ortcutt" at t' in w' is not a spy at t' in $w'\}$

Note that (8) and (9) are consistent: there are centered worlds in which the man the center sees sneaking around on the waterfront is a spy, while the man the center knows under the name "Bernard Ortcutt" is not a spy.

9 This is an analysis of a simplified *analysandum* (see footnote 7). To handle the general case, we need an analysis of:

> Agent x believes, at t in w, that y_1 (relative to R_1), ..., y_n (relative to R_n) stand in relation T.

But I take it that it is straightforward to generalize the above analysis to this case.

10 Definite descriptions that occur in the metalanguage will be assumed to have a Russellian semantics. So (8) is the set of centered worlds (w', t', x') such that there is a y' such that y' is the unique individual that x' sees sneaking around on the waterfront at t' in w', and y' is a spy at t' in w'.

A second virtue of Lewis's account is that it avoids the problem of essential properties. The individual who represents Ortcutt (relative to Q) in one of Ralph's doxastic alternatives (w', t', x') is the individual to whom x' bears Q at t' in w'. Since that individual needn't be Ortcutt himself, that individual needn't have any of Ortcutt's essential properties.

Before I discuss the semantics of attitude ascriptions, let me mention something about the relationship between the *de se* and the *de re* on Lewis's account. Note that if Ralph believes *de se* that he himself is a spy, then (4), repeated here as (10), is true on its *de re* reading:

(10) Ralph believes that someone is a spy.

That suggests that a *de se* belief is also a *de re* belief about oneself. This suggestion is vindicated by Lewis's theory if we count the relation of *identity* as a relation of acquaintance, and take a *de se* belief to be a *de re* belief about oneself relative to the relation of identity. To see how this works, note that on Lewis's account, Ralph believes *de se* (at t in w) that he is a spy iff (11) holds.

(11) Every centered world (w', t', x') compatible with what Ralph believes at t in w is such that x' is a spy at t' in w'.

And Ralph believes *de re* (at t in w) that Ralph is a spy relative to the relation of identity iff (12a) and (12b) both hold.

(12) a. Ralph bears the relation of identity uniquely to Ralph at t in w.
b. Every centered world (w', t', x') compatible with what Ralph believes at t in w is such that the individual to whom x' bears the relation of identity uniquely at t' in w' is a spy at t' in w'.

(11) is equivalent to the conjunction of (12a) and (12b). To see this, first note that (12b) entails (12a): (12b) is only true if Ralph exists at t in w, and if he exists at t in w then he is identical to Ralph at t in w, in which case (12a) is true. So the conjunction of (12a) and (12b) is equivalent to (12b). Second, note that (12b) is just a long-winded way of saying what (11) says, since the individual to whom x' bears the relation of identity uniquely at t' in w' is just x' herself. So (11) and (12b) are equivalent. Since (12b) is equivalent to the conjunction of (12a) and (12b), (11) is equivalent to that conjunction: belief *de se* is belief *de re* about oneself relative to the relation of identity.

All of the acquaintance relations we have talked about so far are relations that relate an individual x to exactly one other individual y. For example, Q

is the relation that x bears to y just in case y is *the unique individual* that x has seen sneaking around on the waterfront. Thus, I will use expressions like "$Q(w,t,x)$" as an abbreviation for "the unique individual to whom x bears Q at t in w". And as I mentioned earlier (see footnote 10), I assume that definite descriptions used in the metalanguage have a Russellian semantics.

2.3 Attitude ascription

Note that the relativized notion of *de re* belief — believing something about someone relative to an acquaintance relation — can be used to characterize an absolute notion of *de re* belief as follows:

(13) Agent x believes, at t in w, that y is F *simpliciter* iff there is an acquaintance relation R such that:

(i) $R(w,t,x) = y$, and

(ii) x believes *de se* (at t in w) that the one to whom he bears R is F.

Note that this means that, in the double-vision scenario, Ralph counts both as believing *simpliciter* that Ortcutt is a spy and as believing *simpliciter* that Ortcutt is not a spy.

The standard approach to *de re* ascription in the formal semantics literature uses this absolute notion in stating the truth conditions of *de re* reports.[11] Start with Quine's sentence (4), repeated here as (14):

(14) Ralph believes that someone is a spy.

The standard account says that this sentence is true (on its *de re* reading) just in case there is an individual y and an acquaintance relation R such that Ralph bears R to y, and Ralph believes that the one to whom he bears R is a spy. Slightly more formally:

$[\![(14)]\!]^{c,(w,t)} = 1$ iff

there is an individual y and an acquaintance relation R such that:

(i) $R(w,t,\text{Ralph}) = y$, and

11 See, for example, Abusch 1997, Aloni 2001, Anand 2006, Cresswell & von Stechow 1982, Heim 1994, Percus & Sauerland 2003a,b, von Stechow 1982, and Maier 2006, 2009.

(ii) every centered world (w', t', x') compatible with what Ralph believes at t in w is such that $R(w', t', x')$ is a spy at t' in w'.

The above semantic account of (14) is usually extended to cover *de re* ascriptions containing singular terms — names, indexicals, pronouns, definite descriptions — in the way one would expect. For example, the truth conditions of (15) would be given as follows:

(15) Ralph believes that Ortcutt is not a spy.

$[\![(15)]\!]^{c,(w,t)} = 1$ iff there is an acquaintance relation R such that:

(i) $R(w, t,$ Ralph$) =$ Ortcutt, and

(ii) every centered world (w', t', x') compatible with what Ralph believes at t in w is such that $R(w', t', x')$ is not a spy at t' in w'.

Note that these truth conditions are *coarse-grained*, in the sense that the *de re* reading of (15) is true so long as there is *some* acquaintance relation relative to which Ralph believes that Ortcutt is a spy. So the truth conditions of (15) would be satisfied in the double-vision scenario in virtue of the following two facts: (a) that Ortcutt is the individual Ralph has heard of under the name "Bernard Ortcutt", and (b) that Ralph believes the individual he has heard of under the name "Bernard Ortcutt" is not a spy. That prediction seems fine, but the fact that the account is coarse-grained in this way means that it has consequences that some might find problematic. For example, the *de re* reading of (16) would also be true in Quine's scenario:

(16) Ralph believes that Ortcutt is a spy.

$[\![(16)]\!]^{c,(w,t)} = 1$ iff there is an acquaintance relation R such that:

(i) $R(w, t,$ Ralph$) =$ Ortcutt , and

(ii) every centered world (w', t', x') compatible with what Ralph believes at t in w is such that $R(w', t', x')$ is a spy at t' in w'.

The truth of (16) follows from the following two facts: (a) that Ortcutt is the man Ralph sees sneaking around on the waterfront, and (b) that Ralph believes that the man he sees sneaking around on the waterfront is a spy.

One might accept that (16) has *a* reading on which it is true — or that there are some contexts in which it is true — but one might also think that the sentence has another, more salient reading on which it is false (or that in most ordinary contexts, it is false). After all, Ralph himself would not accept (16) as a description of his belief state, nor would he nod approvingly if he heard someone utter the words, "Ortcutt is a spy".

The fact that (16) seems false might be accounted for in a number of different ways. For example, one might hold that the domain of acquaintance relations over which we quantify in stating the truth conditions of a *de re* ascription can be restricted by the utterance context, and that a salient restriction is often suggested by the noun phrase used in the ascription itself (cf. Heim 1992).[12] Thus, if (16) is evaluated relative to a context in which the domain of acquaintance relations only contains relation S, then (16) will be false relative to that context.

A more detailed treatment would be needed to compare this account of Frege's Puzzle (and the many subtle variations on it discussed in the philosophical literature) with its many competitors. My only aim here was to show how this account of *de re* ascriptions could be modified to account for some familiar intuitions about attitude reports. But since these issues are only tangentially connected to the phenomena I wish to focus on, I will for the most part ignore them and utilize the coarse truth conditions account in what follows.

Although I have only been talking about belief reports so far, I assume, of course, that this proposal is intended as a general account of attitude reports. In particular, I assume parallel accounts for counterfactual attitude reports (e.g. imagination reports, wish reports).

There remains the question of how these truth conditions are compositionally generated. One approach holds that the *res* expression is moved out of the complement clause into an argument position of the attitude verb. On this view, in a sentence like (16) — repeated here as (17) — *believes* takes three arguments: *Ralph*, *Ortcutt*, and *is a spy*:

(17) Ralph believes that Ortcutt is a spy.
≈ Ralph believes [Ortcutt] [λx x is a spy]

12 Advocates of structured proposition approaches to attitude content have also developed contextualist treatments of Frege's Puzzle. See, for example, Crimmins & Perry 1989 and Richard 1990.

Believes then receives the following lexical entry:[13]

> $[\![\text{believes}]\!]^{c,(w,t)} = \lambda res_e.\lambda p_{(s(r,et))}.\lambda att_e.$ there is an acquaintance relation R such that:
>
> (i) $R(w,t,att) = res$, and
>
> (ii) every centered world (w',t',x') compatible with what *att* believes at t in w is such that that $p(w',t',y') = 1$, where $y' = R(w',t',x')$.

To handle attitude reports with more than one *res* expression, we would need something more sophisticated, along the lines of Cresswell & von Stechow 1982. On that treatment, the first argument of the verb is not a single individual, but a sequence of individuals, and the second argument is not a monadic property but an n-ary relation.[14]

But even with this modification, accounts of this sort are widely disliked on the grounds that *res* movement is thought to be syntactically implausible. Maier (2009: p. 459) offers a round-up of recent complaints about *res* movement; for example, he cites Schlenker (2004: p. 190) who refers to this process as '*De Re* Magic'. While there are accounts in the literature which avoid positing such movement,[15] I won't investigate them here. For I wish to raise another problem for this approach, one which targets the *truth conditions* that this theory assigns to attitude reports. Any account that predicts the Lewisian truth conditions is thus subject to this objection, even if it manages to avoid the problems associated with *res* movement.

3 The problem of counterfactual attitudes

3.1 The problem stated

The problem I wish to discuss is quite simple and can be made vivid with an example. Let us alter our original Ralph/Ortcutt example as follows. Ralph is only acquainted with Ortcutt in one way at t in w, having seen Ortcutt sneaking around on the waterfront. (So in this version of the story, Ralph hasn't heard of the man called "Bernard Ortcutt", town mayor.) But, as

13 Individuals are of type e, worlds of type s, times of type r, truth values of type t.

14 Something must also be said about how this entry relates to *believes* as it occurs in *de dicto* attitude reports. On the face it, it looks like this approach must posit a systematic ambiguity in the meaning of attitude verbs.

15 See Anand 2006, Percus & Sauerland 2003a, and Maier 2009.

before, Ralph believes at t in w that this fellow who is sneaking around on the waterfront is a spy. The relevant acquaintance relation is Q, the relation x bears to y just in case y is the unique individual that x sees sneaking around on the waterfront.

The problem is that, once Ralph has a 'cognitive fix' on Ortcutt, he is able to consider scenarios in which he is not acquainted with Ortcutt in this way. For example, Ralph might imagine, at t in w, a scenario in which he sees Ortcutt (relative to Q) flying a kite in an alpine meadow, rather than sneaking around on the waterfront. If he does imagine this, one of the things he counts as imagining about Ortcutt (relative to Q) is that he does not see Ortcutt sneaking around on the waterfront, i.e. he is imagining (relative to Q) that he does not bear Q to Ortcutt. But on Lewis's analysis of what it is to imagine something about someone relative to an acquaintance relation, Ralph imagines, at t in w, that he does not see Ortcutt (relative to Q) sneaking around on the waterfront iff:

(18) a. $Q(w, t, \text{Ralph}) = \text{Ortcutt}$, and
b. every centered world (w', t', x') compatible with what Ralph imagines at t in w is such that x' does not see $Q(w', t', x')$ sneaking around on the waterfront at t' in w'.

