Presupposition Projection and the Semantics of Attitude Verbs

IRENE HEIM
MIT

Abstract

Karttunen observed that, if the complement of an attitude sentence presupposes p, then that sentence as a whole presupposes that the attitude-holder believes p. I attempt to derive some representative instances of this generalization from suitable assumptions about the lexical semantics of attitude predicates. The enterprise is carried out in a framework of context change semantics, which incorporates Stalnaker's suggestion that presupposition projection results from the stepwise fashion in which information is updated in response to complex utterances. The empirical focus is on predicates of desire and on the contribution of counterfactual mood.

1 INTRODUCTION

How are presuppositions projected in propositional attitude sentences? For example, given that *Patrick sells his cello* presupposes that Patrick owns a cello, what does (1) presuppose?

(1) Patrick wants to sell his cello.

At first sight, (1) likewise seems to presuppose that Patrick owns a cello. But then, it can also appear without contradiction in the context of (2).

(2) Patrick is under the misconception that he owns a cello, and he wants to sell his cello.

Karttunen (1973b, 1974) concludes that (1) presupposes, not that Patrick owns a cello, but rather that *Patrick believes* he owns a cello. This projection behavior is not peculiar to the verb *want*, but generalizes, according to Karttunen, to all other non-factive verbs of propositional attitude. (He cites *believe*, *think*, *expect*, *fear*, *intend*, *suspect*, *assume*, and *hope* as further examples (1973b: 4). Explicitly excluded, beside factives, are verbs of saying.) They are all subject to the following rule (his (21) (1974: 189) with trivial changes).

(3) If σ is a verb of propositional attitude, then a context c satisfies the presuppositions of '$a\sigma\phi$' only if $B_a(c)$ satisfies the presuppositions of ϕ; where '$B_a(c)$' stands for the set of beliefs attributed to a in c.

Together with the rule for *and* (1974: 185):

(4) Context c satisfies the presuppositions of 'ϕ and ψ' just in case
    (i) c satisfies the presuppositions of ϕ, and
    (ii) the context that results from c by the assertion of ϕ satisfies the presuppositions of ψ.

(3) accounts for the intuition that (2) as a whole presupposes nothing. To presuppose nothing means to place no particular requirements on the initial context; in other words, a sentence presupposes nothing iff every possible context satisfies its presuppositions. Given (3) and (4), this is predicted for (2): whatever the initial context may have been like, the first conjunct creates from it an intermediate context in which Patrick is attributed the belief that he owns a cello, and that intermediate context thus satisfies the presuppositions of the second conjunct.

But what about the intuition that (1) in isolation commits the speaker to Patrick's owning a cello? Karttunen speculates that this is attributable to

some additional conversational principle to the effect that, unless it has been indicated otherwise, [Patrick] can be assumed to share the speaker's beliefs. In other words, there is a natural spill-over from [c] to [$B_{Patrick}(c)$]. Consequently, in situations where nothing has been said about [Patrick's] beliefs, one tends to think that, if the presuppositions of [(1)] are satisfied, they are satisfied by virtue of the speaker's tacit assumption that [Patrick] shares his beliefs. (1973b: 6)

I think that Karttunen's proposal was basically right, in its description of the facts as well as in its theoretical conception.[1] In fact, the present article does nothing more than spell it out in somewhat greater precision. This should make it easier to assess its merits and the objections against it. Section 2 introduces the theoretical framework. Sections 3 and 4 examine the semantics of verbs of belief and of various types of desire verbs in reports of realistic as well as counterfactual desires. Section 5 elaborates a bit more on the reasons why attitude verbs should superficially appear to be holes rather than filters.

In the examples below, I will use a variety of presupposition triggers, in particular definite descriptions (to which I here give a classical, Fregean, analysis) and *too*. I assume, perhaps simplistically, that for the purposes of this paper there is no relevant difference between the various kinds of triggers. I have nothing original to say about where presuppositions come from in the first place, what the set of triggers is, and what presupposition exactly each trigger contributes.

2 PRESUPPOSITION PROJECTION IN CONTEXT CHANGE SEMANTICS

The framework I will employ, basically a radical elaboration of ideas of Robert Stalnaker, is characterized by four central assumptions.[2] First, the meaning of a sentence is its *context change potential* (CCP). (By 'sentence', I mean not just a string of words, but a full structural description at the level of Logical Form (LF).)[3] A CCP is a function from contexts to contexts. Contexts are here identified with states of information, which in turn are construed as sets of possible worlds,[4] and the change effected by the CCP of a sentence consists of updating that information by what the sentence says. Second, not only complete (matrix) sentences have context change potentials, but so do embedded sentences down to atomic clauses, and the CCPs of complex sentences are compositionally determined by the CCPs of their constituents. Third, the presuppositions of a sentence are requirements on the context, that is, they determine which contexts its CCP can be applied to. Whenever a sentence presupposes something, it must be evaluated in a context that already entails that presupposition.[5] These requirements are uncancellable; under certain conditions, a context may be fixed up to meet them, but never the other way round, i.e. never is the requirement waived or weakened to make it more easily met by a given context.[6] Fourth (as a consequence of the previous three assumptions, and as already urged by Stalnaker (1973, 1974) and Karttunen (1974)), the phenomena of so-called presupposition projection are just a by-product of the way the CCP of a complex sentence is composed from the CCPs of its parts.

Let me illustrate this with concrete examples and at the same time make it a bit more precise. Suppose (unrealistically) we start with the 'empty'[7] context, where nothing is presupposed yet. This is W, the set of all possible worlds. Imagine that in this context W, there occurs a (successful) assertion of the atomic sentence *it is raining*. The result will be a new context, a subset of W, which contains just those worlds where it is raining. More generally, the CCP of *it is raining* is the instruction to conjoin (that is: intersect) whatever the current context may be with the proposition that it is raining. I use the notation '$c + \phi$' to designate the result of executing the CCP of LF ϕ on the context c.[8] The CCP of (the LF of) *it is raining* can thus be defined as in (5).

(5) For any context c, c + *it is raining* $= \{w \in c: \text{it is raining in } w\}$.

The CCPs of complex sentences are determined compositionally on the basis of the CCPs of their parts, so for truthfunctional connectives, for instance, we have semantic rules like the following (where \ is set-theoretic complementation).

(6) c + *not* $\phi = c\backslash(c + \phi)$.

If you apply the CCP of *not* [*it is raining*] (the LF of *it isn't raining*) to W, what you get by (6) and (5) is the set of all worlds in which it isn't raining.

Actually, things are a bit more complicated, since so far we have disregarded presuppositions (taking *it is raining*, simplifying perhaps, to be an example of a sentence that presupposes nothing). The definition of the CCP of a sentence is supposed to encode not just its content but also its presupposition. The CCP of a sentence without any presupposition will be a *total* function from contexts to contexts (like the one defined in (5)), but in general CCPs are partial: they are defined only for those contexts that satisfy the presuppositions of the sentence in question. For example, *John's cat is hungry* presupposes that John has a unique cat, and this is reflected in the fact that the CCP of this sentence is only defined for contexts that entail that John has a unique cat. (The entailment relation between contexts is the subset relation.)

(7) c + *John's cat is hungry* is defined iff
   c ⊆ {w: John has a unique cat in w};
   where defined, c + *John's cat is hungry*
   — {w ∈ c: John has a hungry cat in w}

'Presupposition projection', according to this theory, arises from the way the definedness conditions of the CCPs of elementary sentences affect those of the CCPs of bigger constituents. For instance, (6) is really an abbreviated version of the following more explicit rule:

(8) c + *not* ϕ is defined just in case c + ϕ is,
   in which case c + *not* ϕ — c\(c + ϕ).

and so the combined effect of (7) and (8) is to predict that *not* [*John's cat is hungry*] (presumably one of the Logical Forms of *John's cat isn't hungry*) also presupposes that John has a unique cat. More generally, (8) predicts negation to be a 'hole' in the sense of Karttunen (1973a).

Notice that the spelled-out rule in (8) is, in a sense, fully recoverable from the abbreviated version of (6):[9] the added top line states just what it takes for the expression to the right of the equation sign below to denote a context—no more and no less. This is always so when we are dealing with the lexical entry of an item (such as here *not*) that doesn't contribute any presuppositions of its own. Only items that are themselves presupposition-triggers have in their entries additional, non-recoverable, definedness conditions. The fuller rule format of (8), while more explicit and therefore easier to use in proofs, has the disadvantage of superficially obscuring the difference. Below I will often use a compromise between the two formats: include the recoverable conditions, but in brackets.

3 BELIEF REPORTS

Now what would it mean to give an account within this framework of presupposition projection in attitude reports? Well, the central task is evidently to specify appropriate lexical entries for the attitude predicates, i.e. to give appropriate definitions of the CCPs of sentences of the form '*a believes ϕ*', '*a wants ϕ*', etc. Once these CCPs are defined, the projection behavior of presuppositions originating with the complement sentence ϕ is thereby determined. So what we must do is write definitions of the form (9).

(9) For any context c, c + *a believes ϕ* is defined only if . . .
    Where defined, c + *a believes ϕ* = . . .

Let us begin by recapitulating the standard possible worlds semantics of *believe*, as found, e.g., in Hintikka (1969). A sentence like (10)

(10) John believes that it is raining.

is true in a world w iff it is raining in every world w' that is doxastically accessible for John to w. What does 'doxastically accessible' mean? It means this: world w' is doxastically accessible for person x to world w iff w' is compatible with the beliefs that x holds in w. This familiar analysis is our starting point, and we now try to recast it faithfully in our context–change framework.

First, a technical convenience: accessibility relations (binary relations among possible worlds) correspond one-to-one to accessibility assignments (functions from worlds to sets of worlds):[10]

(11) Let $R \subseteq W \times W$. Then f_R is that function from W to $\mathscr{P}(W)$ such that, for any $w \in W$, $f_R(w) = \{w' \in W: wRw'\}$.

For instance, to the relation of doxastic accessibility for John corresponds the following function Dox_J ('Dox' for 'doxastic' and 'J' for 'John'):

(12) For any $w \in W$,
    $Dox_J(w) = \{w' \in W: w'$ conforms to what John believes in $w\}$.

The choice, has, of course, no substantive import, but we will save space in our CCP-definitions below by directly referring to these accessibility functions instead of the corresponding accessibility relations. Notice that the values of · accessibility functions are the same kind of thing as contexts, namely sets of possible worlds, and as such are suitable arguments for the CCPs of sentences—a fact that will be exploited in our rules below.

So how does the assertion of a belief-sentence like (10) affect the context? What information does it convey, what possibilities does it rule out? According

to the standard analysis just sketched, (10) tells us about the world we are in that it is a world accessible from which (more precisely: doxastically accessible for John from which) are only worlds in which it is raining. In other words, (10) informs us that we are in a world w such that it rains in every element of $Dox_J(w)$. So the CCP of (10) has to be an instruction to eliminate from the original context all but the worlds which fulfill this condition on w. It must be this:

(13) For any c, c + *John believes it is raining*
  = {w ∈ c: for every w' ∈ $Dox_J(w)$, it is raining in w'}.

Now we must figure out how this CCP is determined compositionally, by the interaction of a general rule for structures of the form '*α believes φ*' on the one hand and the CCP of the complement *it is raining* on the other. A first step towards isolating the contribution of each is to express the condition that it rains throughout $Dox_J(w)$ in terms which make explicit reference to the CCP of *it is raining*. It turns out (given (5) above) that the following equivalence holds:

(14) For any set $X \subseteq W$:
  it is raining in every w' ∈ X iff X + *it is raining* = X.