The first condition, (18a), is met: Ralph does bear Q to Ortcutt at t in w. But, supposing at least one centered world is compatible with what Ralph imagines, the second condition, (18b), is not met. No centered world (w', t', x') *whatsoever* is such that x' does not see $Q(w', t', x')$ sneaking around on the waterfront at t' in w', since $Q(w', t', x')$ just is the individual that x' sees sneaking around on the waterfront at t' in w'. *A fortiori*, no centered world (w', t', x') *compatible with what Ralph imagines* is such that x' does not see $Q(w', t', x')$ sneaking around on the waterfront at t' in w'. Thus, (18b) does not hold, and the account makes the false prediction that Ralph cannot imagine (relative to Q) that he does not see Ortcutt sneaking around on the waterfront.

Before I consider some responses to this problem, let me consider a linguistic version of the problem. Relative to the scenario described, the following sentence would appear to be true:

(19) Ralph imagined that he did not see Ortcutt sneaking around on the waterfront.

On the standard approach to *de re* ascription, however, this sentence is true at world w and time t iff there is an acquaintance relation R meeting two conditions:

- $R(w, t,$ Ralph) = Ortcutt, and
- every centered world (w', t', x') compatible with what Ralph imagines at t in w is such that x' does not see $R(w', t', x')$ sneaking around on the waterfront at t' in w'.

But *ex hypothesi* the only acquaintance relation Ralph bears to Orctutt is Q, so Q is the only potential witness for the existential quantifier. But as we've seen, Q fails to satisfy the second of these two conditions. The account incorrectly predicts that (19) is false in this scenario.[16]

But does this case reveal a genuine problem? Perhaps we have simply mis-identified the acquaintance relation under which Ralph thinks of Ortcutt. Perhaps the acquaintance relation is tied to Ralph's visual perception of Ortcutt and so is not easily put into words, either by him or by us. While this suggestion may help to evade the present counterexample, others are easy to find. Consider a slightly different scenario: as above, except that Ralph imagines that Ortcutt died as infant (cf. Kripke 1980). In this case, there may be no interesting similarities between the physical appearance of the man Ralph sees on the waterfront and the unfortunate infant of Ralph's imagining. In such a case, (20) would be true, but it's unclear how its truth could be accounted for by the standard approach.

(20) Ralph imagined that Ortcutt died as infant.

What this version of the case seems to show is that the Ortcutt with whom Ralph is acquainted at t in w and the individual who represents Ortcutt in an arbitrary one of Ralph's imagination alternatives need not be qualitatively similar in any interesting respect.

One way to think of the point is this. On Lewis's approach, an agent identifies the *res* using a relation of acquaintance: the agent thinks of the *res* as 'the one to whom I bear relation R'. But the fact that the agent bears R to the *res* will, in normal cases, be a *contingent* fact, one that might not have obtained. And typically, if a fact is contingent, then one can imagine that

16 Again, this assumes that at least one centered world is compatible with what Ralph imagines. If no centered worlds are compatible with what Ralph imagines, then the sentence is either vacuously true or neither true nor false.

it fails to obtain. Once we appreciate this general point, it seems unlikely that the problem can be resolved simply by being careful about choosing the appropriate acquaintance relation under which the agent thinks of the *res*.

Before I consider some other possible replies on behalf of Lewis's approach, let me make two observations about this problem. First, it arises not just with imagining but with other counterfactual attitudes such as wishing and dreaming. Ralph could *wish* that he had not seen Ortcutt on the waterfront, and he could have a *dream* in which he does not see Ortcutt on the waterfront.[17]

Second, the above problem would seem to show that an agent can imagine that a given *res* is F even if there is no relation of acquaintance R such that (a) the agent bears R to the *res* and (b) the agent imagines that the thing to whom she bears R is F. This shows that the standard account fails in the left-to-right direction, so to speak: the relevant sentence can be true, even while the proposed truth condition is not met. In other words, the proposed truth condition is *not necessary* for the truth of the sentence. But the reverse is also true: the proposed truth condition is *not sufficient* for truth of the sentence.

To see this, suppose again that Ralph is only acquainted with Ortcutt via Q, the relation x bears to y just in case y is the unique individual that x sees sneaking around on the waterfront. Now Ralph might imagine a scenario in which there is a unique individual that he sees sneaking around on the waterfront, and in which that individual is dumping a body into the water. But note that his imagining might be purely *de se*, i.e. it need not be *de re* with respect to Ortcutt. Ralph might simply be imagining that he (Ralph) sees *someone or other* on the waterfront, dumping a body into the water — the dubious character in his imaginary scenario need not be somebody that he identifies with Ortcutt. But note that, if w at t are the world and time of Ralph's imagining in this case, the following is true:

> There is an acquaintance relation R such that:
>
> (i) $R(w, t, \text{Ralph}) = \text{Ortcutt}$, and
>
> (ii) every centered world (w', t', x') compatible with what Ralph imagines at t in w is such that x' sees $R(w', t', x')$ dumping a body into the water at t' in w'.

17 Maier (2006: pp. 34-35) recognizes that counterfactual attitudes raise a problem for the standard approach; his example involves the attitude of hoping. But he sets the problem aside and does not offer a solution to it.

Q is a witness for the existential quantifier. From this it follows according to the standard account that (21) is true in this scenario:

(21) Ralph imagined that he saw Ortcutt dumping a body into the water.

But I put it to you that Ralph's purely *de se* imagining is not sufficient for the truth of (21). Ralph doesn't take himself to be imagining anything about any individual of his acquaintance (other than himself). He's simply imagining that he sees someone or other dumping a body into the water. That purely general imagining does not support the truth of the above *de re* ascription. Thus, (21) is false in this scenario, but the standard approach predicts it to be true.[18]

Having mentioned this 'flip-side' problem, I will set it aside in what follows and focus on the initial problem of counterfactual attitudes.

3.2 Replies (and replies to those replies)

A natural response on behalf of the standard account is to say that the infant in the imaginary scenario isn't the individual the center sees sneaking around on the waterfront *in that scenario*, but the individual that Ralph *actually* sees sneaking around on the waterfront. If t and w are the time and world (respectively) at which Ralph's act of imagining takes place, then the individual Ralph actually sees is the individual he sees at t in w, namely Ortcutt. So according to this line of thought, the content of Ralph's imagining is the set of centered worlds in which the center does not see *Ortcutt* sneaking around on waterfront.

This is a promising suggestion, at least at first glance. The initial problem arose because the standard account predicts that Ralph imagined at t in w (relative to Q) that he did not see Ortcutt sneaking around on the waterfront only if Ralph imagined the following descriptive centered proposition:

(22) $\{(w', t', x')$: x' does not see $Q(w', t', x')$ sneaking around on the waterfront at t' in $w'\}$

18 Percus & Sauerland (2003a: n.19) notice this problem. They note that if (a) John dreams that the person who reviewed his paper was a bald man in his 90's, and (b) Mary is the person who in fact reviewed John's paper, we still would be unwilling to count the sentence *John dreamed that Mary was a bald man in his 90's* as true. They suggest that this is because of a restriction on what sorts of acquaintance relations (or 'concept generators' in their theory) *dreams* can quantify over. But it's hard to see how this could work as a general solution to the problem.

The trouble was that this actually denotes the empty set of centered worlds. But on the present proposal, Ralph imagines at t in w (relative to Q) that he does not see Ortcutt sneaking around on the waterfront only if Ralph imagines the following *singular* centered proposition:

(23) $\{(w', t', x')$: x' does not see $Q(w, t, x)$ sneaking around on the waterfront at t' in $w'\}$
$= \{(w', t', x')$: x' does not see Ortcutt sneaking around on the waterfront at t' in $w'\}$

And this centered proposition is non-empty: there are centered worlds in which the center does not see Ortcutt sneaking around on the waterfront.

Note, though, that this view is similar (if not identical) to the *singular proposition* view that we discussed — and discarded — in §2.2. So it will face the problem of double-vision and the problem of essential properties. For example, this proposal will run into trouble if Ralph imagines a scenario in which Ortcutt (relative to Q) is distinct from Ortcutt (relative to S), where S is the relation x bears to y iff y is the unique individual that x has heard of under the name "Bernard Ortcutt". On this approach, the content of Ralph's imagining would be the set of centered worlds in which Ortcutt is distinct from Ortcutt — the empty set again. And this account will also face the problem of essential properties: if G is an arbitrary one of Ortcutt's essential properties, Ralph wouldn't be able to imagine something about Ortcutt without imagining that Ortcutt possessed G.

Advocates of the philosophical view known as *two-dimensionalism* might argue that there is a more sophisticated way of bringing "actually" in at this point.[19] According to one version of two-dimensionalism, any attitude is associated with *two kinds* of content. To get a handle on what these two contents are, consider the case in which Ralph imagines that Ortcutt (relative to Q) died as an infant. According to two-dimensionalism, Ralph's imagining can be characterized using some semantic properties of a sentence that might plausibly be thought to express the content of that imagining. The sentence in question is (24):

(24) The individual who I actually see sneaking around on the waterfront died as an infant.

19 Advocates of two-dimensionalism of the sort I have in mind include Chalmers (1996, 2002) and Jackson (1998), though the view I outline in the text may not correspond exactly to the views of either of these authors. Stalnaker (1978, 1988) offers a different version of two-dimensionalism, which we might have considered instead. I think that his approach will also face a version of the problem that I discuss in the text.

If one is a descriptivist about the *de re*, it is plausible to think that an utterance of this sentence by Ralph would express the content of Ralph's imagining.

Now note that, in a standard Kaplanian semantics, we can define two notions of the *content* of a sentence. The first is Kaplan's official notion of the content of a sentence at a context c: the set of indices i such that the sentence is true at c and i (Kaplan 1989). This gives us what the two-dimensionalist calls the *horizontal* content of a sentence at a context. If we assume that the index contains only a possible world, the horizontal content of a sentence at a context is a possible worlds proposition:

Horizontal content of sentence ϕ at context c: $\{w' : [\![\phi]\!]^{c,w'}\}$

But we can also define a second notion of the content of a sentence, if we take the set of contexts at which the sentence is true. This gives us what the two-dimensionalist calls the *diagonal* content of a sentence:

Diagonal content of sentence ϕ: $\{c : [\![\phi]\!]^{c,w_c}\}$

If we assume that a context is a centered world (so $c = (w_c, t_c, x_c)$), then the diagonal content of a sentence is a centered proposition.

Now consider (24). According to the two-dimensionalist, Ralph's imagining at t in w determines two contents, one corresponding to the horizontal of (24) at the context $(w, t,$ Ralph), the other to the diagonal of (24). The diagonal content of (24) is the following centered proposition:

(25) $\{(w', t', x') :$ the individual that x' sees sneaking around on the waterfront at t' in w' died as an infant in $w'\}$

This is what Lewis took to be *the* content of Ralph's imagining; we've argued already that this centered proposition is empty. But the horizontal content of (24) at $(w, t,$ Ralph) is a coherent singular possible worlds proposition:

(26) $\{w' :$ the individual that Ralph sees sneaking around on the waterfront at t in w died as an infant in $w'\}$

Since Ortcutt is the individual that Ralph sees sneaking around on the waterfront at t in w, this is the set of possible worlds in which *Ortcutt* died as an infant.