In other words, the sets of worlds throughout which it is raining are precisely those that map onto themselves under the CCP of *it is raining*. Why? Because if it is already raining in every element of a set of worlds X, then eliminating from X any non-rain-worlds won't change it. Whereas, if a set does become genuinely smaller by eliminating non-rain-worlds from it, then it must be a set which didn't already have rain throughout. Thus (13) becomes (15).

(15) For any c, c + *John believes it is raining*
  = {w ∈ c: $Dox_J(w)$ + *it is raining* = $Dox_J(w)$}.

From this it is easy to generalize to arbitrary complements (and subjects):[11]

(16) For any c, c + *α believes φ* = {w ∈ c: $Dox_α(w)$ + φ = $Dox_α(w)$}

The general format of this rule will recur elsewhere, and it will be more transparent if we use an abbreviation: if c is any context, φ any LF, let 'c + φ = same' express the condition that c + φ = c. So we can render (16) as (17).

(17) For any c, c + *α believes φ* = {w ∈ c: $Dox_α(w)$ + φ = same}.

For arbitrary choices of φ, of course, we can no longer take the well-definedness of '$Dox_α(w)$ + φ' for granted. For certain choices of φ, α and w it might happen that $Dox_α(w)$ is not in the domain of the CCP of φ. In other words, (17) implicitly contains a definedness condition that is brought out into the open in the following fuller rendition:

(18) For any context c,
[c + α *believes* φ is defined iff Dox_α(w) + φ is defined for each w ∈ c].
Where defined, c + α *believes* φ = {w ∈ c: Dox_α(w) + φ = same}.

Now our first proposal is in place and we can look at its predictions about presupposition projection. If the complement φ in (18) has non-trivial presuppositions, i.e. a genuinely partial CCP, then what does this imply for the CCP of α *believes* φ? The answer can be read right off (18): if the CCP of φ makes non-trivial demands on its input context, then so does the CCP of α *believes* φ. If the CCP of φ is defined only for contexts that entail a certain proposition p, then the CCP of α *believes* φ is defined only for those c all of whose elements w map onto Dox_α(w) that entail p. Only for those c, in other words, which entail that α believes p. (Recall the definition in (12): 'Dox_α(w) entails p' means nothing more and nothing less than that α believes p in w.)

What we predict, then, is simply (a special case of) Karttunen's generalization: if φ presupposes p, then α *believes* φ presupposes that α believes p. We thus expect to be able to account at least for the data that most directly supported Karttunen's view, e.g. the fact that a sequence of two belief reports in which the content of the complement of the first entails the presupposition of the complement of the second makes a smooth discourse with no presuppositions as a whole. Let us calculate through an example of this kind to see exactly how this works.

(19) John believes that Mary is here, and he believes that Susan is here too.

Before we can get started, I must fill in a brief sketch of my treatment of *too*. Relying on Kripke,[12] I assume that *too* is implicitly deictic or anaphoric, sort of like *in addition to x*, where the intended reference of x is disambiguated at Logical Form by means of a referential index. In the salient construal of (19), for instance, *too* means 'in addition to *Mary*' and is therefore coindexed with *Mary* in the preceding clause. Also, *too* associates with focus, and this too is represented at LF, by means of the customary subscripted 'F'.[13] So the LF of (19), under the reading we want to consider, is (20).

(20) John believes that $Mary_1$ is here, and he believes that $Susan_F$ is here too_1.

The general rule for the interpretation of *too* is (21).

(21) φ[$α_F$]*too*$_i$ presupposes $x_i \neq \alpha$ & φ[x_i].

Transposed into the context change framework and applied to the example at hand, this amounts to (22).

(22) For any c,
c + $Susan_F$ *is here too*$_1$ is defined iff Mary is here in every world in c.
Where defined, c + $Susan_F$ *is here too*$_1$ = {w ∈ c: Susan is here in w}.

The rule for the connective *and*, of course, is (23) (cf. Karttunen's (4) above).

(23) $[c + \phi \text{ and } \psi$ is defined iff $c + \phi$ and $(c + \phi) + \psi$ are defined.]
Where defined, $c + \phi \text{ and } \psi = (c + \phi) + \psi$.

What we want to show is that (20) as a whole presupposes nothing. This means that any context (even the completely information-less W) is in the domain of the CCP of (20). So what we have to show is that $c + (20)$ is always defined, regardless of any special properties of c. Here is the proof:

Let c be an arbitrary context $\subseteq W$. By rule (23), $c + (20)$ is defined just in case both $c + \textit{John believes Mary}_1 \textit{ is here}$ and $(c + \textit{John believes Mary}_1 \textit{ is here}) + \textit{he believes Susan}_F \textit{ is here too}_1$ are. We first show that $c + \textit{John believes Mary}_1 \textit{ is here}$ is defined. This follows trivially by rule (14) from the fact that $\textit{Mary}_1 \textit{ is here}$ has no presuppositions, i.e. a CCP that is always defined. We also know from rule (18) what $c + \textit{John believes Mary}_1 \textit{ is here}$ (henceforth abbreviated as c') is, namely:

(24) $c' := c + \textit{John believes Mary}_1 \textit{ is here} =$
$\{w \in c: \text{Mary is here in all } w' \in Dox_J(w)\}$

We have left to show that $c' + \textit{he believes Susan}_F \textit{ is here too}_1$ is defined. By rule (18) this is so iff $Dox_J(w) + \textit{Susan}_F \textit{ is here too}_1$ is defined for every $w \in c'$. Let w be an arbitrary $w \in c'$. It follows by (24) that Mary is here in all $w' \in Dox_J(w)$. According to (22), this in turn guarantees the definedness of $Dox_J(w) + \textit{Susan}_F \textit{ is here too}_1$. End of proof.

This calculation should have made clear just how the utterance of the first conjunct of (20) is responsible for the fact that the second conjunct's presuppositional requirement is satisfied by the intermediate context against which it is evaluated. (And mind you, it is *satisfied*, not *cancelled*, even though it superficially may appear so!)

For a contrasting case, where presupposition filtering does not occur and in fact the discourse is deviant, consider (25).[14]

(25) John doubts that Mary$_1$ is here and/but believes that Susan$_F$ is here too$_1$.

That (25) doesn't make sense is predicted if we assume that *doubt* means (or at least implies) something like *not believe*. After the first conjunct in (25), we then have a context for all whose elements w $Dox_J(w)$ fails to entail Mary's being here. So not only is $Dox_J(w) + \textit{Susan is here too}_1$ not guaranteed for every such w to be defined, it is actually guaranteed to be undefined for all of them.[15]

I don't mean to suggest that the present analysis is unique in providing an account of the unacceptability of (25). The same prediction is made by anyone who assumes every presupposition to be also an entailment of the minimal sentence that carries it.[16] Given this assumption, the two conjuncts of (25) simply have incompatible contents, and this suffices to explain the deviance.

Notice, however, that this simpler explanation doesn't generalize to slightly more complex examples like (26).

(26) John doubts that Mary$_1$ is here. He believes that if Susan$_F$ were here too$_1$ there would be dancing.

This use of *too* is likewise deviant, but the *content* of complement of the second sentence, viz. that there would be dancing if both Mary and Susan were here, is not at all incompatible with Mary's being absent. The present analysis covers this case along with (25) (provided that the conditional inherits the presupposition of its antecedent, as standardly assumed; see below).

The reader may have been wondering how our rule (18) relates to Karttunen's rule (3), which I cited in the introduction. There is an obvious respect in which (3) says less than (18) (and in which all of Karttunen's rules say less than ours): (3) does not tell us what the *outcome* of incrementing a context by a *believes* ϕ is; it merely states the *prerequisites* of the incrementation operation. So it could at best be equivalent to the first half of (18), the definedness conditions (the part in brackets). But is it equivalent even to that? There is a superficial discrepancy: (18) requires the CCP of ϕ to be defined for each of a set of contexts, namely all the $\text{Dox}_a(w)$ for each $w \in c$, whereas (3), in effect, requires it to be defined for the single context $B_a(c)$. What is the relation between Karttunen's B_a and our Dox_a? Karttunen defines $B_a(c)$ as the set of beliefs attributed to a in c. So, construed as a set of propositions, $B_a(c) = \{p \subseteq W: c \text{ entails that } a \text{ believes } p\}$, or more explicitly: $B_a(c) = \{p \subseteq W: \forall w \in c: a \text{ believes } p \text{ in } w\}$. Rewriting this in terms of the doxastic accessibility function: $B_a(c) = \{p \subseteq W: \forall w \in c: \text{Dox}_a(w) \subseteq p\}$. Now if we form the set of worlds in which all the propositions in this set are true, what we get is $\bigcup_{w \in c} \text{Dox}_a(w)$. In other words (abstracting away from extrinsic differences): whereas I required the CCP of ϕ to be defined for each $\text{Dox}_a(w)$ for $w \in c$, Karttunen required it to be defined for the union of them all. But to the extent that definedness of the CCP of ϕ for a context is a matter of that context entailing a certain proposition (as in the cases of interest, where ϕ has a purely presuppositional CCP),[17] then the two requirements obviously come to the same thing: a proposition is entailed by each element of a set iff it is entailed by its union. So my proposal is not in conflict with Karttunen's, but can be seen as an elaboration of it.[18]

Thus far, we are following in Karttunen's footsteps and, indeed, if we restrict our attention to verbs of belief, our predictions fully mirror his. But differences show up when we extend the same treatment to other attitudes, say to the desire verb *want*. According to the standard Hintikka-style analysis we took as a starting point, the rule for *want* should look just like that for *believe*, except with a different accessibility relation substituted. Here it is *buletic* accessibility

that is relevant, so the pertinent accessibility function is Bul_j: $w \to \{w' \in W: w'$ conforms to what John wants in $w\}$.

(27) $[c + a \ wants \phi$ is defined iff $Bul_a(w) + \phi$ is defined for each $w \in c$.]
Where defined, $c + a \ wants \phi = \{w \in c: Bul_a(w) + \phi = same\}$.

The predictions implied by (27) diverge from Karttunen's in two ways, one good and one bad. First, the good news. (27) predicts correctly what happens in sequences of two desire reports, like (28) or (29) below.

(28) Patrick and Ann both dream of winning cellos. Ann would like one for her own use. Patrick wants to sell his cello for a profit.

(29) John wants $Fred_1$ to come, and he wants Jim_F to come too$_1$.

In these examples, the presupposition originating with the last complement clause also gets 'filtered out': they are felicitous and require no initial presuppositions. (28) makes sense without committing the speaker to the assumption that Patrick believes he has or will ever have a cello, and (29) also doesn't presuppose that John believes Fred will come.[19] Karttunen's rule (3) fails to account for this, but (27) predicts it straightforwardly. In fact, (27) derives the following generalization: if ϕ presupposes p, then $a \ wants \phi$ presupposes that a wants p. This is a welcome result for *want-want* sequences, but—and here comes the down side—it is not suited to capture the analogous filtering effect in *believe-want* sequences like our initial example (2) or (30) below.

(30) John believes that $Mary_1$ is coming, and he wants $Susan_F$ to come too$_1$.

These were better taken care of with Karttunen's generalization that $a \ wants \phi$ presupposes that a *believes* p.

4 DESIRE REPORTS AND COUNTERFACTUALITY

We just saw that the straightforward treatment of desire predicates in (27) fails to account for the ease with which a preceding belief-report can help satisfy the presuppositional requirement of the *want*-complement, as in (2) and (30). The problem is that the sets $Dox_a(w)$ and $Bul_a(w)$ (for a given w) may in principle stand in any relation whatsoever, i.e. they may be mutually disjoint, they may overlap, one may be a subset of the other, or vice versa. After all, which worlds I deem desirable has nothing to do with which I consider candidates for actuality. So there is no way we can be sure, e.g., in the evaluation of (30), that just because all of John's belief-worlds have Mary coming in them, this should also be so in all his desire-worlds.

Now it is no news to anybody who has thought about the semantics of *want-*

sentences that the primitive treatment in (27) has other defects as well (see below for examples). So it may be a good idea to cast around in the literature for a more sophisticted semantic analysis and then see if that perhaps helps with the presupposition projection facts. This is what I have tried to do. In the ideal case, there would have been some independently motivated analysis out there that only needed to be routinely transposed into the context change framework and then would have automatically gotten the projection facts right. Unfortunately, that wasn't quite what I found. But by combining insights from various sources, I have come up with something that does, I hope, throw *some* light on what the projection behavior of desire verbs has to do with their truthconditional semantics, even though not all the choices I had to make were determined by independent evidence.

4.1 A conditional semantics for desire verbs

The analysis of desire verbs I want to pursue here is sketched in Stalnaker (1984: 89): 'wanting something is preferring it to certain relevant alternatives, the relevant alternatives being those possibilities that the agent believes will be realized if he does not get what he wants.' An important feature of this analysis is that it sees a hidden conditional in every desire report. A little more explicitly, the leading intuition is that *John wants you to leave* means that John thinks that if you leave he will be in a more desirable world than if you don't leave.

The main task in implementing this idea is to spell out the conditionals in the above paraphrase. For this I employ a version of the semantics that Lewis (1973) proposed for counterfactual conditionals and Stalnaker (1968) for conditionals in general.[20] The key concept here is that of comparative similarity among worlds, and the basic idea is that a conditional *if ϕ, ψ* is true in a world w iff ψ is true in all ϕ-worlds maximally similar to w. (By a 'ϕ-world maximally similar to w', we mean a world in which ϕ is true and which resembles w no less than any other world where ϕ is true.) The meaning of *want*, as indicated by the paraphrase above, can now be described as follows:

(31) 'α *wants ϕ*' is true in w iff
for every $w' \in \text{Dox}_\alpha(w)$:
every ϕ-world maximally similar to w' is more desirable to α in w than any non-ϕ-world maximally similar to w'.

(31) instructs us, for every belief-world, to compare the set of its closest ϕ-alternatives to the set of its closest non-ϕ-alternatives. In effect, however, one of these two sets of 'alternatives' will always be the singleton of just w' itself: if ϕ is true in w', it is the former set, otherwise the latter. So another way of stating these truthconditions is in the following disjunctive form: For every belief

world w', either ϕ is true in w' and w' is more desirable than its closest non-ϕ-alternatives, or else ϕ is false in w' and w' is less desirable than its closest ϕ-alternatives. When a sentence like *I want you to call me on Monday* is used, there typically are doxastic alternatives where you do call me on Monday as well as those where you don't. For it to be true, then, each of the former must be more desirable than minimally different ones where you don't call, and each of the latter less desirable than minimally different ones where you do.

Independently of my present concern with presupposition projection, what motivation is there for this semantic analysis? In what respects is it more successful than the primitive treatment in (27)?

First, the new rule, unlike the old, no longer predicts that when ϕ entails ψ, *a wants ϕ* therefore entails *a wants ψ*. This is welcome in light of certain intuitively fallacious instances of this inference pattern. Here is an example I owe to Asher (1987): imagine that Nicholas is not willing to pay the \$3,000 that he believes it would cost him if he flew to Paris on the Concorde, but he would love to fly on the Concorde if he could get the trip for free. Under these circumstances (32a) is true, yet (32b) is false, despite the fact that taking a free trip on the Concorde, of course, implies taking a trip on the Concorde.

(32) (a) Nicholas wants a free trip on the Concorde.
     (b) Nicholas wants a trip on the Concorde.

The prediction of our rule (31) conforms to Asher's intuition: (32b) is false because many of Nicholas's doxastic alternatives in which he flies on the Concorde (in fact, all to which he assigns a high degree of subjective probability) are such that he flies for \$3,000 there and is therefore worse off than in minimally differing worlds where he doesn't fly at all. Yet this does not prevent (32a) from being true: those (relatively unlikely) belief-worlds where he does get a free trip are better than similar worlds where he doesn't, and the other (more likely) belief-worlds, where he doesn't fly, or flies and pays, are each worse than otherwise similar free-ride-worlds.

Stalnaker (1984: 89) discusses a different type of counterexample to the same inference pattern:[21] 'Suppose I am sick. I want to get well. But getting well entails having been sick, and I do not want to have been sick. Suppose there was a murder. I want to know who committed the murder. But my knowing who committed the murder entails that the murder was committed, and I never wanted the murder to have been committed.' These fallacious inferences, too, would have been validated by the old rule (27): if I get well in all the worlds that conform to my desires, then I have been sick in all the worlds that conform to my desires.

What does the new rule (31) predict for these examples? We would like to show that it allows for the premise, *I want to get well*, to be true and the conclusion, *I want to have been sick*, to be false at the same time. Here is how this

can be. There are basically three kinds of worlds: w_1, where I am healthy all the time; w_2, where I am sick first and then get well; and w_3, where I am sick and stay sick. In terms of their desirability to me in the actual world w_0, they are ordered as follows: w_1 is better than w_2, which is better than w_3. My beliefs in w_0 are such that I believe that I have been sick, i.e., w_2, $w_3 \in \text{Dox}_I(w_0)$ but $w_1 \notin \text{Dox}_I(w_0)$. Now in w_2 I get well and the closest world where I don't is w_3, which is less desirable. And in w_3, I don't get well and the closest world where I do is w_2, which is more desirable. Hence *I want to get well* is true in w_0. On the other hand, in both w_2 and w_3 I have been sick, but these are not better (rather: worse) than the closest world, w_1, where I haven't been. So *I want to have been sick* is false in w_0.

For a related point, consider a statement like (33).

(33) I want to teach Tuesdays and Thursdays next semester.

Suppose this sentence is intuitively true as spoken by me today. Is it therefore the case, as the old rule (27) would have it, that I teach Tuesdays and Thursdays next semester in all the worlds that are compatible with everything I desire? No. In worlds that are compatible with everything I desire I actually don't teach at all. But if this is so, (27) predicts (33) to be false, and likewise for the majority of *want*-sentences that we accept as true in everyday conversation. Rule (31) has no such problem: as it happens, I believe that I will teach (a regular course load) next semester. This means there are no doxastically accessible worlds in which I don't teach at all. In all doxastically accessible worlds, I either teach Tuesdays and Thursdays, or else I teach the same load on different weekdays. Among these, the former are more desirable than the latter, and this makes (33) true by (31).

4.2 CCP and presupposition projection

Supposing that the analysis of desire reports in (31) is on the right track, what form will it take in context change semantics? Let me approach this question via a detour and look first at the CCP of indicative conditional statements.

4.2.1 Excursion: context change with indicative conditionals

First, a few technical devices and abbreviations. The relation of comparative similarity among worlds can be encoded by a family of selection functions; for each world w, there is a selection function Sim_w from propositions to propositions which maps each p to the set of p-worlds maximally similar to w:

(34) $\text{Sim}_w(p)$: —
    $\{w' \in W : w' \in p$ and w' resembles w no less than any other world in $p\}$

In a truthconditional semantics, where each sentence ϕ expresses a proposition $[[\phi]]$, the semantic rule for conditionals can now be stated as follows:

(35) $w \in [[if\phi, \psi]]$ iff $\text{Sim}_w([[\phi]]) \subseteq [[\psi]]$.

When we try to transpose this into the context change framework, the main hurdle is to find a proposition that will serve as the argument for the selection function. We can't just make reference to 'the proposition expressed by ϕ'; rather, we get a proposition only by applying the CCP of ϕ to some argument. What should that argument be?

An often voiced intuition is that it is the main context, i.e., the input context to the CCP of the whole conditional. Compare, e.g., Stalnaker (1975: 276): 'when a speaker says *if A*, then everything he is presupposing to hold in the actual situation is presupposed to hold in the hypothetical situation in which A is true.'[22] This suggests the following CCP definition:

(36) $c + if\phi, \psi = \{w \in c: \text{Sim}_w(c + \phi) + \psi = \text{same}\}$

Suppose, for example, Mary calls us and tells us she is calling from a phone booth. So it is part of our common ground c that Mary is in the phone booth. If I now say *If John is in the phone booth*. . ., the hypothetical situations I am asking you to consider are all situations where John and Mary are in the phone booth *together*, as opposed to those where he is there *instead of her*. For instance, if I continue . . . *then the door doesn't close*, this will in effect give you information about how the size of the booth relates to the *combined* volume of Mary and John, and it won't tell you anything about how John's size *alone* relates to the booth's. This interpretation is forced even if worlds with two people in a phone booth at once are relatively far-fetched in comparison to the worlds in c; in other words, if the selection function is such that the closest worlds *with John in the booth* that it would pick out for any $w \in c$ are worlds where he is there alone. So it is not (or at least not necessarily) a property of the similarity relation that leads us to consider worlds with John added to Mary rather than worlds where he replaces her. Rather, it seems to be due to the fact that, in evaluating this conditional, the selection function must apply to a proposition that retains all the information in c along with that contributed by the antecedent. Rule (36) guarantees this.

(36) also makes welcome predictions about presupposition projection in conditionals. In particular, it directly derives the familiar generalization[23] that conditionals inherit the presuppositions of their antecedents. In our terms: unless $c + \phi$ is defined, $c + if\phi, \psi$ won't be either. And it also derives the fact that presuppositions of the consequent which are entailed by the antecedent get 'filtered out'. (This is because the CCP of ψ is applied to $\text{Sim}_w(c + \phi)$ and this, by the general definition of selection functions, is a subset of $c + \phi$.)[24]

4.2.2 Back to *want*

Now we return to the analysis of *want*. The truthconditional-semantics version from (31) above is reformulated below:

(37) $w \in [[\alpha \text{ wants } \phi]]$ iff for every $w' \in \text{Dox}_\alpha(w)$,
$\text{Sim}_{w'}([[\phi]]) <_{a,w} \text{Sim}_{w'}(W\backslash[[\phi]])$.

Apart from the notation employed in the previous section, this uses an abbreviation for the ranking of possible worlds in terms of desirability. $<_{a,w}$ is meant to be primarily a relation between worlds (defined in (38a)), but it is employed in (37) in an extended sense (defined in (38b)), as a relation between sets of worlds.

(38) (a) For any $w, w', w'' \in W$,
$w' <_{a,w} w''$ iff w' is more desirable to α in w than w''.
  (b) For any $w \in W$, $X \subseteq W$, $Y \subseteq W$,
$X <_{a,w} Y$ iff $w' <_{a,w} w''$ for all $w' \in X$, $w'' \in Y$.

Once again, the main hurdle in transposing (37) into the context change framework is to find propositions for $\text{Sim}_{w'}$ to apply to that will take the place of $[[\phi]]$ and $W\backslash[[\phi]]$ in (37) respectively. This time, the w' on which the selection functions $\text{Sim}_{w'}$ are based are all drawn from $\text{Dox}_\alpha(w)$, so this set is a natural choice to take the place of c in (36) if we want to make the rules analogous. So I propose (39).