That's a nice result, but so far we've just replicated what was achieved by the first proposal that used an 'actualized' description. The advantage

of the two-dimensionalist approach arises in connection with the problem of double-vision. Suppose, for example, that Ralph imagines a scenario in which Ortcutt (relative to Q) is distinct from Ortcutt (relative to S). This time the horizontal content of Ralph's imagining is empty, since it is the set of possible worlds w' such that the individual Ralph saw sneaking around on the waterfront at t in w is distinct from the individual that Ralph has heard of under the name "Bernard Ortcutt" at t in w, aka the set of worlds in which Ortcutt is distinct from Ortcutt. But this time the *diagonal* content will not be empty, for the diagonal content will be the following centered proposition:

(27) $\{(w', t', x')$: the individual that x' sees sneaking around on the waterfront at t' in w' is distinct from the individual that x' has heard of under the name "Bernard Ortcutt" at t' in $w'\}$

This is the same centered proposition that Lewis would assign as the content of this imagining, and it is a perfectly good, non-empty centered proposition.

Suppose we allow the two-dimensionalist to claim that an imagining is coherent if *either* its diagonal content *or* its horizontal content is non-empty. Then we can say that the two-dimensionalist has succeeded in showing that Ralph's imaginings in the two cases of interest (the original problem case, the double-vision case) are coherent. Of course, there are still some important questions facing this account, such as: How exactly does a single episode of imagining relate to, or determine, these two contents? What sort of semantics for *de re* attitude ascriptions goes along with this account?[20]

But even if we waive these questions, a further problem awaits. The problem arises when we *combine* the problem of double-vision with the problem of counterfactual attitudes. To see the problem, let's again suppose that Ralph has double-vision with respect to Ortcutt. He sees Ortcutt sneaking around on the waterfront (he bears Q to Ortcutt), and he knows of Ortcutt via the name "Bernard Ortcutt" (he bears S to Ortcutt). Still, Ralph might have an imagining which he could report as follows:

> I'm imagining a situation in which that guy [Ralph points at the man on the waterfront] is distinct from Bernard Ortcutt, and in which I never saw that guy, and in which Ortcutt never goes by the name "Bernard Ortcutt".

20 The view may also run into a version of the problem of essential properties.

In other words, Ralph is imagining that Ortcutt (relative to Q) is distinct from Ortcutt (relative to S), and that he doesn't bear Q to Ortcutt (relative to Q), and that he doesn't bear S to Ortcutt (relative to S).

The problem is that both the horizontal content and the diagonal content of Ralph's imagining are empty. The diagonal content is the following centered proposition:

(28) $\{(w', t', x')$: $Q(w', t', x') \neq S(w', t', x')$ at t' in w', and x' has never seen $Q(w', t', x')$ at t' in w', and $S(w', t', x')$ never goes by the name "Bernard Ortcutt" at t' in $w'\}$

That (28) denotes the empty centered proposition is over-determined: there is no centered world (w', t', x') in which x' fails to see $Q(w', t', x')$ at t' in w', and there is no centered world (w', t', x') in which $S(w', t', x')$ never goes by the name "Bernard Ortcutt" at t' in w', given that "$Q(w', t', x')$" abbreviates "the individual that x' sees sneaking around on the waterfront at t' in w'", and that "$S(w', t', x')$" abbreviates "the individual that x' has heard of under the name "Bernard Ortcutt" at t' in w'". This is just the problem of counterfactual attitudes once again.

If w and t are the world and time of Ralph's imagining, then the horizontal content of his imagining is the following possible worlds proposition:

(29) $\{w'$: $Q(w, t, \text{Ralph}) \neq S(w, t, \text{Ralph})$ in w', and Ralph has never seen $Q(w, t, \text{Ralph})$ in w', and Ralph has not heard of $S(w, t, \text{Ralph})$ under the name "Bernard Ortcutt" in $w'\}$

But since $Q(w, t, \text{Ralph}) = \text{Ortcutt} = S(w, t, \text{Ralph})$, this is a set of possible worlds in which Ortcutt is distinct from Ortcutt. Since there aren't any possible worlds like that, (29) is empty. So the trouble remains.[21]

21 Note that the *two-dimensional intension* — which one can think of as a function from contexts to horizontal contents — of (24) is non-empty. So one might attempt to use this object in characterizing the content of Ralph's imagining. It may prove fruitful to pursue this idea in conjunction with the idea that counterfactual attitudes are somehow 'parasitic' on the base attitude of belief; see Asher 1987, Kamp 1990, and Heim 1992 for work in this vein. I explore these ideas sympathetically in Ninan 2008: Ch. 2; Yanovich 2011 offers a similar proposal.

4 A solution

4.1 Questions of identification

I begin my search for a solution with the following observation: the problem with counterfactual attitudes does not arise for the centered worlds treatment of *de se* attitudes. To see this, it will be useful to contrast the centered worlds account with a theory that *does* face a ***de se*** version of the problem of counterfactual attitudes. Recall the *descriptive (possible worlds) proposition view* of *de se* attitudes discussed — and discarded — in §2.1. On that account, there is a property F such that Lingens is the unique F in w, and Lingens believes *de se* in w that he has property G just in case every possible world w' compatible with what he believes is such that the F in w' has property G in w'. I argued against that view earlier, but here I want to note that this approach also faces an additional problem: it faces a ***de se*** version of the problem of counterfactual attitudes.

To see what I mean by this, suppose that the property that plays the relevant role for Lingens in w is the property of being the tallest philosopher in Palo Alto. So according to the account in question, for Lingens to imagine *de se* in w that he is G is for all of his imagination worlds w' to be such that the tallest philosopher in Palo Alto in w' is G in w'. Note that we immediately have the result that Lingens cannot imagine anything incompatible with his being the tallest philosopher in Palo Alto. Lingens cannot, for example, imagine that his growth was stunted and his career path diverted, so that he ended up as a diminutive dentist in Poughkeepsie, rather than as the tallest philosopher in Palo Alto. For according to the descriptivist approach in question, for Ralph to imagine this in w is for every possible world w' compatible with what he imagines in w to be such that the tallest philosopher in Palo Alto in w' is a diminutive dentist in Poughkeepsie in w' and not the tallest philosopher in Palo Alto in w'. But no possible worlds meet this condition.

The centered worlds theorist avoids this particular problem. Even if the property of being the tallest philosopher in Palo Alto is the only property that Lingens uniquely possesses and that he believes himself to possess uniquely, that property plays no role in the centered worlds account of the content of Lingens's *de se* imaginings. For Lingens to imagine that he is not the tallest philosopher in Palo Alto is simply for all his centered imagination worlds (w', t', x') to be such that x' is not the tallest philosopher in Palo Alto at t' in w'. There is no problem here, since there certainly are centered worlds

like that; the one centered on me right now in the actual world is an example.

We can think of the key difference between possible worlds descriptivism and the centered worlds theory in terms of how they answer the following **question of *de se* identification**:

> Let s be a potential imagination alternative for Lingens in w.
> Who represents Lingens in s?

A *potential* imagination alternative for Lingens in w is a possibility that we can assess for compatibility with what he imagines in w. For the moment, I am relying on our intuitive notion of a possibility that we can assess for compatibility with what someone imagines. So I am not presupposing that potential imagination alternatives are possible worlds nor that they are centered worlds nor that they are something else; part of the point of this section is to propose a new way to represent potential imagination alternatives (and potential attitudinal alternatives more generally). Thus, to answer this question, one should first say what sort of entity one takes s to be, and then say who in s represents Lingens there. Of course, if Lingens is suffering from identity confusion, his imagination alternatives might contain more than one individual who corresponds to Lingens there, in which case our question is not well-defined. But, to simplify our discussion, I put this complication aside for the moment. So let us suppose that Lingens does not suffer from identity confusion, and that his only attitudes about himself are *de se* attitudes.

The possible worlds descriptivist answers our question of *de se* identification as follows: s is a possible world w', and there is a property F such that (i) Lingens is the unique F in w, and (ii) the individual who represents Lingens in w' is the unique F in w'. (Should w' fail to contain a unique F, we automatically rule it out as incompatible with what Lingens imagines at t in w.) Assuming F is a qualitative property — as it needs to be in order to avoid the problem of essential properties — this view requires the individual x' in w' who represents Lingens there to be qualitatively similar to Lingens (as he is in w) in the following respect: x' must be the unique F in w', just as Lingens is the unique F in w. And this means that any potential imagination alternative for Lingens in w that manages to contain a representative for Lingens will end up representing Lingens as being the unique F. As a result, the view predicts (incorrectly) that Lingens cannot imagine a scenario in which he is not the unique F.

The centered worlds theorist, on the other hand, offers a different answer to our question of *de se* identification. She says: s is a centered world (w', t', x'), and the individual who represents Lingens in that centered world is simply x', the center. Note that this answer does not require the individual in (w', t', x') who represents Lingens to be qualitatively similar to Lingens (as he is at t in w) in any interesting respect. So even if Lingens believes, correctly, that he is the tallest philosopher in Palo Alto at t in w, neither that property (*being the tallest philosopher in Palo Alto*) nor any other plays a role in determining who represents Lingens at one of his potential imagination alternatives. When we are assessing a centered world for compatibility with what Lingens imagines at t in w, the center of that centered world is simply *stipulated* to be Lingens's representative there.

This means that, on the centered worlds approach, the '*de se* representation relation' is extremely thin: anything anywhere in logical space represents a way for Lingens to be, a way that we can assess for compatibility with what Lingens imagines about himself at t in w. Consider any individual x' who exists at any time t' in any world w'. No matter how qualitatively dissimilar x' (as she is at t' in w') is from Lingens (as he is at t in w), the centered world (w', t', x') represents a way for Lingens to be, a way that we can assess for compatibility with what Lingens imagines or believes.

This may seem like a strange feature of the account, but it is worth keeping in mind two things. First, it is evidently this feature of the account that enables it to solve the *de se* problem of counterfactual attitudes. Second, if an agent has a normal amount of self-knowledge, she will have a large number of correct beliefs about the qualitative properties she possesses. This means that her centered belief worlds will be centered on individuals who *are* qualitatively similar to her in important respects. Furthermore, when she imagines things about herself, she will usually import a great deal of this information into the imaginary scenario, and thus the individual who represents her in one of her imagination alternatives will also be similar to her in many respects. For example, when I imagine myself scaling K2, I hold fixed a great number of facts about myself: facts about how I look, how many siblings I have, what my name is, and so on. So the individual who represents me in one of my imagination alternatives *will* be similar to me in many respects. But that this is so reflects two facts: (i) the fact that I possess a good deal of accurate information about myself, and (ii) the fact that I normally import a lot of information about myself into my *de se* imaginings. The fact that my representatives are similar to me in these ways

is *not* a result of inherent features of the framework for representing my *de se* attitudes. The framework itself places essentially no constraint on which individuals in logical space represent a way for me to be.

Despite this minimalist answer to the question of *de se* identification, the centered worlds theorist begins to sound more like the possible worlds descriptivist when it comes to *questions of* ***de re*** *identification*. Recall the case that led to the ***de re*** problem of counterfactual attitudes: The only acquaintance relation Ralph bears to Ortcutt at t in w is Q, the relation x bears to y just in case y is the unique individual who x sees sneaking around on the waterfront. And Ralph imagines at t in w that he did not see Ortcutt sneaking around on the waterfront. Now consider the following **question of *de re* identification**:

> Let s be a potential imagination alternative for Ralph at t in w.
> Who represents Ortcutt in s?

The centered worlds theorist answers this question as follows: s is a centered world (w', t', x'), and the individual y' who represents Ortcutt in that centered world is the one to whom x' bears Q at t' in w', i.e. the individual who x' sees sneaking around on the waterfront at t' in w'. (Again, if there is no such individual in (w', t', x'), then that centered world is automatically ruled incompatible with what Ralph imagines at t in w.)

Note that this answer *does* require the individual y' in (w', t', x') who represents Ortcutt there to be qualitatively similar to Ortcutt (as he is at t in w) in a certain respect. More precisely, it requires the *pair* (x', y') to be similar to the pair (Ralph, Ortcutt) in the following way: x' must bear Q to y' at t' in w', just as Ralph bears Q to Ortcutt at t in w. So any potential imagination alternative for Ralph at t in w in which Ralph and Ortcutt both have representatives will represent Ralph as seeing Ortcutt sneaking around on the waterfront. This is what leads the account to predict (incorrectly) that Ralph cannot imagine a scenario in which he *does not* see Ortcutt sneaking around on the waterfront.