(39) $c + \alpha \text{ wants } \phi =$
{$w \in c$: for every $w' \in \text{Dox}_\alpha(w)$:
$\text{Sim}_{w'}(\text{Dox}_\alpha(w) + \phi) <_{a,w} \text{Sim}_{w'}(\text{Dox}_\alpha(w) + \text{not } \phi)$}

(Where $\text{Dox}_\alpha(w) + \text{not } \phi$ is, of course, defined as $\text{Dox}_\alpha(w)\backslash\text{Dox}_\alpha(w) + \phi$.) (39) implies that all the desirability comparisons that enter into determining the truth of a *want*-sentence are entirely among the subject's *belief worlds*. If a belief world w' has ϕ true in it, it must be more desirable than otherwise similar *belief* worlds where ϕ is false, and if a belief world has ϕ false in it, then it must be less desirable than otherwise similar *belief* worlds where ϕ is true. The desirability of non-belief-worlds never has any bearing on the truth of a *want*-report. We will see shortly that this may be an excessively strong claim. But let us suppose it is correct for the moment and examine the implications of rule (39) for presupposition projection.

Not surprisingly at this point, (39) straightforwardly predicts Karttunen's generalization: the presuppositions of α *wants* ϕ are satisfied just in case the subject α is presupposed to believe the presuppositions of the complement ϕ. Let's calculate quickly through an illustrating example. To be shown: (30) as a whole presupposes nothing, in other words, c + (30) is defined regardless of the

choice of c. The first part of the proof parallels the one regarding (20) in the section on belief reports: we establish that c + *John believes Mary$_1$ is here* (=: c') is well defined for all c and equals {w ∈ c: Mary is here in all w' ∈ $Dox_J(w)$} (see (24) above). It remains to demonstrate the definedness of c' + *he wants Susan$_F$ to be here too$_1$*. By (39), we must show that $Dox_J(w)$ + *Susan$_F$ to be here too$_1$* and $Dox_J(w)$ + *not [Susan$_F$ to be here too$_1$]* are defined for all w ∈ c', which (by the *not*-rule (8) and the *too*-rule (21)) means that, for each w ∈ c', Mary is here in all w' ∈ $Dox_J(w)$. But this we have just shown.

4.2.3 Amendments

As it stands, the context change version (39) of our analysis of *want* loses one of the welcome predictions of the truthconditional version in (31). Recall again Stalnaker's concern with blocking the inference from *I want to get well* to *I want to have been sick* (and from *I want to know who committed the murder* to *I want the murder to have been committed*). I showed above how (31) made the premise true and the conclusion false because I believe that I have been sick. In the same scenario, (39) unfortunately predicts the conclusion to be trivially true instead of false. If I believe in w that I have been sick, then $Dox_J(w)$ + *not [PRO to have been sick]* is empty, and so is $Sim_{w'}$ applied to it. Since it is trivially true that all the worlds in the empty set are worse than any others, this suffices to make the conclusion true. Stating the problem more generally, (39) predicts that, whenever α believes ϕ or believes *not*ϕ, it trivially follows that both α wants ϕ and wants *not*ϕ.

A natural move to prevent these trivial truths is to make all selection functions undefined for the impossible proposition.[25] In other words, amend (34) above to (40) below.

(40) p is in the domain of Sim_w only if p ≠ ∅; where defined, Sim_w(p): = {w' ∈ W: w' ∈ p and w' resembles w no less than any other world in p}

With (40), (39) implies, in effect, that *want*-sentences have an additional presupposition (above and beyond those projected from the complement according to Karttunen's generalization), namely that the subject does not believe the complement nor its negation. More formally, c + α *wants* ϕ will be undefined whenever $Dox_\alpha(w)$ + ϕ = ∅ or $Dox_\alpha(w)$ + ϕ = $Dox_\alpha(w)$.

Regarding Stalnaker's fallacious inferences, (39) still doesn't predict the same as (31). When the premise is true, (31) allowed the conclusion to be downright false; (39) only allows it to be a presupposition failure. This disagrees with Stalnaker's stated judgment, but for his examples, at least, I think it is quite defensible. *I want to have been sick* (as well as *I want not to have been sick*, or its more colloquial Neg-Raising variant *I don't want to have been sick*) is a strange

sentence indeed to use for someone who takes for granted that she has been sick.[26] One would much rather say *I am glad that I have been sick* or some such thing, with a *factive* desire predicate. Stalnaker too spontaneously avoids *want* as he further comments on the murder-example: 'Given that there was a murder, I would rather know who committed it than not know. The question of whether or not I *look with favor on the fact* that there was a murder—whether I *am glad* that it happened or *wish* that it had not—does not arise in that context' (1984: 89; emphases added). (This raises the question of how these other desire-predicates differ from *want*, which I will take up below.)[27]

Still, even if you agree that Stalnaker's examples are appropriately classified as infelicitous rather than false, it doesn't seem right that one can never speak of wanting things one is convinced will happen or convinced won't happen. (41), for instance, certainly does not suggest in any way that John has the slightest doubt about where he will be tonight, nor do we have difficulty making sense of utterances like (42).

(41) (John hired a babysitter because) he wants to go to the movies tonight.
(42) I want this weekend to last forever. (But I know, of course, that it will be over in a few hours.)

These observations are a serious threat to the present analysis, and I am persuaded at least by (41) that a genuine modification is called for. To see what I have in mind, consider briefly the semantics of a related verb, *intend*, which displays the behavior of *want* in (41) even more strikingly. What one intends is typically, not just occasionally, something that one is convinced will happen. So our rule for *want*, generalized as it stands to *intend*, would systematically predict presupposition failure for perfectly appropriate *intend*-sentences. But the correct rule, I think, is only a little bit different. What seems to be going on when we assess someone's intention is that we don't take into account *all* his beliefs, but just those that he has about matters unaffected by his own future actions. More precisely, what should take the place of Dox_a in the rule for *intend* is the following accessibility function F_a:[28]

(43) For any $w \in W$: $F_a(w) = \{w' \in W$: w' is compatible with everything that a in w believes to be the case no matter how he chooses to act}

$F_a(w)$ is always a superset of $Dox_a(w)$. If we substitute F_a for Dox_a as we adapt our *want*-rule (39) for *intend*, we no longer predict *John intends to go to the movies* to be a presupposition failure just because John is convinced he will in fact go. (We would only predict it to be inappropriate if he were convinced he'll go no matter how he chooses to act. *This* prediction seems right.) The substitution also implies different predictions for presupposition projection, in fact, a subtle departure from Karttunen's generalization. No longer do we predict that (44) presupposes just that Patrick believes he has a cello tomorrow, but that it

presupposes, more specifically, that he believes he has a cello *independently of what he does*.

(44) Patrick intends to sell his cello (right now).

This is certainly plausible for the example at hand, and I hope that a closer look at the data will bear it out in general.[29] Returning now to *want* and example (41), I suggest that *want* has a reading more or less equivalent to *intend* and this is what we witness here. Probably this is not really an ambiguity but indicates a broader sort of vagueness. But this question is just one of numerous loose ends that I am leaving here.

What about (42)? I am even less sure how to respond to this example. One strategy might be similar to the one I just took with (41): maybe for some reason not all the subject's beliefs are taken into account here either, but only a subset too weak to imply that Monday is right around the corner. Alternatively, (42) might be seen as reporting the attitudes of a mildly split personality.[30] The reasonable part of me knows and is resigned to the fact that time passes, but the primitive creature of passion has lost sight of it. Another loose end.

4.3.4 Further predictions about presupposition filtering

Rule (39) captures Karttunen's generalization so 'well' that it also shares its more dubious and downright inadequate predictions. As for the latter, (39) loses the one thing that was nice about (27), namely the prediction that presuppositions are filtered in *want-want*-sequences, such as (28) and (29) (repeated here).

(29) John wants Fred$_1$ to come, and he wants Jim$_F$ to come too$_1$.

Supposing that (as the felicity of the first conjunct requires, by our present analysis) the possibility of Fred not coming is compatible with John's beliefs, the CCP of the second conjunct is not defined for its context. I have no solution to this important problem. All I can do is point out that under the present perspective it falls together with an analogous type of counterexample to Karttunen's generalization about conditionals.[31] (45) is likewise fully acceptable, though predicted a presupposition failure by rule (36).

(45) If Mary$_1$ comes, we'll have a quorum. If Susan$_F$ comes too$_1$, we'll have a majority.

The only way to treat this case—and thus the analogous one in (29)—that I know of is by invoking accommodation of an inexplicit restriction.[32] Once this mechanism is invoked, of course, the question arises to what extent it could also have been employed to yield some of the predictions that I took pains to

make follow directly from the CCP definitions.[33] A serious exploration of this alternative must await another occasion.

Another prediction shared with Karttunen's original proposal is worth noting: the order of *believe* and *want* is not interchangeable if we want the presupposition of the second complement to get filtered. While (2) presupposes nothing, the same is not predicted for (46).

(46) Patrick wants me to buy him a cello, although he believes that his cello is going to take up a lot of space.

Does intuitive judgment bear this prediction out? Not strikingly so, but (46) does seem a bit less good than (2) or (28). The same pattern shows up in the following judgments from Asher (1987) (the judgments pertain to the possibility of an anaphoric reading for *it* while the intended antecedent (*a car, a Porsche*) has narrow scope with respect to the first attitude verb):

(47) (a) Fred believes that his wife will buy him a car. He hopes that *it* will be a Porsche.
      (b) ?Fred hopes that he will get a Porsche. He believes that his wife will buy *it* for him.
      (c) *John wants to have a Porsche. He believes his mother will buy *it* for him.

Supposing that the pronouns on the intended anaphoric readings would be E-Type pronouns, equivalent to the definite descriptions *the car his wife will buy him*, *the Porsche he will get (have)*, the acceptability of these pronouns turns on the satisfaction of the corresponding definite description's presuppositions. In this light, (47a–c) support our prediction. On the other hand, Asher accepts (48), and Cresswell (1988) offers (49).

(48) John wants a woman to marry him. He believes he can make her happy.
(49) Susan wants a pet. She believes she will look after it.

Both authors comment that the meaning of the *believe*-complement here is an implicit conditional: John believes that, *if* a woman marries him, he can make her happy, and Susan believes that, *if* she gets a pet, she will look after it. So once again, we must invoke accommodation (modal subordination). I would like to note, however, that certain examples of an analogous form do seem to fit transparently with our (i.e., Karttunen's) analysis. Suppose I had to miss the last set of the Wimbledon women's final because of a hairdresser's appointment, so I don't know who won, though I do know the game is over and decided by now. In such a situation, I might say (50)

(50) I want Gabriela to have won.

However, it would be strange to continue as in (51), even though—in analogy with the examples above—this should be just another way of saying that I am convinced that Steffi cried *if* Gabriela won.[34]

(51) *... and I am sure that it made Steffi cry hard.

So there do seem to be some limitations (however obscure) on the availability of accommodated restrictions and, when these apply, the workings of the CCPs themselves are seen more directly and tend to confirm the present approach.

4.3 *Counterfactual and factive desire predicates*

Want-sentences are felicitous in contexts where it has already been established that the subject believes the presuppositions of its complement. Hence the naturalness of sequences like (2) and (30). Our current analysis suggests that this is so because of a fact about the meaning of *want*: *want*-sentences are interpreted with respect to a doxastic modal base: to want ϕ means to find the ϕ-worlds among the worlds compatible with one's beliefs more desirable than comparable non-ϕ-worlds compatible with one's beliefs. Thus the truth of a *want*-sentence never turns on the desirability of any worlds which contradict the subject's beliefs. Therefore, if only such non-belief-worlds violate the presuppositions of ϕ, we can be guaranteed that we won't need to consider them in evaluating a *wants* ϕ.

Apart from the subtler doubts we already raised above, there is a rather obvious reason why this explanation cannot be right. If we consider a wider range of desire predicates, we find that the majority of them do not require or even permit such doxastic modal bases. This is particularly clear in an example like (52), where the use of *wish* with the irrealis mood[35] suggests strongly that John is pessimistic, perhaps that he is even certain he won't teach Tuesdays.

(52) John wishes he would teach on Tuesdays.

(52) cannot be analyzed as meaning that John teaches on Tuesdays in his most desirable belief-worlds; to the contrary, it suggests that he doesn't teach on Tuesdays in *any* of his belief worlds. Nevertheless, presuppositions triggered in the complements of *wish*-sentences appear to be satisfied by previous *believe*-sentences. (53) is just as felicitous as (30).

(53) John believes that Mary$_1$ is the only one here, and he wishes Susan$_F$ were here too$_1$.

If we did insist on interpreting *wish* here with a doxastic modal base, i.e., followed essentially our rule for *want* from above, we could not provide an adequate interpretation for (53): after processing the *believe*-sentence, we would have a context c′ such that for all its elements w, $Dox_J(w)$ entails that Mary is

here and nobody else is. According to rule (39) with amendment (40), c' + *he wishes Susan$_F$ were here too$_I$* would then always be undefined. This is because $Dox_J(w)$ + *Susan$_F$ were here too$_I$* is empty for all $w \in c'$, and thus not in the domain of Sim_w.[36] In short, (39) would predict the second sentence of (53) to be inappropriate whenever the first has been accepted, and this is clearly wrong.

Counterfactual desire reports like those involving *wish* + irrealis complement are not the only ones that create problems. Our present analysis of *want* also fails to generalize to the *factive* members in the family of desire predicates. Consider (54).

(54) John is glad he will teach on Tuesdays.

The predicate *be glad*, like all factive desire predicates, introduces a presupposition to the effect that the subject believes in the truth of its complement.[37] Again, this trivializes the truthconditions that would be predicted for (54) if we simply used the same rule (39) as for *want*. This time, it is the set to the right of $<_{a,w}$ that is automatically undefined. Again, that's evidently wrong.

Here, too, the intuitive source of the problem is that (39) excludes all doxastically inaccessible worlds from consideration. Evidently, one cannot assess the truth of a *is glad that* ϕ without reference to facts about the desirability of certain non-ϕ-worlds, and thereby of certain non-belief-worlds. But then it becomes mysterious again why the presuppositions of ϕ (if any) need only be satisfied in the belief-worlds, as the felicity of (55) indicates.

(55) John$_1$ thought he$_1$ was late and was glad that Bill$_F$ was late too$_1$.

Consideration of such an extended range of desire predicates suggests that our current analysis of *want* constitutes at best a special case. How come, then, the facts about presupposition projection are exactly the same for those other desire predicates as for *want*? Not only does a preceding belief-sentence satisfy the presuppositions of the subsequent desire-complement: (56a,b) presuppose nothing as a whole; but we also spontaneously accommodate the assumption that the subject believes the presupposition when we hear (57a,b) out of context.

(56) John believes Mary$_1$ is coming, and
    (a) he is glad Susan$_F$ is coming too$_1$
    (b) he wishes Susan$_F$ were coming too$_1$.
(57) (a) Patrick is glad he sold his cello.
    (b) Patrick wishes he had sold his cello.

If the modal base for these verbs is not doxastic, then why should the presuppositions of their complements be satisfied just because the subject is known to *believe* them? We have no guarantee of this.

4.3.1 Excursion: counterfactual conditionals

It is instructive to see that a similar dilemma arises with presuppositions in the antecedents of counterfactual conditionals. We couldn't just use the same rule (36) for them that we gave above for indicative conditionals, because their antecedents are typically inconsistent with the common ground and they would thus come out undefined. But what should we use instead of c in (36) to apply the CCP of ϕ to? Most discussions of counterfactuals in the literature suggest that it should just be W, i.e., a context devoid of all information.[38] That way, the CCP for subjunctive conditionals would be as in (58).

(58) $c + $ *if* ϕ *would* $\psi = \{w \in c: \mathrm{Sim}_w(W + \phi) + \psi = \mathrm{same}\}$

But then we have an unwelcome prediction: counterfactuals whose antecedents have presuppositions should never be interpretable,[39] because the modal base, being W, can't have the required entailments. It is surprising, then, that counterfactuals with presuppositional antecedents are so common and that they seem perfect under the same condition that their indicative variants are, viz. when the previous (primary) context entails the presupposition. For instance, when it is already in the common ground that Mary attended, that seems to license (60) as much as (59).

(59) If John attended too, . . .
(60) If John had attended too, . . .

These examples suggest that the antecedent of a counterfactual is not really added to an '*empty*' context, but to one which is in some sense a *revision* of the common ground c. It results from c by suspending some of the assumptions in c; i.e., it is a superset of c. But since the specific purpose of the revision is to create an input context for the CCP of the antecedent, there are limits to what can be suspended: presuppositions required by the antecedent must stay. Let's assume, for concreteness, that the result of the revision is always the biggest ($-$ least informative) context within those limits.[40] This leads to the following definition.

(61) For any context c, LF ϕ:
    $\mathrm{rev}_\phi(c)$, the *revision of c for* ϕ, is $\bigcup\{X \subseteq W: c \subseteq X$ and $X + \phi$ is defined$\}$.

The CCP definition for counterfactual conditionals can then be given as in (62).[41]

(62) $c + $ *if* ϕ *would* $\psi = \{w \in c: \mathrm{Sim}_w(\mathrm{rev}_\phi(c) + \phi) + \psi = \mathrm{same}\}$

(62) solves our dilemma regarding the inheritance of presuppositions from the antecedents of counterfactuals. Not in a particularly exciting way, of course; I

have simply stipulated the appropriate constraint on the revision process in (61). At any rate, it follows directly from (61) that $rev_\phi(c)$ will entail the presuppositions required by ϕ if and only if c does.[42] For our purposes, where we restrict attention to CCPs whose definedness conditions are purely presuppositional,[43] this means that $rev_\phi(c) + \phi$ is defined iff c $+ \phi$ is. Given (62), it then follows that a context is in the domain of the CCP of a counterfactual conditional only if it is in the domain of the CCP of its antecedent. In other words, counterfactual conditionals inherit the presuppositions of their antecedents.

Let me close this excursion with a remark on the effect of presuppositional requirements in the antecedent of a counterfactual's truth conditions. Recall the context where Mary is presupposed to be in the phone booth. We noted above that an indicative *if*-clause like *If John is in the phone booth . . .* in this context amounts to the supposition that *both* John and Mary are in the booth. This is otherwise for a minimally different subjunctive *if*-clause: If we say *If John WERE in the phone booth*, then it depends on the actual facts and the selection function whether the hypothetical situations under consideration have both people in the booth or have John there *instead of* Mary. . . . *then Mary would be outside* is a felicitous and possibly true continuation. (As opposed to the deviant indicative variant *If John IS in the phone booth, then Mary is outside*. This is acceptable only if we are ready to conclude that Mary's being in the phone booth isn't presupposed after all.) This difference, of course, is predicted by (62). But what is also predicted is that if we add to the subjunctive antecedent a *too*, as in *If John were in the phone booth too . . .*, then the meaning is in a certain respect more like that of the indicative again: no matter what the selection function and facts of the world, we only get to consider hypothetical worlds with both people in the booth together. So *If John were in the phone booth TOO, then Mary would be outside* is also deviant.[44]

4.3.2 Back to *wish* and *be glad*

I want to propose that *wish* and *be glad* have the same core semantics as *want*, but there is a difference that is analogous to that between indicative and subjunctive conditionals. To get the intuitive idea, recall our initial conditional paraphrase for *John wants you to leave*: John thinks that if you leave he will be in a more desirable world than if you don't leave. If we try to construct similar paraphrases for sentences with *wish* and *be glad*, here is how they come out: *John wishes you were gone* means 'John thinks that if you *were* gone he *would* be in a more desirable world than he is in *because* you are not gone'. *John is glad you are gone* means 'John thinks that *because* you are gone he is in a more desirable world than he *would* be in if you *were* not gone.' The common pattern is apparent, and the differences are in the choice of indicative vs. subjunctive mood and of *if* vs. *because*.

So what should the CCP definitions look like? The counterfactual conditional in the paraphrase for the *wish*-sentence, together with what we just said about counterfactual conditionals in the last section, suggests the following minimal variant of the rule for *want*:

(63) $c + \alpha$ wishes ϕ =
$\{w \in c$: for every $w' \in Dox_\alpha(w)$:
$Sim_{w'}(rev_\phi(Dox_\alpha(w)) + \phi) <_{a,w} Sim_{w'}(Dox_\alpha(w) + not\,\phi)\}$

Here I just replaced $Dox_\alpha(w)$ by $rev_\phi(Dox_\alpha(w))$ on the left side of $<_{a,w}$. Actually, the right side can be simplified: if we take for granted that ϕ is incompatible with α's beliefs,[45] then $not\,\phi$ is entailed by them, and therefore $Sim_{w'}(Dox_\alpha(w) + not\,\phi)$ for each $w' \in \cup_{w \in c} Dox_\alpha(w)$ is just $\{w'\}$. So we get away with (64).

(64) $c + \alpha$ *wishes* ϕ =
$\{w \in c$: for every $w' \in Dox_\alpha(w)$: $Sim_{w'}(rev_\phi(Dox_\alpha(w)) + \phi) <_{a,w} w'\}$

The paraphrase of the *glad*-sentence, on the other hand, has the counterfactual on the opposite side of 'more desirable than', so this is where we should substitute $rev_\phi(Dox_\alpha(w))$ for $Dox_\alpha(w)$ in (39). And because of factivity, i.e., the fact that α's beliefs can be assumed to entail ϕ, we can this time simplify the left side. The result is (65).

(65) $c + \alpha$ *is glad* ϕ =
$\{w \in c$: for every $w' \in Dox_\alpha(w)$: $w' <_{a,w} Sim_{w'}(rev_\phi(Dox_\alpha(w)) + not\,\phi)\}$

Presupposition projection from the complements of *wish* and *be glad* sentences, of course, works as desired now. (64) and (65) both imply (by reasoning parallel to that above) that the CCPs of α *wishes/is glad* ϕ will be defined for an initial context c iff $Dox_\alpha(w) + \phi$ is defined for all $w \in c$—Karttunen's generalization. Recall, e.g., our dilemma with (53). Our context c' after processing the *believe*-sentence was c' = $\{w$: Mary and nobody else is here in all $w' \in Dox_J(w)\}$. For each such $w \in c'$, what can we say about $rev_{Susan-_F-were-here-too_I}(Dox_J(w))$? By definition of the revision process, this set is not so big as to include any world where Mary isn't here, and so we can be assured that the CCP of *Susan_F were here too_I* is defined for it. The whole incrementation process thus goes through smoothly, and the presupposition filtering in (53) is accounted for.

5 WHY DO ATTITUDE VERBS APPEAR TO BE HOLES?

Why is it that a sentence like (66), uttered in isolation, seems to presuppose that it actually was raining, rather than merely that John believes so?

(66) John believes that it stopped raining.

Observations of this sort led Gazdar to conclude that attitude verbs are essentially holes (with apparent exceptions due to cancellation).[46] For us as for Karttunen, they are filters (as a matter of their intrinsic semantics), and such hole-behavior is unexpected.

Let us be clear about what exactly the problem is. There is no dispute about the interpretation of (66) in contexts where it is presupposed that John believes it had been raining. This presupposition—as our analysis predicts and as everyone agrees—suffices to make (66) interpretable and nothing else is accommodated. What, however, if it isn't yet presupposed that John believes it was raining? Then something must be accommodated. What will this be? Our analysis as it stands, it would seem, leads us to expect the minimal accommodation required to make the sentence interpretable. This would be accommodation of the assumption that *John believes* it to have been raining. But in point of fact, we spontaneously accommodate something else, namely that it had *in fact* been raining.