The trouble for possible worlds descriptivism about the *de se* and for Lewis's view of the *de re* seems to stem from a common source: the demand that the relevant 'representing individual' in a potential imagination alternative be qualitatively similar in a particular way to the corresponding 'represented individual'. The centered worlds theorist avoids the *de se* problem of counterfactual attitudes precisely because the agent's representative in one of her potential imagination alternatives (i.e. the center of a

centered world) needn't be qualitatively similar in any interesting respect to the agent herself. This suggests a strategy for solving the *de re* problem of counterfactual attitudes: find a way of modelling imagination alternatives which does not require '*de re* representatives' to be qualitatively similar to the individuals that they represent.

4.2 Pair-centered worlds

In order to simplify our discussion, I want to continue to put aside the complications that arise in connection with the double-vision phenomenon. So until further notice, I will ignore cases in which an agent is acquainted with an individual in more than one way. And I will assume that the way in which any agent is acquainted with herself is via the relation of identity. We will come back to the problem of double-vision in the next subsection.

On the centered worlds account, the agent's representative in an alternative is a *separate coordinate* of the alternative. A centered world is a kind of structured possibility, a possibility with three parameters or coordinates: world, time, and center. One of the coordinates — the center — is specified as being the individual that represents the agent in that alternative. This seems to be what ensures that there is no requirement of qualitative similarity between the individual who represents the agent at a potential alternative and the agent herself.

Can we treat individuals *other than the agent* in the same way? For example: can we treat Ortcutt's representative in one of Ralph's potential imagination alternatives as a further coordinate of that alternative? Let a *pair-centered world* be a triple $(w, t, \{x, y\})$ consisting of a world, a time, and a *pair* of individuals that exist at that world and time. (A 'pair' here is just an unordered set with exactly two members.) If a centered world represents a possible way that an individual might be, a pair-centered world represents a possible way that a *pair* of individuals might be. We might then try to represent Ralph's potential imagination alternatives using pair-centered worlds, allowing one of the centers in a pair-centered world to represent Ralph, the other to represent Ortcutt.[22]

But notice an issue that arises as soon as we add additional centers to our representation of potential imagination alternatives: we need some sort of mechanism for saying who each center represents. Suppose, for example, that

22 Something similar to pair-centered worlds have been used in accounts of '*de te*' attitude ascription; see Schlenker 1999, Anand 2006, and Ninan 2010.

Ralph imagines at t in w that he is short and that Ortcutt is tall. And suppose that we want to assess a particular pair-centered world $(w', t', \{x', y'\})$ for compatibility with what Ralph imagines at t in w. If x' is short at t' in w' and y' is tall at t' in w', is $(w', t', \{x', y'\})$ compatible with what Ralph imagines? The question does not admit of a categorical answer, since we do not know whether x' represents Ralph and y' Ortcutt, or vice-versa. The best we can say is: $(w', t', \{x', y'\})$ is compatible with what Ralph imagines if x' represents Ralph and y' Ortcutt, but it is not compatible with what he imagines if the reverse is true.

Note that no analogous problem arises in the 'singly-centered' worlds framework. A centered world contains only *one* distinguished individual — the center — and we have a background stipulation in place that tells us who the center of a centered world represents. The stipulation says that, when we are assessing a centered world for compatibility with what an agent x imagines (or believes, etc.), the center represents x. But once we moved to a 'multi-centered' framework, we need some sort of mechanism for saying who a given center in a potential attitudinal alternative is supposed to represent.

Here is a fanciful way of thinking about what such a mechanism would look like. Imagine that we could take a pair-centered world $(w', t', \{x', y'\})$ and hang *labels* or *tags* around the necks of both x' and y'; for example, we might hang the *Ralph* tag around x''s neck and the *Ortcutt* tag around y''s neck. Tagging the individuals in this way would make it completely clear to us as theorists who each center was being used to represent. If we put the *Ralph* tag on x' and the *Ortcutt* tag on y', then x' would represent Ralph, and y' would represent Ortcutt. Then if x' was short at t' in w', and y' tall, we would know that $(w', t', \{x', y'\})$ represented Ralph as being short and Ortcutt as being tall. So if Ralph had imagined at t in w that he was short and that Ortcutt was tall, we would know that this 'tagged' pair-centered world was compatible with what he had imagined.

But what is a *tagged* pair-centered world? Since we can't actually travel around logical space and drop tags around the necks of various possible persons, here is how I propose to 'tag' the centers of a pair-centered world. Take the pair-centered world $(w', t', \{x', y'\})$, and replace x' with the *ordered pair* (Ralph, x'), and y' with the ordered pair $(\text{Ortcutt}, y')$. This procedure yields the following object:

$$(w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y')\})$$

x' and y' play the same role as they did before: they are individuals that exist

at t' in w' who are being used to represent individuals with whom Ralph is acquainted at t in w, namely himself and Ortcutt. But it should now be clear who each of x' and y' represent: x' represents Ralph, y' Ortcutt. Each center represents its corresponding tag. I now re-define the term "pair-centered world" to refer to objects of this sort, and I will say that (Ralph, x') and $(\text{Ortcutt}, y')$ are the two *tagged centers* of the above pair-centered world.

To see how this helps with the problem of counterfactual attitudes, we need to say something about what it is for a pair-centered world to be compatible with what Ralph imagines at t in w. Recall that, at t in w, Ralph's imagining concerns only himself and Ortcutt. So, intuitively speaking, a possibility s should be compatible with what Ralph imagines at t in w just in case: (i) the individual x' who represents Ralph in s has all the properties that Ralph imagines himself to have; (ii) the individual y' who represents Ortcutt in s has all the properties that Ralph imagines that Ortcutt has, and (iii) the two individuals, x' and y', who represent Ralph and Ortcutt respectively in s should stand in all the relations that Ralph imagines that he and Ortcutt stand in. This suggests the following:

(30) A pair-centered world world $(w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y')\})$ is compatible with what Ralph imagines at t in w iff:

- x' has, at t' in w', all the properties Ralph imagines that he himself has at t in w;
- y' has, at t' in w', all the properties Ralph imagines that Ortcutt has at t in w; and
- x' and y' stand, at t' in w', in all the relations that Ralph imagines that he and Ortcutt stand in at t in w.

Note that since any one-place property can be represented as a (degenerate) two-place relation, we really only need the third of these three clauses.

Suppose, for example, that at t in w, Ralph imagines all and only the following things:

- that he is tall,
- that Ortcutt is short, and
- that he and Ortcutt are friends.

Then a pair-centered world $(w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y')\})$ will be compatible with what Ralph imagines at t in w iff:

- x' is tall at t' in w',
- y' is short at t' in w', and
- x' and y' are friends at t' in w'.

Now to apply this idea to the problem of counterfactual attitudes. Recall: Ralph is acquainted with Ortcutt at t in w in only one way, via Q, the relation x bears to y just in case y is the unique individual that x sees sneaking around on the waterfront. Nevertheless, it seems possible that Ralph might imagine at t in w a scenario in which he does not see Ortcutt sneaking around on the waterfront. On the present approach, we would represent Ralph's episode of imagining by saying that all of the pair-centered worlds $(w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y')\})$ compatible with what Ralph imagines at t in w are such that x' does not see y' sneaking around on the waterfront at t' in w'. And there are pair-centered worlds that meet this condition: for example, since I do not see Barack Obama sneaking around on the waterfront on January 1, 2012 in the actual world α, $(\alpha, 01/01/2012, \{(\text{Ralph}, \text{Dilip}), (\text{Ortcutt}, \text{Obama})\})$ is a pair-centered world that meets this condition.

Note that because *any* pair-centered world is a potential imagination alternative for Ralph at t in w, there needn't be any interesting sort of qualitative similarity between the individual who represents Ortcutt in a pair-centered world and Ortcutt himself (as he is at t in w). Similarly, consider (x', y'), the pair of individuals who represent the pair (Ralph, Ortcutt) in the pair-centered world $(w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y')\})$. Again, there is no requirement that (x', y') be qualitatively similar to (Ralph, Ortcutt) (as they are at t in w) in any interesting respect. In particular, even if the only acquaintance relation Ralph bears to Ortcutt at t in w is Q, (x', y') can represent (Ralph, Ortcutt) in $(w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y')\})$ even if x' does not bear Q to y' at t' in w'. So this approach appears to solve the ***de re*** problem of counterfactual attitudes for the same reason the centered worlds theory solved the ***de se*** problem of counterfactual attitudes. In both cases, the solution is to relinquish the demand that the representing individuals in a potential imagination alternative be qualitatively similar to the individuals that they represent.

4.3 Multi-centered worlds

To handle attitudes towards more than two individuals we need to consider *multi-centered worlds*, i.e. possibilities with more than two tagged centers. If, for example, Ralph were to imagine that he was standing between Ortcutt and Jones, we could use the following sort of object to represent one of Ralph's potential imagination alternatives:

$$(w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y'), (\text{Jones}, z')\})$$

We could then gloss the notion of what it is for a multi-centered world like this to be compatible with what Ralph imagines in the way one would expect given the above account of pair-centered compatibility, i.e. (30). More generally, when an agent has an attitude towards n individuals, we would use a multi-centered world with n tagged centers to represent one of her potential attitudinal alternatives. I'll provide more details on how this works in a moment, but first I want to address the problem of double-vision.

Recall our original case of double-vision, and suppose for simplicity that Ralph only has attitudes (at t in w) about himself (relative to identity), about Ortcutt (relative to Q), and about Ortcutt (relative to S), where Q is the relation x bears to y just in case y is the unique individual that x sees sneaking around on the waterfront, and S the relation that x bears to y just in case y is the unique individual that x has heard of under the name "Bernard Ortcutt". Suppose now that Ralph imagines at t in w a scenario in which each of these three characters figures, e.g. he imagines a scenario he could report by saying, "I'm imagining that I'm refereeing a boxing match between that man [he points at the man on the waterfront] and Bernard J. Ortcutt. That man is in the red corner, while Ortcutt is in the blue corner." Intuitively, each of Ralph's imagination alternatives at t in w ought to contain at least three individuals: one who represents Ralph himself (relative to identity), and two who both correspond Ortcutt, one representing Ortcutt (relative to Q), the other representing Ortcutt (relative to S). Ortcutt needs two distinct representatives in each of Ralph's imagination alternatives, since Ralph is imagining that Ortcutt (relative to Q) is distinct from Ortcutt (relative to S).

We might try to simply represent Ralph's potential imagination alternatives using multi-centered worlds with three tagged centers, two of which have *Ortcutt* tags on them, e.g.:

$$s = (w', t', \{(\text{Ralph}, x'), (\text{Ortcutt}, y'), (\text{Ortcutt}, z')\})$$

But this isn't quite adequate. While s does contain two representatives for Ortcutt — y' and z' — we don't know which representative is which. Which of y' and z' represents Ortcutt *relative to* Q? Which represents Ortcutt *relative to* S? Suppose that, at t' in w', x' is refereeing a boxing match between y' and z', and that y' is in the red corner and that z' is in the blue corner. Is s compatible with what Ralph imagines? It is if y' represents Ortcutt (relative to Q) and z' represents Ortcutt (relative to S), but it's not if the reverse is true.