So there is a prima facie discrepancy between the observed facts and our (Karttunen's) predictions. The purpose of this section is to consider some independent factors on which we might blame these facts without abandoning our basic analysis. I will consider two hypotheses in particular:[47] first, that all cases where attitude verbs seem to be holes result from *de re* construals of (a constituent containing) the presupposition trigger. Second, a version of Karttunen's 'spill-over' story cited in the introduction.

5.1 De re *readings*

If a presupposition trigger in the complement clause is not really interpreted in the scope of the attitude verb, then it will be unsurprising on any theory that the relevant presupposition must be satisfied or accommodated in the main context. Take, for instance, the existence and uniqueness presupposition associated with the italicized definite description in (67).

(67) John thought that *the person who was going to kill him* had come to read the gas meter.

In the salient reading of (67), this definite is interpreted *de re*. Analyses of this phenomenon vary, but somehow or other they all imply that it is the speaker of (67), not John, who is 'responsible' for the definite description. For concreteness, assume a Quine-Kaplan-Lewis analysis of *de re* reports along the following lines.[48]

(68) There is an acquaintance relation D such that
    (i) John bore D to the person who was going to kill him, and
    (ii) John thought that whoever he bore D to had come to read the gas meter.

However this is implemented in detail, the definite description is outside the attitude complement in this paraphrase, and it is only to be expected that its presupposition must be entailed by the common ground. Nothing is asserted or presupposed about whether John believes there is somebody that will kill him. (The invited inference here is, of course, that he lacks this belief, but the sentence itself doesn't say one way or the other.)

Now it is quite uncontroversial that *some* cases of apparent presupposition inheritance from an attitude complement—such as in this salient reading of (67)—should be explained away in this manner. But would it be plausible to speculate that *all* presuppositions that percolate to the top from what seems to be a complement-internal trigger are really riding piggy-back on a *de re* construal of their trigger (or an expression containing it)? There are at least two prima facie obstacles to such a claim. First, the notion of a *de re* reading does not so obviously generalize to presupposition triggers other than definite descriptions. Second, one would be committed to the view that *de re* readings are *ceteris paribus* always preferred over *de dicto* readings, which contradicts superficial evidence. Let's take a closer look at each of these two points.

5.1.2 *De re* construals for other presupposition triggers?

When we move beyond definite descriptions, we have to clarify, on a case-by-case basis for each type of presupposition trigger, what a *de re* reading would even consist in. To the extent that the constructions in question can be analyzed as involving covert definite descriptions, this is relatively easy.[49] Take, e.g., the presuppositions of aspectual verbs like *stop*. Uncontroversial instances of *de re* readings like in (67) are harder to come by here, but they can be found if one looks.

(69) John thought I had stopped proof-reading.

By (69) I could conceivably mean that John thought of the activity of mine that was in fact a proof-reading, but that he may not have recognized as such, that it had stopped. For instance, John may have seen me from a distance and thought I was reading a magazine, then (after he had looked away) heard my step, at which point he concluded I must have stopped reading the magazine. In fact, I was proof-reading my article and continued doing this even as I was walking around. To be a true report about John's attitudes in this story, (69) would have to be represented along the following lines:

(70) There is an acquaintance relation D such that
    (i) John bore D to my proof-reading, and
    (ii) John thought that the activity he bore D to had stopped.

Here I have, in effect, treated the -*ing*-complement of *stop* as a definite description of a process. A verifying value for D in our scenario could be the

relation that person x bears to activity y if x observed such and such visible manifestations of y (i.e., the visible manifestations of my proof-reading that John perceived when he looked). So analyzed, (69) evidently presupposes that I had *in fact* been proof-reading, and it doesn't presuppose that *John thought* I had been.

But what about presupposition triggers like *again*, *even*, or the *too* in many of our examples throughout this article? What might it mean for an occurrence of one of those to receive, or be part of a constituent that receives, a *de re* construal? Here is an example which might help us clarify this question and where such a construal might be independently motivated. Imagine two kids talking to each other on the phone:

(71) John: I$_1$ am already in bed.
     Mary: My parents think I$_F$ am also$_1$ in bed.

The point about example (71) is that it is quite clearly felicitous even if Mary's parents cannot be assumed to have any beliefs about John. Mary is not committed to the presupposition predicted by Karttunen, i.e., that her parents believe John to be in bed.[50]

I bring up this example here because it seems promising to try to account for its apparently exceptional projection behavior by analyzing Mary's utterance along the lines of a *de re* paraphrase like (72).

(72) Of the property of also being in bed, my parents think that I have it.

The idea behind this paraphrase is that 'the property of also being in bed' (more accurately here: 'the property of [PRO$_F$ also$_1$ being in bed]') is just another way of describing the property of being in bed, and that it is a description which fits that property only contingently: it is true of it just in case John happens to be in bed. And since the latter is a fact known to Mary but unknown to her parents, she, but not they, can describe it in those words. This would have to be worked out further, and I am not convinced it is the right approach to this type of example,[51] but it deserves consideration.

5.1.3   A general preference for *de re* readings?

Suppose we can overcome the first obstacle and posit plausible *de re* construals for all kinds of presupposition triggers. Would this amount to an alternative explanation of Gazdar's observation that when attitude reports with presuppositional complements are presented out of context, we always accommodate the presupposition in the main context? Not all by itself. We would have to defend the further claim that *de re* readings are *ceteris paribus* preferred wherever there is a choice between a *de re* and a *de dicto* construal. Without this assumption, we would merely predict that presuppositions *sometimes* percolate

to the top, and we would expect this to correlate with independent clues in favor of a *de re* reading (such as the overall plausibility considerations that encourage us to flesh out the story in (67) in such a way that John is an unsuspecting murder victim).

Now common wisdom certainly has it the other way round: *de dicto* readings are the unmarked choice. For one thing, this is what you'd expect under standard analyses relying on quantifying-in, polyadic homophones of the attitude verb, or another such special mechanism to generate the *de re* reading; *de dicto* readings are somehow simpler and conceptually prior on all these approaches. Moreover, it seems to be confirmed by intuitive judgment. For instance, what is the unmarked reading of (73)?

(73) Does Ralph think that *the man he saw at the beach* is a spy?

Evidently the *de dicto* reading, because the spontaneous answer (given the facts of Quine's story) is: 'No, he thinks that the man he saw at the beach is a pillar of the community.'

Despite these initial deterrents, I think that a basic preference for *de re* readings may be defensible. Here is how. First, I propose a slight refinement of the standard *de re* analysis:[52] replace existential quantification over acquaintance relations by reference to a contextually salient particular acquaintance relation. For instance, (67) means (74) rather than (68).

(74) (i) John bore D to the person who was going to kill him, and
      (ii) John thought that whoever he bore D had to come to read the gas
           meter;
      where D is the acquaintance relation supplied by the utterance context.

Like all context-dependency, the selection of an appropriate acquaintance relation for the interpretation of a given utterance of a *de re* belief report depends on a miscellany of pragmatic factors. Sometimes general background knowledge plays the major part, as when we hear (67) and somehow guess that the intended D is the relation of visual contact between John and his killer as he lets him into the house. But it is plausible that the speaker's description of the *res* will usually be one important factor among others, and in the absence of other clues often the decisive one. This implies that, everything else being equal, the speaker's decision in (73) to refer to Ortcutt as 'the man Ralph saw at the beach' (rather than as 'Ortcutt' or 'the man he saw in the shadows') will bias the hearer towards the assumption that the intended acquaintance relation between Ralph and Ortcutt is the one established in the beach-encounter. And with this disambiguation, *Ralph thinks the man he saw at the beach is a spy* is false, *even though read de re*.

Another way of summarizing the suggestion I just made is this: there is not really just one *de re* reading (for a given constituent), but there are many—one

for each acquaintance relation that the context might supply. And some of those many, namely those where the acquaintance relation happens to include the subject's awareness that the *res* fits the same description used by the speaker, are very similar to the *de dicto* reading: more precisely, they entail it. In a way, I am blurring the distinction between *de re* and *de dicto* readings. But that may not be such a bad thing.[53] More often than not, the two are impossible to tell apart in practice anyway. When we hear somebody say that John thinks his dog is sick, do we understand that John takes himself to be in a world where the dog he has there is sick, or do we rather understand that John ascribes illness to his dog under some acquaintance relation or other? Under ordinary circumstances, where people know whether they own dogs, are acquainted with their dogs, and rarely encounter them unrecognizably disguised, one is true of John just in case the other is. So we couldn't really tell whether we construe the utterance *de dicto* and infer the truth of a *de re* reading, or the other way round. The present proposal, which implies that the unmarked reading is a *de re* reading that entails the *de dicto* reading, is equally compatible with our intuitions about those ordinary cases. And it also accounts for (73), where it looked at first like we prefer *de dicto*.

   This is all very sketchy. I am not yet ready to really endorse the view that *de re* construals are *ceteris paribus* preferred wherever possible, and that all presupposition inheritance from the complements of attitude verbs is due to this preference. But I think it is not a hopeless line to pursue. In the next section, I sketch an alternative which likewise strikes me as promising, and I will not attempt to choose at this point.

5.2 *Other people's beliefs and the nature of accommodation*

Recall the problem as we stated it with respect to sentence (66). When this sentence is uttered in a context where it isn't yet presupposed that John believes it had been raining, then we spontaneously accommodate the presupposition that it had (in fact) been raining. Suppose (in distinction to the alternative hypothesis of the preceding section) that the LF of (66) is what it appears to be, with the presupposition trigger genuinely in the scope of *believe*. And suppose further our account of the CCP of (66) is correct, i.e., it is defined exactly for those contexts in which it is presupposed that John believes it had been raining. Then we have a double puzzle of sorts: accommodating the presupposition that it had in fact been raining is predicted to be *neither necessary nor sufficient* to turn the context into one for which the CCP of (66) is defined. Yet, in practice it seems to be sufficient as well as highly preferred (if perhaps not downright necessary). Why?

   Part of the answer, if our analysis is at all to be saved, has to be that we actually accommodate at once *both* the presupposition that it has been raining

and the presupposition that John believes so. For without the latter, the CCP of (66) just wouldn't be defined. This is consistent with intuitive judgment in so far as we certainly wind up assuming that John believes so when we accept (66).[54] So the puzzle can be restated this way: why is it somehow easier to accommodate both that it rained and that John thought so at once, than to accommodate the latter alone?

It is useful here to recall a general point about accommodation:[55] assumptions to be accommodated are supposed to be uncontroversial and unsurprising. One may explicitly assert controversial and surprising things (in fact, one should), but to expect one's audience to accept them by way of accommodation is not good conversational practice. So when we hear (66) out of the blue, we know two things: first, as a matter of the semantics of this sentence, we know that it requires the presupposition that John believes it was raining. Second, we know that the speaker takes this to be uncontroversial and unsurprising. Now why would it be unsurprising that John has such a belief? The most natural guess is that it would be unsurprising because it was in fact raining and John was in an appropriate position to find out. Of course, these are not the only possible conditions under which someone might form a belief that it was raining; but they are the most *normal* conditions. Therefore, if accommodation is generally accompanied by a suggestion of unsurprisingness, then it is not so puzzling that these are the conditions which we spontaneously imagine to obtain. (This, I think, is what Karttunen had in mind in the passage I quoted in the Introduction.) Again, I am not confident that this is the right story, but it is prima facie plausible, and it gives us another way of maintaining our semantic analysis in spite of superficial appearances that attitude reports inherit the presuppositions of their complements.

6 CONCLUSION

Karttunen (1974: 188), having classified complementizable verbs into a number of subgroups according to their permeability for presuppositions of their complements, wrote: 'These distinctions are of course not arbitrary but presumably follow from the semantics of verb complementation in some manner yet to be explained.' What sort of an explanation was he hoping for? Presumably the kind that Stalnaker (1974, 1985, and elsewhere) proposed explicitly for the connective *and* and sketched for conditionals and belief-predicates. In the case of *and*, this was a simple and satisfying explanation indeed (1974: 210–11):

Karttunen defends the following [generalization]: ... the presuppositions of a conjunction are the presuppositions required by either of the conjuncts, *minus any required by the second conjunct which are entailed by the first* ... we can explain [this] generalization without postulating ad hoc

semantic or pragmatic rules. The explanation goes like this: ... when a speaker says something of the form A *and* B, he may take it for granted that A ... after he has said it. The proposition that A will be added to the background of common assumptions before the speaker asserts that B. Now suppose that B expresses a proposition that would, for some reason, be inappropriate to assert except in a context where A, or something entailed by A, is presupposed. Even if A is *not* presupposed initially, one may still assert A *and* B since by the time one gets to saying that B, the context has shifted, and it is by then presupposed that A.

The mere commonplace that asserting ϕ *and* ψ consists of asserting first ϕ and then ψ was sufficient to explain presupposition projection in conjunctions. In our terminology: one naïve look at *and* reveals what its CCP is and that this CCP makes the correct predictions.

That, of course, was an exceptionally easy case. Already the case of the conditional is much less obvious, notwithstanding Stalnaker's optimism in the following passage (1974: 211):

The analogous generalization about conditional statements is explainable on equally simple assumptions. Here we need first the assumption that what is explicitly *supposed* becomes (temporarily) a part of the background of common assumptions in subsequent conversation, and second that an *if* clause is an explicit supposition. Again, Karttunen's generalization is derived from these obvious assumptions.

What exactly is the role of supposing the antecedent in the overall context change, i.e., why is it necessary to do so in order to calculate the information conveyed by the whole conditional? What does this supposing amount to when the conditional is a counterfactual? What else happens after the supposition of the antecedent, in particular, what do we do with the consequent? In short, unlike the earlier story about *and*, this description of the CCP of *if* leaves a lot unsaid, and it is not so immediately evident how it should be completed in such a way that it predicts both the informational content of conditionals and their presuppositions. The following description of the CCP of *believe*, though still not complete,[56] is more nearly so (Stalnaker 1988):

What Phoebe believes, or is assumed to believe, may be different from, or incompatible with, what a speaker talking about Phoebe's beliefs believes or assumes. The relevant derived context will be ... the set of all possible situations that might, for all the *speaker* presupposes, be compatible with Phoebe's beliefs. This set of possible situations is the derived context for interpreting the clauses that are intended to express the contents of Phoebe's beliefs ... All of the ways that ordinary contextual information constrains and guides the interpretation of assertions ... will also be ways in which derived contexts constrain and guide the interpretation of embedded sentences which ascribe or deny beliefs . . . [for example], presupposition requirements: Just as 'Harry regrets accepting the bribe' is appropriate only in a context in which it is presupposed that a bribe was offered, and that Harry accepted it, so the statement 'Phoebe believes that Harry regrets accepting the bribe' requires a derived context in which it is presupposed that a bribe was offered and Harry accepted it. That is to say, it must be presupposed—taken by the speaker to be common ground—that Phoebe believes that a bribe was offered, and that Harry accepted it.

But numerous questions arise here as well when we try to extrapolate to other attitude verbs.

Well, I tried to complete these sketches, and it turned out to be harder than I thought. I had to set my sights low and got around only to two or three of the many verb types that Karttunen included in the classification referred to above. Even with those two or three, I barely scratched the surface and left many open problems and dangling stipulations. I set out to support the hypothesis that all presupposition projection was just a by-product of an independently plausible account of context change. I don't know if I have done more to support it than to cast doubt on it, but at least I have given it more concrete shape.

Acknowledgements

IRENE HEIM
Department of Linguistics and Philosophy,
MIT 20D-219
Cambridge, MA 02139
USA

NOTES

1 The main challenge was the alternative theory of Gazdar (1979), which relies heavily on cancellation of presuppositions. See Soames (1979, 1982, 1989) for critical discussion of that approach, as well as (more recently) van der Sandt (1989) and Zeevat (1991).

2 I follow here primarily the theory of Heim (1988: chapter 3, 1983). The framework of van der Sandt (1989, 1990), based on Discourse Representation Theory (Kamp (1981)), is very similar and probably equivalent in all respects relevant to this article. All current versions of such theories are descendants, in some sense, of the approach to presupposition projection that was urged by Stalnaker (1973, 1974) and Karttunen (1974).

3 Or whatever grammatical level(s) is (are) relevant to semantic interpretation.

4 Karttunen takes a context to be a set of logical forms (1974) or a set of propositions (1973b). In that sense, a context is not identical to a set of worlds, but it uniquely determines one, namely the set of worlds where all its elements are true. Stalnaker (1979) calls this the *context set*. For our purposes, there is no need to distinguish between different contexts that determine the same set of worlds, so we might as well identify contexts with their context sets.

5 The general framework leaves open that the definedness of a CCP might depend on properties of the input context other than the fact that it has certain entailments. I don't know if there are actual instances of this. In this article, at any rate, I only consider sentences whose CCPs have what I call 'purely presuppositional'

definedness conditions. These are the sentences ϕ for which there is some proposition p such that, for any c, c $+ \phi$ is defined iff c entails p (including sentences with 'no presuppositions', in which case the p in question is W).

6 Thus what I mean by 'uncancellability' is, in the terminology of Soames (1989), that there is no 'de jure accommodation'. Any accommodation there might be is 'de facto accommodation'.

7 By an 'empty' context, I mean one that is empty of any information. In other words, it is the set W of all possible worlds—*not* the empty set $\varnothing$; hence the quotation marks. (The terminology makes literal sense when contexts are construed as sets of propositions, and that's where it comes from.)

8 This corresponds to Karttunen's (1974) 'c ∪ {ϕ}'—recall that he takes contexts to be sets of logical forms.

9 This is not to be confused with my rash claim in Heim (1983) that CCPs are fully predictable from truthconditional properties, for which I was rightly taken to task by Soames (1989) and Mats Rooth (personal communication in a letter dating from 1986).

10 See Hintikka (1969) and Lewis (1973: 7).

11 This formulation of the rule is due to Hans Kamp (personal communication, Saturday afternoon, November 9, 1985, Cognitive Science Center seminar room at the University of Texas, Austin). See below for discussion of how it relates to Karttunen's rule (3).

12 Kripke (1990; and as cited in Soames (1989: note 54)). A common alternative assumption is that *too* triggers an *existential* presupposition, e.g. *Susan$_F$ is here too* would presuppose that someone other than Susan is here. But Kripke has argued persuasively that this is not quite correct and that words like *also*, *too*, *again* (and maybe many other presupposition triggers) have an essentially anaphoric-deictic semantics. (Some older treatments, e.g. Green (1968) as cited in

Karttunen (1974: 184), agree with Kripke, but the existential version somehow became more widespread.)

13 A problem which I set aside here: unlike *only* and *even*, *too* sometimes associates with a 'focus' that couldn't possibly be prosodically prominent because it is phonetically null. For instance, we are forced to this analysis for one of the readings of *John wants to come too*, the one where it means that John wants it to be the case that he comes in addition to so-and-so. Given the semantic rule in the text, the representation for this reading must be *John wants [PRO$_F$ to come too$_j$]*. (If we designate the overt NP *John* the focus, *too* must be attached in the matrix clause, and the meaning is that John, in addition to so-and-so, wants to come, which is also present but different.)

14 The deviance is, of course, only under the construal indicated, with *too* disambiguated as 'in addition to Mary'. If there is an alternative antecedent instead of *Mary* available, the sentence may be fine.

15 For this reason, it is not even possible to rescue the example by accommodation. What we would have to accommodate to ensure definedness, i.e. that John believes Mary to be here, contradicts what we have just been explicitly told.

16 This assumption follows directly under 'semantic' accounts of presupposition (3-valued or with truth-value gaps), and it is also part of Gazdar's (1979) theory.

17 See note 5.

18 Stalnaker's (1988) notion of the 'derived context' for *believe*-sentences also corresponds to the union of the $Dox_a(w)$ across c; in fact he explicitly defines it this way: 'The relevant derived context will be determined by the basic context in the following way: for each possible situation in the basic context, Phoebe will be in a definite belief state which is itself defined by a set of possible situations—the ones compatible with what Phoebe believes in that possible situation. The union of all the possible belief states will be the set of

all possible situations that might, for all the *speaker* presupposes, be compatible with Phoebe's beliefs. This set of possible situations is the derived context for interpreting the clauses that are intended to express the contents of Phoebe's beliefs.' But Stalnaker also does not give a rule that actually specifies the *outcome* of the context change effected by a belief-sentence. As far as I can see, one cannot write such a rule without referring to the individual sets $Dox_a(w)$ rather than just to their union.

19 The significance of this type of counter-examples to Karttunen is heightened by the fact that it seems to support Gazdar: his mechanism of cancellation by a conflicting conversational implicature covers (28) along with (2).

20 My version here is actually a cross between the two authors' versions. I depart from Lewis in making the Limit Assumption, but I don't make what Lewis calls Stalnaker's Assumption. In other words, I assume that for each world w and each contingent proposition p there is a non-empty set of p-worlds which are maximally similar to w, but this need not be a singleton. Evidently, it is important either to justify these choices or show that my aims in this paper do not really depend on them. I must leave both to future work.

21 Similar examples were also discussed by Janet Dean Fodor (1979).

22 Stalnaker goes right on to say that this is a specific property of indicative, as opposed to subjunctive, conditionals. See below.

23 See, e.g., Karttunen (1973a, 1974).

24 (36) is not fully equivalent to the standard rule for presuppositions of conditionals: it predicts that a presupposition of ψ may in principle get filtered away even though not entailed by $c + \phi$, as long as it is entailed by $\cup\{Sim_w(c + \phi): w \in c\}$. I am unable to give a concrete example of diverging predictions, however, because it is not intuitively clear to me at present what a context c has to be like in order for

$\cup\{Sim_w(c + \phi): w \in c\}$ to be a proper subset of $c + \phi$.

25 This is actually what Stalnaker does, but in his case, he is forced to it by Stalnaker's Assumption, which we did not adopt.

26 In the murder example, Stalnaker spontaneously changes tense: 'I never want*ed* the murder to have been committed.' But, of course, the question of what I wanted in the past, before I knew that the murder had been committed, is quite separate. (39) has no difficulty with the possibility that *a wants ϕ* is true or false at one time and then *a believes ϕ* is true at a later time. Nor is there a problem, of course, with wanting ϕ before one comes to believe ~ϕ. *I didn't want him to do it, but I saw he was doing it anyway* is fine, but this is presumably because the reference time of the *want*-clause precedes that of the *see*-clause. Such an example does not show that '*a wants ϕ*' and '*a believes ~ϕ*' can be true *at the same time*. Likewise, it is not a counterexample to our analysis that I can say coherently: *I know he is in and I want him out.* Here it is important to make the reference times of the embedded clauses explicit: this sentence says that I know that he is in *now* and I want him to be out *in the immediate future*. (In other words, the complements of the two clauses are not negations of each other.) Complements of *want* always have a futurate interpretation. (Perhaps this is a general property of *for*-infinitivals; see, e.g., Stowell 1982.)