The solution is to tag our centers not merely with *individuals*, but with *individual-acquaintance relation pairs*. In other words, we should represent one of Ralph's potential alternatives at t in w not with an object like s, but with an object like s':

$$s' = (w', t', \{((\text{Ralph}, \text{identity}), x'), ((\text{Ortcutt}, Q), y'), ((\text{Ortcutt}, S), z')\})$$

Now it is clear that y' is the individual in s' that represents Ortcutt relative to Q and that z' is the individual in s' that represents Ortcutt relative to S.

s' is the sort of thing we can assess for compatibility with what Ralph imagines at t in w: it is compatible with what he imagines at t in w just in case x', y', and z' stand, at t' in w', in all the relations that Ralph imagines that he (relative to identity), Ortcutt (relative to Q), and Ortcutt (relative to S) stand in at t in w. I shall continue to use the term "multi-centered world", but now I use it to refer to objects like s', possibilities whose centers are tagged by individual-relation pairs, rather than simply by individuals. It is objects of this sort that we will use to represent potential attitudinal alternatives.[23]

We are now in a position to state a more general theory. I begin with an observation about multi-centered worlds like s' above: the third coordinate of s', viz.:

$$\{((\text{Ralph}, \text{identity}), x'), ((\text{Ortcutt}, Q), y'), ((\text{Ortcutt}, S), z')\}$$

is a *function*, i.e. a set of ordered pairs such that if (a, b) and (a, c) are in the set, then $b = c$. The domain of the function is a set of individual-acquaintance relation pairs:

23 I am assuming that to believe that one is F relative to the relation of identity is to believe *de se* that one is F. The real status of this claim is worth investigating, but in the present context, it can simply be regarded as a stipulation. (Thanks to Josh Dever for discussion here.)

$$\{(\text{Ralph}, \text{identity}), (\text{Ortcutt}, Q), (\text{Ortcutt}, S)\}$$

and its range is a set of individuals:

$$\{x', y', z'\}$$

To characterize this general type of function, it will help to define the notion of *an agent's acquaintance set*:

> DEFINITION. A set A of individual-acquaintance relation pairs is agent x's *acquaintance set* at time t in world w iff A contains all and only those pairs (y, R) such that x bears R to y at t in w.

An acquaintance set *simpliciter* is the acquaintance set of some agent at some time in some world. We can then define the notion of a *tagging function* as follows:

> DEFINITION. A function f is a *tagging function* iff the domain of f is an acquaintance set, and there is a world-time pair (w, t) such that the range of f is included in the set of individuals that exist at t in w.

Using this notion, we can say that a multi-centered world (w, t, f) is a triple consisting of a possible world w, a time t, and a tagging function f, where the range of f is included in the set of individuals that exist at t in w.

We can characterize what it is for an arbitrary multi-centered world to be compatible with what an agent imagines as follows:[24]

24 The fact that (w', t', f') is compatible with what x imagines at t in w only if the domain of f' is x's acquaintance set at t in w seems to imply that whenever an agent imagines something, her imaginary scenario contains (representatives of) all of the individuals with whom she is acquainted. One solution to this problem is to add a 'null' individual to the domain of each world; to say that $f'(y, R)$ is the null individual at t' in w' is to say that (w', t', f') represents y (relative to R) as not existing (cf. Lewis 1983a: p. 398). Alternatively, we might re-define the notion of a multi-centered world so that the tagging function f' of a multi-centered world (w', t', f') can map elements in its domain to individuals that do not exist at t' in w'. If $f'(y, R)$ does not exist at t' in w', then (w', t', f') represents y (relative to R) as not existing. I am not sure which of these options is preferable, and in any case, I gloss over this subtlety in what follows. (Thanks to Josh Dever for discussion of this point.)

(31) Let $\{(g_1, G_1), \ldots, (g_n, G_n)\}$ be x's acquaintance set at t in w.

A multi-centered world (w', t', f') is compatible with what x imagines at t in w iff

- the domain of f' is $\{(g_1, G_1), \ldots, (g_n, G_n)\}$; and
- $f'(g_1, G_1), \ldots, f'(g_n, G_n)$ stand, at t' in w', in all the relations that x imagines that g_1 (relative to G_1), …, g_n (relative to G_n) stand in at t in w.

A *multi-centered proposition* is a set of multi-centered worlds, or the characteristic function thereof. A multi-centered proposition p is true at a multi-centered world (w, t, f) iff $(w, t, f) \in p$ (or $p(w, t, f) = 1$). An agent x imagines a multi-centered proposition p at t in w iff p is true at all the multi-centered worlds compatible with what x imagines at t in w. And we offer the following account of *de re* imagining:

(32) Agent x imagines, at t in w, that y_1 (relative to R_1), …, y_n (relative to R_n) stand in relation T iff
every multi-centered world (w', t', f') compatible with what x imagines (at t in w) is such that $f'(y_1, R_1), \ldots, f'(y_n, R_n)$ stand in relation T at t' in w'.

In the monadic case, this amounts to the following:

(33) Agent x imagines, at t in w, that y (relative to R) is F iff
every multi-centered world (w', t', f') compatible with what x imagines (at t in w) is such that $f'(y, R)$ is F at t' in w'.

Of course, the foregoing all applies, *mutatis mutandis*, to attitudes other than imagining.

One way in which the present account differs from the centered worlds approach concerns what each account takes as its basic notion of belief (imagining, etc.). For Lewis, the basic notion of belief is belief *de se*. This is the basic notion in two (related) senses: First, this is the only notion of belief that one needs when assessing a potential doxastic alternative for compatibility with what an agent believes. Second, the notion of *de*

re belief — believing something about someone relative to an acquaintance relation — is defined in terms of the notion of *de se* belief (cf. (7)). On the multi-centered approach, in contrast, the basic notion of belief is belief *de re*, or the notion of what an agent believes about an individual relative to an acquaintance relation, or, more generally, what an agent believes about a plurality of individuals-relative-to-acquaintance-relations. Again, this is the basic notion in two senses. First, it is this notion of belief that we use when we characterize the notion of doxastic compatibility (cf. (31)). Second, *de se* belief is understood as a special case of *de re* belief — it is a *de re* belief where the *res* is the agent and the relation of acquaintance is identity. Thus, we have not tried to reduce the notion of *de re* belief to something more basic; unlike Lewis's, our account of the *de re* is *non-reductive*.

While Lewis's approach might be more economical, there does not seem to be anything problematic about taking the notion of believing something about an individual relative to an acquaintance relation to be the basic, intuitive notion of belief on which to build a semantic account of attitude ascriptions. It seems to be at least as clear as the notion of *de se* belief, the notion of believing something about oneself in the first-person way. Lewis (1979, 1983a) essentially takes the *de se* notion as primitive, elucidating it by way of example. Our approach to the notion of *de re* belief is similar. Although talking of believing something about someone relative to an acquaintance relation is not an everyday way of talking, it is not hard to see how this sort of talk is intended to be taken when we reflect on cases like Quine's original Ortcutt case. And we are again on a par here with Lewis: talking of 'believing something about oneself in the first-person way' is also not an everyday way of speaking, but it can be explained by pointing to examples of the phenomenon in which the theorist is interested.[25]

25 There is more to say about the differences between the centered and the multi-centered theories, but it would take us too far afield to explore them here. One point worth mentioning is that, on the multi-centered approach (at least as I have developed it here), the content of a *de se* attitude is what Perry (1979: 45ff.) calls a "proposition of limited accessibility". When I believe that I am hungry, I believe a multi-centered proposition that only I can believe, viz.:

$$\{(w, t, f) : f(\text{DN, identity}) \text{ is hungry at } t \text{ in } w\}$$

(Since only I am identical to DN, my acquaintance set is the only one that contains the pair (DN, identity).) In this respect, the account is similar to Fregean approaches to *de se* thought cf. Frege 1918/1956, Kripke 2008. Relatedly, *de se* contents are not 'relativistic' on the multi-centered approach, the way they are on Lewis's. I hope to discuss these issues in future work.

The multi-centered approach allows us to resolve the problem of double-vision as follows. Since Ralph believes, at t in w, that Ortcutt (relative to Q) is a spy, he believes the following multi-centered proposition:

(34) $\{(w', t', f')\colon f'(\text{Ortcutt}, Q) \text{ is a spy at } t' \text{ in } w'\}$

Since he also believes, at t in w, that Ortcutt (relative to S) is not a spy, he also believes (35):

(35) $\{(w', t', f')\colon f'(\text{Ortcutt}, S) \text{ is not a spy at } t' \text{ in } w'\}$

Note that (34) and (35) are not disjoint, since there are multi-centered worlds (w', t', f') in which f'(Ortcutt, Q) is distinct from f'(Ortcutt, S). In other words, each of Ralph's doxastic alternatives (w', t', f') will contain one individual who represents Ortcutt (relative to Q) there, and another who represents Ortcutt (relative to Q) there. The first — $f'(\text{Ortcutt}, Q)$ — will be a spy at t' in w', and will have whatever other properties Ralph attributes to Ortcutt (relative to Q). The second — $f'(\text{Ortcutt}, S)$ — will not be a spy at t' in w', and will have whatever other properties that Ralph attributes to Ortcutt (relative to S).

We should also reassure ourselves that our solution to the problem of counterfactual attitudes has been preserved. As in our original problem case, suppose that Ralph is acquainted with Ortcutt only via Q at t in w, and that he imagines at t in w that he does not see Ortcutt sneaking around on the waterfront. On our present account, this means that all the multi-centered worlds (w', t', f') compatible with what Ortcutt imagines at t in w are such that f'(Ralph, identity) does not see f'(Ortcutt, Q) sneaking around on the waterfront at t' in w'. And there are, of course, multi-centered worlds that meet that condition.

Although we have said what it is for a multi-centered proposition to be true at a multi-centered world, we haven't said what it is for a belief to be true *simpliciter*. In the possible worlds framework, if x believes possible worlds proposition p in world w, her belief is true *simpliciter* iff p is true at w. In the centered worlds framework, if x believes the centered proposition p at t in w, then x's belief is true *simpliciter* iff p is true at (w, t, x). Note that in the centered worlds case, when we assess x's belief in p at t in w for truth, we check whether p is true at the centered world centered on the 'represented individual' (x) at the time and world at which the believing is taking place. This suggests the following extension to the multi-centered worlds case:

> Suppose $\{(g_1, G_1), \ldots, (g_n, G_n)\}$ is x's acquaintance set at t in w.
>
> Then if x believes the multi-centered proposition p at t in w, then x's belief is true *simpliciter* iff p is true at (w, t, f), where $f = \{((g_1, G_1), g_1), \ldots, ((g_n, G_n), g_n)\}$.

The reader can verify that this yields intuitively correct results.[26]

5 Assignment-sensitivity

We now turn to the task of giving a compositional semantics that generates multi-centered worlds truth conditions for *de re* attitude reports. Note that the multi-centered worlds account of *de re* attitudes is an account of what it is for an agent to imagine (believe, etc.) something about someone *relative to an acquaintance relation*. But, like Lewis, we can use this relativized notion to characterize an *absolute* notion of *de re* imagination (belief, etc.), as follows (cf. (13)):

(36) Agent x imagines, at t in w, that $y_1, .., y_n$ stand in relation T *simpliciter* iff
there is a set of acquaintance relations $\{R_1, \ldots, R_n\}$ such that x imagines, at t in w, that y_1 (relative to R_1), …, y_n (relative to R_n) stand in relation T.

As before, I will state the semantics so that we predict coarse truth conditions for *de re* reports. But as before, the account is compatible with the idea that the sets of acquaintance relations over which we quantify can be restricted by the context (cf. §2.3).

26 Hazen (1979) uses something similar to multi-centered worlds in giving a counterpart-theoretic semantics for a language of quantified modal logic that includes an actuality operator. Proposals closely related to Hazen's are discussed in Dorr 2010, Stalnaker 2012, Bacon forthcoming, Kment forthcoming, and Russell forthcoming. The main difference between Hazen's proposal and mine is that his "representative functions" are functions mapping individuals to individuals, whereas my tagging functions are functions mapping individual-acquaintance relation pairs to individuals.

Lewis (1983a, 1986: §4.4) offers a slightly different implementation of Hazen's basic idea, analysing metaphysical possibility in terms of a relation between sequences of possible individuals. In other work (Ninan 2008, 2010), I represent attitudinal alternatives using *sequenced worlds*, formal objects which are closely related to Lewis's sequences of possibilia (see Torre 2010 as well). There are some subtle differences between the sequenced worlds approach and the present account, most of which do not matter for present purposes.

There may be more than one way to formulate a semantic account that will predict multi-centered worlds truth conditions for *de re* reports, but the account I wish to explore has two central features. First, it treats pronouns and proper names as *assignment-sensitive*, that is, as sensitive to a variable assignment. And second, it treats attitude verbs as *assignment shifters* (unselective quantifiers).[27] I'll begin with third-person singular pronouns, expressions that are widely thought to be assignment-sensitive. After showing how the account works in that case, I will show how to extend the account to first- and second-person singular pronouns, proper names, definite descriptions, and quantifiers.

5.1 Third-person pronouns

Consider a two-dimensional semantic framework, along the lines of Kaplan 1989, in which the circumstance of evaluation contains a world, a time, and a variable assignment. A variable assignment is a (possibly partial) function from numerical indices ("1", "1462", etc.) to individuals. According to a familiar picture, third-person singular pronouns can be treated in much the same way variables are treated in logic. So suppose that pronouns like *he* always bear a numerical index at LF. Where j is a numerical index, the meaning of the pronoun is then determined by the variable assignment of the circumstance as follows (cf. Heim & Kratzer 1998: p. 241):

$$[\![\text{he}_j]\!]^{c,(w,t,g)} = g(j)$$[28]

Such a semantics allows us to predict the fact that *he* can occur bound, as it might in *Every boy loves his mother*.

This semantics also allows us to treat deictic occurrences of pronouns, if we assume that the utterance context c determines a variable assignment g_c. Consider the sentence *He is tall*. In accordance with our indexing requirement,

27 Cumming (2008) and Santorio (2012) both treat attitude verbs as assignment shifters, but offer more standard accounts of the truth conditions of *de re* ascriptions. (Cumming's main interest is in giving an account of certain kinds of *de dicto* attitude ascriptions.)

28 If g is undefined for j, then $[\![\text{he}_j]\!]^{c,(w,t,g)}$ is undefined. So technically, our entry should be as follows:

$[\![\text{he}_j]\!]^{c,(w,t,g)}$ is defined only if j is in the domain of g.
Where defined, $[\![\text{he}_j]\!]^{c,(w,t,g)} = g(j)$

I leave this out for readability in what follows.

this sentence will be represented as *He$_j$ is tall*, for some numeral j. Consider this sentence in a context c in which I utter that sentence while intending to refer to Obama by my use of *he*. To represent this, the theorist should say that c determines a variable assignment g_c that maps j to Obama.[29] Now recall the standard definition of truth at a context for two-dimensional systems of the sort we are presupposing (cf. Kaplan 1989: p. 547):

> A sentence ϕ is *true at a context c* iff $[\![\phi]\!]^{c,(w_c,t_c,g_c)} = 1$

Truth at a context *simpliciter* is defined in terms of the recursive notion of truth at a context and circumstance, aka truth at a point of evaluation. A sentence is true at a context c just in case it is true at the corresponding *proper* point of evaluation $(c,(w_c,t_c,g_c))$, which results from setting all values of the parameters in the circumstance to the corresponding values of the context. Since *He$_j$ is tall* is true relative to a point of evaluation $(c,(w,t,g))$ just in case $g(j)$ is tall at t and w, that sentence will be true relative to a context c just in case $g_c(j)$ is tall at t_c and w_c. If I intended in c to refer to Obama, $g_c(j)$ would be Obama, and my utterance would be true just in case Obama is tall at t_c in w_c. It is the truth-at-a-context conditions of *de re* ascriptions that we want to predict in what follows.

In order to do that, we need to state a lexical entry for *believes*. But in order to do *that*, we need to put in place one more definition.

> DEFINITION. Let g be a variable assignment with domain $\{i_1,..,i_n\}$; let $G = \{G_1,\ldots,G_n\}$ be a set of acquaintance relations; and let f be a tagging function.
>
> Then $\boldsymbol{g^{f,G}}$ is a variable assignment s.t. for all k $(1 \leq k \leq n)$,
>
> $g^{f,G}$ is defined for i_k iff f is defined for $(g(i_k),G_k)$;
>
> where defined, $g^{f,G}(i_k) = f(g(i_k),G_k)$.

29 Note that the context determines a variable assignment only relative to an assignment of numerical indices to index-bearing expressions. So it might be more accurate to say that, given an unindexed sentence S, the context determines a set of admissible pairs (ϕ,g) consisting of an indexed LF ϕ corresponding to S and a variable assignment g. See Büring 2005: §2.3.2 for related discussion.

Using this notation, we can state our semantics for *imagines* as follows:

> $[\![x \text{ imagines } \phi]\!]^{c,(w,t,g)} = 1$ iff
>
> there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what x imagines at t in w is such that $[\![\phi]\!]^{c,(w',t',g^{f',G})} = 1$[30]

This is actually only a first-pass entry; we will need to revise it slightly when we turn to indexicals and proper names. But this will do for now. I state the semantics for *imagines* only, but parallel accounts for the other attitude verbs can be easily inferred.

Consider a *de re* report that contains a third-person pronoun:

(37) Ralph imagines that he_2 is a spy.

Note that someone who utters (37) will typically have in mind a referent for the pronoun *he*. Suppose, for example, that I utter (37) in a context c in which I am pointing at Ortcutt. This means that the contextually determined assignment g_c will map 2 to Ortcutt. So that sentence will be true at c just in case Ralph imagines (at t_c in w_c) that Ortcutt is a spy (relative to some acquaintance relation or other).

To see how our semantics predicts this truth-at-a-context condition, first consider the truth-at-a-point-of-evaluation conditions for (37):

- $[\![\text{Ralph imagined } \text{he}_2 \text{ was a spy}]\!]^{c,(w,t,g)} = 1$ iff
- there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $[\![\text{he}_2 \text{ was a spy}]\!]^{c,(w',t',g^{f',G})} = 1$ iff
- there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $g^{f',G}(2)$ is a spy at t' in w' iff
- there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $f'(g(2), G_2)$ is a spy at t' in w'.

30 Here, and in what follows, n represents the number of numerical indices in the domain of g.

We get the truth-at-a-context conditions by taking that last line and setting all the circumstance parameters — w, t, g — to their corresponding context values, yielding:

> there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t_c in w_c is such that $f'(g_c(2), G_2)$ is a spy at t' in w'.

Since c is a context in which I utter (37) while pointing at Ortcutt, $g_c(2)$ is Ortcutt. So (37) is true at c iff:

> there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t_c in w_c is such that $f'(\text{Ortcutt}, G_2)$ is a spy at t' in w'.

In other words, (37) will be true at c just in case there is an acquaintance relation relative to which Ralph imagines that Ortcutt is a spy.

I note that the account does not require *res* movement; the *res* expression is interpreted *in situ*. There are of course other accounts in the literature which also avoid *res* movement — such as Percus & Sauerland 2003a and Anand 2006 — but the present account is more economical than those insofar as it does not require positing any covert material in the *res*-phrase.

5.2 *I* and *you*

What about other pronouns? Although philosophers tend to follow Kaplan (1989) in assuming that the first- and second-person pronouns are indexicals (expressions that receive their semantic value directly from the context parameter), certain examples suggest that these pronouns can be bound:

(38) a. Only I did my homework. (von Stechow 2002, attributed to Heim)
Bound reading ≈ "I did my homework and for all individuals x ($x \neq$ me → x didn't do x's homework)."

b. You're the only one who forgot your wallet.
Bound reading ≈ "You forgot your wallet and for all individuals x ($x \neq$ you → x didn't forget x's wallet)."

This suggests that, like their third-person cousins, first- and second-person pronouns should be treated as assignment-sensitive expressions, not as straight indexicals.

A first-pass implementation would simply attach numerical indices to these pronouns, indices whose values are determined by the variable assignment of the circumstance:

$$[\![\mathrm{I}_3]\!]^{c,(w,t,g)} = g(3)$$

$$[\![\mathrm{you}_4]\!]^{c,(w,t,g)} = g(4)$$

But this treatment overlooks the most salient feature of the interpretation of these expressions: that when they occur free, *I* refers to the speaker and *you* to the addressee. One familiar way to account for this, while still assuming an assignment-sensitive semantics for these terms, is to treat this aspect of pronominal meaning as presuppositional, where presuppositions are understood as conditions on definedness (cf. Cooper 1983, Heim 2008, Kratzer 2009).[31] I adopt a slight variation on the standard version of this approach; on my account, the semantics of *I* and *you* are given as follows:[32]

$$[\![\mathrm{I}_3]\!]^{c,(w,t,g)}\colon\ g_c(3) = x_c.\ g(3)$$

$$[\![\mathrm{you}_4]\!]^{c,(w,t,g)}\colon\ g_c(4) = a_c.\ g(4)$$

Note also that a similar treatment of third-person pronouns can be given in order to capture the semantic contribution of gender features:

$$[\![\mathrm{he}_2]\!]^{c,(w,t,g)}\colon\ g_c(2) \text{ is male at } t \text{ in } w.\ g(2)$$

What does this predict about the truth conditions of sentences like (39)?

(39) Ralph imagined that I_3 was a spy.

Before we can answer this question, we need to say a bit more about how undefinedness 'projects'. I assume that if I_3 is undefined at a point of

31 I should note that I am not entirely convinced that the presuppositional treatment of person features is the right one. The 'indexicals-as-variables' approach needs to be able to constrain the reference of free occurrences of the first- and second-person pronouns. The presuppositional approach offers one way of achieving this, but others may be possible. I adopt the presuppositional approach here simply for the sake of familiarity.

32 A note on this notation: "$[\![\mathrm{I}_3]\!]^{c,(w,t,g)}\colon\ g_c(3) = x_c.\ g(3)$" says that $[\![\mathrm{I}_3]\!]^{c,(w,t,g)}$ is defined iff $g_c(3) = x_c$, and that, if defined, $[\![\mathrm{I}_3]\!]^{c,(w,t,g)} = g(3)$.

evaluation $(c,(w,t,g))$, then any simple sentence containing it — e.g. I_3 *am a spy* — is likewise undefined at $(c,(w,t,g))$. This suggests that the intension of I_3 *am a spy* at a context c should be regarded as a partial function from circumstances to truth values. It is defined at a circumstance (w,t,g) just in case I_3 *am a spy* is defined at $(c,(w,t,g))$:

$\lambda(w,t,g).[\![\text{I}_3 \text{ am a spy}]\!]^{c,(w,t,g)} =$

$\lambda(w,t,g)\colon\ g_c(3) = x_c.\ g(3)$ is a spy in at t in w

What happens when a presupposition-bearing sentence is embedded under an attitude verb? I assume the following: if ϕ presupposes ψ, then x *imagines* ϕ presupposes x *imagines* ψ (cf. Heim 1992, Karttunen 1974). That means that an imagination ascription is defined only if the presuppositions of the complement clause are satisfied at each of the relevant imagination possibilities. In our system, this means the following:

> $[\![x \text{ imagines } \phi]\!]^{c,(w,t,g)}$ is defined iff there is a set of acquaintance relations $G = \{G_1,\ldots,G_n\}$ such that every multi-centered world (w',t',f') compatible with what x imagines at t in w is such that $[\![\phi]\!]^{c,(w',t',g^{f',G})}$ is defined.

> Where defined, $[\![x \text{ imagines } \phi]\!]^{c,(w,t,g))} = 1$ iff there is a set of acquaintance relations $G = \{G_1,\ldots,G_n\}$ such that every multi-centered world (w',t',f') compatible with what x imagines at t in w is such that $[\![\phi]\!]^{c,(w',t',g^{f',G})} = 1$.[33]

Here is what this means for our target sentence, (39):

> $[\![\text{Ralph imagined that I}_3 \text{ was a spy}]\!]^{c,(w,t,g)}$ is defined iff there is a set of acquaintance relations $G = \{G_1,\ldots,G_n\}$ such that every multi-centered world (w',t',f') compatible with what Ralph imagines at t in w is such that $g_c(3) = x_c$.

33 A peculiarity of this semantics is that we existentially quantify over a set of acquaintance relations in stating the definedness condition, and then existentially quantify *again* over such a set in stating the truth conditions. But as far as I can tell, this doesn't lead to any untoward predictions.

Note that this reduces to the following:

> $[\![\text{Ralph imagined that } I_3 \text{ was a spy}]\!]^{c,(w,t,g)}$ is defined iff $g_c(3) = x_c$.

In other words, (39) is defined at a point of evaluation $(c,(w,t,g))$ iff the complement clause I_3 *was a spy* is defined at c.

This may seem odd: were we not following Karttunen and Heim in thinking that if ϕ presupposes ψ, then *x imagines* ϕ presupposes *x imagines* ψ? And yet here we have a case in which ϕ presupposes ψ, but in which *x imagines* ϕ also presupposes ψ. Although this might seem puzzling, nothing has actually gone wrong. The result in this particular case is due entirely to the fact that the particular presupposition that we associated with I_3 places no constraint on the circumstance parameter of the point of evaluation, only on the context parameter — we have an 'indexical presupposition'. What is more important is whether our prediction of the definedness conditions for (39) are correct; it seems to me that they are.[34]

We then predict the following truth-at-a-point conditions for (39):

> Where defined, $[\![\text{Ralph imagined that } I_3 \text{ was a spy}]\!]^{c,(w,t,g)} = 1$ iff there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $f'(g(3), G_3)$ is a spy at t' in w'.

If the semantic value of the sentence is defined at context c, then $g_c(3)$ will be the speaker of c. Where defined, the sentence will be true at c iff

> there is a set of acquaintance relations G such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $f'(g_c(3), G_3)$ is a spy at t' in w'.

If the sentence is defined at c, $g_c(3)$ is the speaker of c. So the sentence will be true at c just in case Ralph imagines that the speaker of c is a spy (relative to some acquaintance relation or other), which is the right result.

34 Note that our semantics predicts that the presuppositions of pronouns alway project to the global context (unless the pronoun is embedded under a monstrous operator). Given the empirical facts discussed in Heim 2008, this may be a welcome prediction.

5.3 Proper names

The idea that names are assignment-sensitive has been defended in a number of places (Büring 2005, Cumming 2008, Dever 1998, Elbourne 2005, Geurts 1997, Yagisawa 1984). Some of the authors of these works are motivated by the observation that names can be bound, as in (40):

(40) If a child is christened "Goofy", and the CEO of Disney hears about it, he'll sue Goofy's parents. (Geurts 1997: p. 322)

Intuitively, (40) has a reading on which it is true just in case:

> For every x, if (x is a child christened "Goofy" and the CEO hears that x is christened "Goofy"), then (the CEO sues x's parents).

If that's right, then it appears that the occurrence of *Goofy* in the consequent of (40) is bound, presumably by material in the antecedent (or perhaps by the conditional operator).

Cumming (2008: p. 535) gives another example:

(41) There is a gentleman in Hertfordshire by the name of "Ernest". Ernest is engaged to two women.

As Cumming notes, this pair of sentences appears to be true just in case:

> There is an x such that (x is a gentleman in Hertfordshire named "Ernest" and x is engaged to two women).

If that's right, then the occurrence of *Ernest* in the second sentence of the discourse in (41) appears to be bound by material found in the first sentence.

Now I won't offer a semantic theory that predicts the truth conditions of (40) and (41). That would take us to far afield into the semantics of indefinite descriptions and conditionals. But I will assume that any such theory will need to invoke the idea that proper names are assignment-sensitive, as the authors mentioned above suggest.

How would such a theory of proper names go? A natural idea is to generalize our earlier treatment of pronouns. On this approach, we assume that each proper name bears a numerical index at LF. We place a definedness condition on the semantic value of the name-plus-index to the effect that it is defined at a point of evaluation $(c, (w, t, g))$ only if g_c maps the index to the intuitive referent of the name. Here is the idea:

$[\![\text{Ortcutt}_2]\!]^{c,(w,t,g)}$: $g_c(2)$ = Ortcutt. $g(2)$

Given this semantics for names, and our analysis of attitude verbs, it's not hard to see that the resulting account generates the correct multi-centered worlds truth conditions for reports like (42):

(42) Ralph imagined that Ortcutt_2 was a spy.

Since our treatment of this case isn't substantially different from our treatment of reports with pronouns in the scope of the attitude verb, I'll simply record the predicted truth-at-a-point conditions of this sentence:

- $[\![\text{Ralph imagined that Ortcutt}_2 \text{ was a spy}]\!]^{c,(w,t,g)}$ is defined iff there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $g_c(2)$ = Ortcutt.
- Where defined, $[\![\text{Ralph imagined that Ortcutt}_2 \text{ was a spy}]\!]^{c,(w,t,g)} = 1$ iff there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $f'(g(2), G_2)$ is a spy at t' in w'.

Let c be a context at which this sentence is defined, i.e. $[\![(42)]\!]^{c,i_c}$ is defined. When we assess this sentence for truth at c, the assignment of the circumstance will be identical to the assignment of the context g_c, which maps 2 to Ortcutt. So the sentence will be true at c just in case Ralph imagines that Ortcutt is a spy (relative to some acquaintance relation or other).

5.4 Definite descriptions and quantifiers

As is well-known, definite descriptions occurring in the scope of an attitude verb often give rise to two readings. Consider (43), for example:

(43) Ralph imagined that the man who orchestrated the GM bailout was flying a kite.

Suppose Ralph imagined that Steve Rattner, his friend from the tennis club, was flying a kite. Ralph has no idea that Rattner is the man who orchestrated the GM bailout, and he's not imagining that Rattner is the man who orchestrated the bailout either. Now Rattner *is* the man who orchestrated the GM bailout. So is (43) true or false? It is natural to say that it is and it isn't, i.e.

the sentence has two readings, one on which it is false and one on which is true. The reading on which it is false is the *de dicto* reading; the reading on which it is true is the *de re* reading. For the *de re* reading to be true, Ralph needn't think of Rattner as the man who orchestrated the bailout. We want an account of (43) that will predict both these readings.

Although there is evidence that definite descriptions can be bound, our analysis doesn't actually require us to treat definite descriptions as assignment-sensitive. We can in fact generate both readings given fairly familiar (if not uncontroversial) assumptions about the semantics of definite descriptions. Still, our account of *de re* ascriptions with definites fits in with our general theme of assignment-sensitivity, since the account exploits the assignment-sensitivity of the *trace* left behind by a wide-scoped definition description.

The account has two main parts. First, we treat the definite article à la Frege and Strawson, i.e. as presupposing, rather than asserting, existence and uniqueness. In our system, this means giving the following semantics for the definite article (letting k be the type of our circumstances, i.e. triples consisting of a world, a time, and a variable assignment):

- $[\![\text{the}]\!]^{c,(w,t,g)} = \lambda F_{(k,et)}\colon \exists!x(F(w,t,g)(x) = 1).\ \iota x F(w,t,g)(x) = 1$
- $[\![\text{the man who orchestrated the GM bailout}]\!]^{c,(w,t,g)}$ is defined iff $\exists!x$ (x orchestrated the GM bailout at t in w).

  Where defined, $[\![\text{the man who orchestrated the GM bailout}]\!]^{c,(w,t,g)} =$ the man who orchestrated the GM bailout at t in w.

Second, we shall suppose that the *de dicto/de re* ambiguity is generated by a scope ambiguity: when the description has narrow scope with respect to the attitude verb, the *de dicto* reading results; when it has wide scope, the *de re* reading results.[35]

To illustrate this, begin with the *de dicto* reading. Computing this reading is entirely straightforward, and so I here record only the predicted truth-at-a-point conditions:

35 Analysing the *de dicto/de re* ambiguity in terms of scope is controversial, though not wholly unmotivated (see Keshet 2008 for a recent discussion of the issues involved). The broad approach taken in this paper is, I think, compatible with an alternative treatment of the interaction between definites and attitude verbs, one which doesn't involve appeal to scope. But I leave working out the details of this as a matter for future work.

- $[\![$Ralph imagined that the man who orchestrated the GM bailout was flying a kite$]\!]^{c,(w,t,g)}$ is defined iff there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that there is a unique man who orchestrated the GM bailout at t' in w'.
- Where defined, $[\![$Ralph imagined that the man who orchestrated the GM bailout was flying a kite$]\!]^{c,(w,t,g)} = 1$ iff there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that the unique man who orchestrated the GM bailout at t' in w' was flying a kite at t' in w'.

The *de re* reading of (43) is generated when that sentence has something like the following LF:

> The man who orchestrated the GM bailout [λ_2 Ralph imagines that x_2 was flying a kite]

Note that x_2 is a trace left behind by the definite, which has moved above the attitude verb. Importantly, the semantic value of a trace is determined by the variable assignment of the circumstance.

It's easiest to see how the truth-at-a-point conditions of the *de re* reading are computed if we split the calculation into two parts, the definite (which we've already computed above) and the property abstract. We compute the value of the property abstract thus:[36]

- $[\![\lambda_2$ Ralph imagines that x_2 was flying a kite$]\!]^{c,(w,t,g)} =$
- $\lambda y.[\![$Ralph imagines that x_2 was flying a kite$]\!]^{c,(w,t,g^{y/2})} =$
- (λy. there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $[\![x_2$ was flying a kite$]\!]^{c,(w',t',g^{y/2^{f',G}})} = 1$) =
- (λy. there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $g^{y/2^{f',G}}(2)$ was flying a kite at t' in w') =

36 For simplicity, I ignore the possibility that x_2 *was flying a kite* might be undefined at $(c, (w, t, g))$.

- λy. there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $f'(g^{y/2}(2), G_2)$ was flying a kite at t' in w'

Since $g^{y/2}(2)$ is just y, the last line reduces to:

- λy. there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') compatible with what Ralph imagines at t in w is such that $f'(y, G_2)$ was flying a kite at t' in w'

Now suppose that Steve Rattner is the man who organized the GM bailout at t in w. Then (43) will be true at $(c, (w, t, g))$ iff:

> there is a set of acquaintance relations $G = \{G_1, \ldots, G_n\}$ such that every multi-centered world (w', t', f') is compatible with what Ralph imagines at t in w is such that $f'(\text{Rattner}, G_2)$ was flying a kite at t' in w'

And this is exactly what we want.

This last calculation shows how *de re* readings of quantificational examples would work, like the sentence with which we began the paper, repeated here as (44):

(44) Ralph believes that someone is a spy.

The *de re* reading can again be generated by scoping out the quantifier:

> Someone [λ_2 Ralph believes that x_2 is a spy]

To calculate the truth conditions of this, one applies the meaning of the quantifier to the meaning of the property-abstract which we computed above. Again, the assignment-sensitivity of the trace left behind by the quantifier plays a crucial role in generating the right reading. The *de dicto* reading results from interpreting the quantifier *in situ*.

References

Abusch, Dorit. 1997. Sequence of tense and temporal *de re*. *Linguistics and Philosophy* 20(1). 1–50. http://dx.doi.org/10.1023/A:1005331423820.

Aloni, Maria. 2001. *Quantification under conceptual covers*. University of Amsterdam PhD thesis. http://staff.science.uva.nl/~maloni/tesi/tesi.pdf.

Anand, Pranav. 2006. De De Se. MIT PhD thesis. http://hdl.handle.net/1721.1/37418.

Asher, Nicholas. 1987. A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy* 10(2). 125–197. http://dx.doi.org/10.1007/BF00584317.

Bacon, Andrew. forthcoming. Representing counterparts. *Australasian Journal of Logic*. http://philpapers.org/rec/BACRC.

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

Castañeda, Hector-Neri. 1966. 'He': A study in the logic of self-consciousness. *Ratio* 8. 130–157.

Castañeda, Hector-Neri. 1967. Indicators and quasi-indicators. *American Philosophical Quarterly* 4. 85–100. http://www.jstor.org/stable/20009232.

Chalmers, David. 1996. *The conscious mind*. New York: Oxford University Press.

Chalmers, David. 2002. The components of content. In David Chalmers (ed.), *Philosophy of mind: Classical and contemporary readings*. New York: Oxford University Press.

Cooper, Robin. 1983. *Quantification and syntactic theory*. Dordrecht: Reidel.

Cresswell, Maxwell J. & Arnim von Stechow. 1982. *De re* belief generalized. *Linguistics and Philosophy* 5(4). 503–535. http://dx.doi.org/10.1007/BF00355585.

Crimmins, Mark & John Perry. 1989. The prince and the phone booth: Reporting puzzling beliefs. *Journal of Philosophy* 86(12). 685–711. http://dx.doi.org/10.2307/2027014.

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

Dever, Josh. 1998. *Variables*. University of California, Berkeley PhD thesis. https://webspace.utexas.edu/deverj/personal/papers/Variables.pdf.

Dorr, Cian. 2010. How to be a modal realist. Unpublished manuscript, Oxford University. http://users.ox.ac.uk/~sfop0257/papers/ModalRealism.pdf.

Elbourne, Paul. 2005. *Situations and individuals*. MIT Press.

Frege, Gottlob. 1918/1956. The thought: A logical inquiry. *Mind* 65(259). 287–311. http://dx.doi.org/10.1093/mind/65.1.289.

Geurts, Bart. 1997. Good news about the description theory of names. *Journal of Semantics* 14(4). 319–348. http://dx.doi.org/10.1093/jos/14.4.319.

Hawthorne, John & David Manley. 2012. *The reference book*. Oxford: Oxford University Press.

Hazen, Allen. 1979. Counterpart-theoretic semantics for modal logic. *The Journal of Philosophy* 76(6). 319–338. http://dx.doi.org/10.2307/2025472.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221. http://dx.doi.org/10.1093/jos/9.3.183.

Heim, Irene. 1994. Comments on Abusch's theory of tense. *Ellipsis, Tense, and Questions*. http://www.semanticsarchive.net/Archive/Tg5ZmI1N/heim-comments-Abusch-94.pdf.

Heim, Irene. 2008. Features on bound pronouns. In Daniel Harbour, David Adger & Susana Bèjar (eds.), *Phi-theory: Phi-features across modules and interfaces*. Oxford: Oxford University Press.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hintikka, Jaakko. 1962. *Knowledge and belief*. Ithaca, N.Y.: Cornell University Press.

Jackson, Frank. 1998. *From metaphysics to ethics*. Oxford: Clarendon Press.

Jeshion, Robin (ed.). 2010. *New essays on singular thought*. Oxford: Oxford University Press.

Kamp, Hans. 1990. Prolegomena to a structural account of belief and other attitudes. In C. Anthony Anderson & Joseph Owens (eds.), *Propositional attitudes: The role of content in logic, language, and mind*, 27–90. Stanford: CSLI Publications.

Kaplan, David. 1968. Quantifying in. *Synthese* 19(1-2). 178–214. http://dx.doi.org/10.1007/BF00568057.

Kaplan, David. 1989. Demonstratives. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes from Kaplan*, 481–563. New York: Oxford University Press.

Karttunen, Lauri. 1974. Presuppositions and linguistic context. *Theoretical Linguistics* I. 181–194. http://dx.doi.org/10.1515/thli.1974.1.1-3.181.

Keshet, Ezra. 2008. *Good intensions: Paving two roads to a theory of the* de re/de dicto *distinction*. MIT PhD thesis. http://hdl.handle.net/1721.1/45622.

Kment, Boris. forthcoming. Individuals, supervenience, and fundamentality. *Philosophical Review*.

Kratzer, Angelika. 2009. Making a pronoun: Fake indexicals as windows into the properties of pronouns. *Linguistic Inquiry* 40(2). 187–237. http://dx.doi.org/10.1162/ling.2009.40.2.187.

Kripke, Saul. 1980. *Naming and necessity*. Oxford: Blackwell.

Kripke, Saul. 2008. Frege's theory of sense and reference: Some exegetical notes. *Theoria* 74. 181–218. http://dx.doi.org/10.1111/j.1755-2567.2008.00018.x. Reprinted in Kripke 2011: pp. 254-291. Page references are to the 2011 reprint.

Kripke, Saul. 2011. *Philosophical troubles, volume 1*. New York: Oxford University Press.

Lewis, David K. 1979. Attitudes *de dicto* and *de se*. *Philosophical Review* 88(4). 513–543. http://dx.doi.org/10.1093/0195032047.003.0010. Reprinted in Lewis 1983b: pp. 133-159. Page references are to the 1983 reprint.

Lewis, David K. 1981. What puzzling Pierre does not believe. *Australasian Journal of Philosophy* 59. 283–289. http://dx.doi.org/10.1080/00048408112340241. Reprinted in Lewis 1999: pp. 408-417. Page references are to the 1999 reprint.

Lewis, David K. 1983a. Individuation by acquaintance and by stipulation. *Philosophical Review* 92(1). 3–32. http://dx.doi.org/10.2307/2184519. Reprinted in Lewis 1999: pp. 373-402. Page references are to the 1999 reprint.

Lewis, David K. 1983b. *Philosophical papers, volume 1*. New York: Oxford University Press.

Lewis, David K. 1986. *On the plurality of worlds*. Oxford: Blackwell.

Lewis, David K. 1999. *Papers in epistemology and metaphysics*. Cambridge: Cambridge University Press.

Maier, Emar. 2006. *Belief in context: Towards a unified semantics of* de re *and* de se *attitude reports*. Radboud University Nijmegen PhD thesis. http://www.ru.nl/ncs/~emar/diss/.

Maier, Emar. 2009. Presupposing acquaintance: A unified semantics for *de dicto*, *de re*, and *de se* belief reports. *Linguistics and Philosophy* 32(5). 429–474. http://dx.doi.org/10.1007/s10988-010-9065-2.

Ninan, Dilip. 2008. *Imagination, content, and the self*. MIT PhD thesis. http://hdl.handle.net/1721.1/45621.

Ninan, Dilip. 2010. *De se* attitudes: Ascription and communication. *Philosophy Compass* 5(7). 551–567. http://dx.doi.org/10.1111/j.1747-9991.2010.00290.x.

Percus, Orin & Uli Sauerland. 2003a. On the LFs of attitude reports. *Sinn und Bedeutung 7*. 228–242. http://semanticsarchive.net/Archive/jI4NmJlY/PercusSauerland03a.pdf.

Percus, Orin & Uli Sauerland. 2003b. Pronoun movement in dream reports. *Proceedings of NELS 33*. http://semarch.linguistics.fas.nyu.edu/Archive/jQoMDc2M/PercusSauerland03b.pdf.

Perry, John. 1977. Frege on demonstratives. *Philosophical Review* 86(4). 474 – 497. http://dx.doi.org/10.2307/2184564. Reprinted in Perry 1993: pp. 3-32. Page references are to the 1993 reprint.

Perry, John. 1979. The problem of the essential indexical. *Noûs* 13(1). 3–21. http://dx.doi.org/10.2307/2214792. Reprinted in Perry 1993: pp. 33-52. Page references are to the 1993 reprint.

Perry, John. 1993. *The problem of the essential indexical and other essays*. New York: Oxford University Press.

Quine, W.V.O. 1956. Quantifiers and propositional attitudes. *Journal of Philosophy* 53(5). 177–187. http://dx.doi.org/10.2307/2022451. Reprinted in Quine 1976: pp. 185-196.

Quine, W.V.O. 1976. *The ways of paradox and other essays*. Cambridge, MA: Harvard University Press.

Richard, Mark. 1990. *Propositional attitudes: An essay on thoughts and how we ascribe them*. New York: Cambridge University Press.

Russell, Jeff Sanford. forthcoming. Actuality for counterpart theorists. *Mind*. http://users.ox.ac.uk/~sfop0399/actuality.web.pdf.

Santorio, Paolo. 2012. Reference and monstrosity. *Philosophical Review* 121(3). 359–406. http://dx.doi.org/10.1215/00318108-1574427.

Schlenker, Philippe. 1999. *Propositional attitudes and indexicality: A cross-categorial approach*. MIT PhD thesis.

Schlenker, Philippe. 2002. Indexicality, logophoricity, and plural pronouns. In J. Lecarme (ed.), *Research in Afroasiatic grammar 2*, 409–428. Amsterdam: Benjamins. http://www.semanticsarchive.net/Archive/Dg4ODI2M/Indexicality.html.

Schlenker, Philippe. 2004. Person and binding (a partial survey). *Italian Journal of Linguistics/Rivista di Linguistica* 16(1). 155–218. http://www.linguistics.ucla.edu/people/schlenker/Person.pdf.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics 9: Pragmatics*, 315–332. New York: Academic Press.

Stalnaker, Robert. 1981. Indexical belief. *Synthese* 49(1). 129–149. http://dx.doi.org/10.1093/0198237073.003.0008. Reprinted in Stalnaker 1999: pp. 130-149. Page references are to the 1999 reprint.

Stalnaker, Robert. 1988. Belief attribution and context. In Robert Grimm & Daniel Merrill (eds.), *Contents of thought*. Tucson: University of Arizona Press. http://dx.doi.org/10.1093/0198237073.003.0009.

Stalnaker, Robert. 1999. *Context and content*. New York: Oxford University Press.

Stalnaker, Robert. 2008a. Comment on “Stalnaker on Sleeping Beauty," Brain Pains Blog. http://brainpains.com/wordpress/?p=54.

Stalnaker, Robert. 2008b. *Our knowledge of the internal world*. Oxford: Oxford University Press.

Stalnaker, Robert. 2012. *Mere possibilities*. Princeton, NJ: Princeton University Press.

von Stechow, Arnim. 1982. Structured propositions. Unpublished manuscript. http://www.sfs.uni-tuebingen.de/~astechow/Aufsaetze/Structured%20Prop%201.pdf.

von Stechow, Arnim. 2002. Binding by verbs: Tense, person and mood under attitudes. Unpublished manuscript. http://philpapers.org/rec/VONBBV.

Strawson, Peter. 1959. *Individuals: An essay in descriptive metaphysics*. Garden City, N.Y.: Doubleday.

Torre, Stephan. 2010. Centered assertion. *Philosophical Studies* 150(1). 97–114. http://dx.doi.org/10.1007/s11098-009-9399-1.

Yagisawa, Takashi. 1984. Proper names as variables. *Erkenntnis* 21(2). 195–208. http://dx.doi.org/10.1007/BF00166385.

Yanovich, Igor. 2011. The problem of counterfactual *de re* attitudes. *Semantics and Linguistic Theory* 21. http://elanguage.net/journals/salt/article/view/21.56.

Dilip Ninan
Department of Philosophy
Tufts University
Miner Hall, Second Floor
Medford, MA 02155, USA
dilip.ninan@tufts.edu
www.dilipninan.org