27 One may object to my reanalysis of Stalnaker's judgments as follows: if we *have* to say which one of the two sentences *I want to have been sick* and *I want not to have been sick* is true, we have a firm intuition that it is the latter. What little strangeness there may be in both of them does not impede this judgment, but our current proposal does not account for it, because it predicts exactly the same status—undefined—for the two sentences' CCPs. I am not sure how best to respond to this objection. Perhaps what is going on here is that we tacitly reinterpret these

28 sentences with something like the semantic rule for factive and counterfactual desire predicates (see below) and base our truthvalue judgments on this reinterpretation.

28 (43) suppresses the temporal parameter. Taking that into account, we would have: for any $w \in W$, time t, $F_a(w, t) = \{w' \in W$: w' is compatible with everything that a in w at t believes to be the case no matter how he chooses to act after $t\}$.

29 There is a problem: *Patrick intends to sell his cello tomorrow* is predicted to presuppose that he will have a cello tomorrow no matter how he acts. But he *could* have decided to sell it today, in which case he wouldn't have one tomorrow anymore. This must be fixed somehow by restricting the relevant actions to those at or after the reference time of the complement.

30 Cf. Lewis (1986: 34-5).

31 As I was reminded by Carl Pollard (personal communication); (45) is his example.

32 As in Roberts' treatment of modal subordination. See Roberts (1989, 1991).

33 This is, in effect, the proposal of Cresswell (1988).

34 Another example of this kind: ??*John prefers for you to have already received his letter, even though he is sure you were very upset by it.* Perhaps *will*-sentences are particularly easy to read as containing covert restrictions. Or perhaps there is a blurring of indicative and counterfactual mood in the future.

35 What I mean here by the 'irrealis mood' is, in morphological terms, homophonous with the past tense, except for the 1st and 3rd persons singular of the verb *be*, whose irrealis forms are *were* rather than *was*. It seems to be the same mood that shows up in the antecedents of counterfactual conditionals. Wherever I refer to the verb *wish* from now on, I mean the *wish* that governs a tensed complement clause in the irrealis. (As opposed to, say, *wish* + infinitive, which I disregard here.) The irrealis mood itself (i.e., the suffix of

non-concatenative morpheme that realizes it on the embedded verb) will receive no semantic interpretation at all; I treat it as a mere surface phenomenon (analogous to, say, case on an NP), and only the superordinate verb that governs it is a semantic unit. (Likewise, the mood marking in the antecedent of a counterfactual conditional is semantically redundant; only the counterfactual modal—*would* or *might*—in the consequent is interpreted.) I don't thereby mean to deny, of course, that the lexical items which govern the irrealis mood form a natural semantic class.

36 An alternative might be to undo the amendment in (40) and return to (34). Then $c' + he\ wishes\ Susan_F\ were\ here\ too_1$ would always be c' again. Still, the truthconditions for (53) would be trivialized, so this is not a way out.

37 By the usual definition of factivity, it moreover presupposes that the complement is in fact true (not just believed to be so by the subject)—though some have argued against that stronger presupposition on the grounds that one can say things like: *Mary, who was under the illusion that it was Sunday, was glad that she could stay in bed* (Klein 1975, as cited in Gazdar (1979: 122)). What matters here is only the assumption that a *is glad that ϕ* presupposes *at least* that a believes ϕ, which has not, to my knowledge, been disputed.

38 For instance, Kratzer (1981: 69) proposes that the modal base for counterfactual conditionals is 'empty'.

39 Except perhaps by way of local accommodation in the sense of Heim (1983).

40 An alternative assumption would be that it is just *some* context within these limits, and other contextual clues determine which particular one it is for each given utterance of a counterfactual conditional. This would then be yet another source of vagueness, on top of that already due to the flexibility of criteria for similarity. I don't see at this point how this option

could be empirically distinguished from the one adopted in the text.

41 As David Dowty (personal communication) reminded me, (62) and (36) do not predict that we must use a subjunctive conditional when the antecedent is presupposed to be false and an indicative one otherwise. A closer look reveals that we predict one direction of this generalization: if the antecedent is incompatible with the common ground, the indicative conditional is infelicitous and only the subjunctive one is permitted. (This follows because of the amendment in (40), which makes the selection function undefined for the inconsistent proposition.) On the other hand, nothing I have said so far implies that one couldn't use the subjunctive conditional even when the antecedent is compatible with the common ground. Interestingly, there are some cases where this systematically occurs; see Karttunen & Peters (1979) and especially Stalnaker (1975) for examples. But it is not an option that is always freely available, and to capture this we must impose an additional felicity condition on the choice of the counterfactual. A rough proposal, inspired by Stalnaker (1975), is that *if ϕ would ψ* is felicitous in a context c only if there is at least one world $w \in c$ such that $Sim_w(rev_\phi(c) + \phi)$ is not a subset of $c + \phi$. (A rationale for this might be that the counterfactual is the marked choice and thus pre-empted by the indicative conditional when one might as well have used the latter.)

42 Proof: suppose the CCP of ϕ is defined for exactly those contexts which entail p. The definition in (61) thus amounts to $rev_\phi(c) - \cup \{X: c \subseteq X \subseteq p\}$. We need to show that $rev_\phi(c)$ entails p iff c does. First, assume that c does not entail p. Then no superset of c does, so $\{X: c \subseteq X \subseteq p\} - \varnothing$ and its union is W and doesn't either. Second, assume that c does entail p. Then $\cup\{X: c \subseteq X \subseteq p\} - p$, which of course entails p.

43 Cf. n. 5.

44 This deviance is, of course, unsurprising if every presupposition is also an entailment of the minimal sentence that carries it (as in Gazdar 1979), because this assumption alone suffices to predict the sentence to be contradictory. But notice that we get the same deviance in slightly more complex cases, such as *If John weren't also in the phone booth, Mary would be outside. John is also in the phone booth* may both presuppose and entail that Mary is, but *not (John is also in the phone booth)* presumably only presupposes it.

45 Why exactly should we be allowed to take this for granted? One possibility is to stipulate a further felicity condition in sentences of the form *a wishes ϕ*, namely that they fit only in contexts where it is presupposed that *a* believes *not ϕ*. Alternatively, we might explore weaker conditions analogous to Stalnaker's proposal for subjunctive conditionals; cf. n. 41. So (63) might not always reduce to (64), but it will in typical contexts.

46 Unlike Karttunen, Gazdar (1979) works with a theory according to which presuppositions are cancellable. Specifically, they get cancelled whenever they contradict an assertion, conversational implicature, or other presupposition of the same or preceding sentence(s). This is what he claims happens in (2). The initial sentence of this text, *Patrick is under the misconception that he owns a cello*, entails that Patrick does not have a cello, hence conflicts with the potential presupposition of the subsequent sentence, and thus cancels it. In other cases, a conversational implicature or other potential presupposition might be responsible for the cancellation.

47 There may be other factors; e.g., otherwise non-factive verbs might sometimes have factive readings. I don't know whether this occurs with attitude verbs though. It does seem to happen with verbs of saying. Take Gazdar's example (i).

(i) The salesman didn't tell me that my camera was suitable for color too.

It is true that we conclude that the speaker has a camera and that it is suitable for black-and-white photography. But, as Berman (1989) points out, we spontaneously conclude even more, namely that the speaker's camera is suitable for color, which is not a presupposition but the content of the complement clause. In other words, *tell* seems to be read factively in this example. But once a predicate is factive, then its being also a hole presents no additional mystery. Consider, for instance, the following meaning rule for the factive verb *know*, which implements the common idea that factives presuppose their complements.

(ii) c + *John knows that* ϕ is undefined unless $c - c + \phi$.

Where defined, c + *John knows that* $\phi - \{w \in c: Dox_j(w) + \phi - same\}$.

This requires, among other things, that $c + \phi$ be defined, thus that c satisfy any presuppositions of ϕ. This is not a story, however, that could be extended to account for all cases where attitude verbs act like holes. If we control for factivity, as in (iii) below, there is still a spontaneous interference that the speaker has a camera and it's suitable for black-and-white.

(iii) (This salesman told me a lot of lies, but at least) he didn't tell me that my camera was suitable for color too.

48 See Quine (1956), Kaplan (1969), Lewis (1979), and others. I leave open here how *de re* construals are represented at LF; something along the lines of Creswell & Stechow (1982) should suit my purposes. Of course, transposing their proposal into the present framework would first require an account of variables and quantification in a context change framework. This also goes beyond the scope of the present article, but see Heim (1983, 1988).

49 See von Stechow (1981) for an explicitly descriptional analysis of factives, for instance.

50 As Rob van der Sandt pointed out to me, examples of this sort are discussed in Fauconnier (1984); see also Zeevat (1991).

51 One reason for my scepticism is that I don't see off-hand how this approach throws light on the appropriateness of *also* in the following minimal variation of (71): imagine John and Mary competed for one job, and everybody, including the parents, knew this.

(i) John: I_1 got the job.
    Mary: My parents think that I_F also$_1$ got it.

Why isn't *also$_i$ get the job* a description that fits the property of getting the job when it happens that x_i gets the job? We might amend the proposal so that a property only falls under the description *also$_i$ ζ* if it is true of x_i and at least one other individual. But this is not quite what we need here, since the intuitive reason why (i) is out is not that *Mary* believes only one person got the job, but that *her parents* believe this.

52 This has also been argued for by von Stechow (1984) and it is tacitly taken for granted in Higginbotham (1989).

53 It may also make it easier to reconcile the two-way *de re/de dicto* ambiguity of the standard theory with finer classifications such as the four-way distinction in Fodor (1979: 229). (I owe this reference to Angelika Kratzer.)

54 In a theory like Gazdar's, this could be attributed to the mere fact that (66) also *asserts* that John believed there was rain. But we also infer such a belief in an analogous sentence like (i), where it can't have come as an entailment of the assertion. (That way, we'd only get that John *wants* there to have been rain.)

(i) John wants it to stop raining soon.

So there is evidence independent of our analysis that the presupposition that the subject believes the complement's presupposition is generally accommodated in addition to the presupposition that it is true in fact.

55 See, e.g., Soames (1989: 567).

56 It is not quite complete for the reason given in n. 18.

REFERENCES

Asher, Nicholas (1987), 'A typology for attitude verbs and their anaphoric properties', *Linguistics and Philosophy*, **10**, 125–98.

Berman, Stephen (1989), 'An analysis of quantificational variability in indirect *wh*-questions', in *Proceedings of the Eighth West Coast Conference on Formal Linguistics*, Stanford Linguistics Association, Stanford, 29–43.

Cresswell, Maxwell J. (1988), 'Anaphoric attitudes', in *Semantical Essays: Possible Worlds and Their Rivals*. Kluwer, Dordrecht.

Cresswell, Maxwell J. & Arnim von Stechow (1982), 'De re belief generalized', *Linguistics and Philosophy*, **5**, 4: 503–35.

Fauconnier, Gilles (1984), *Mental Spaces*, MIT Press, Cambridge, MA.

Fodor, Janet Dean (1979), *The Linguistic Description of Opaque Contexts*, Garland, New York (reprinted from 1970 MIT Ph.D. thesis).

Gazdar, Gerald (1979), *Pragmatics*, Academic Press, New York.

Green, Georgia (1968), 'On *too* and *either*, and just on *too* and *either*, either', in B. Darden *et al*. (eds), *Papers from the Fourth Regional Meeting of the Chicago Linguistics Society*, University of Chicago, Chicago.

Heim, Irene (1983), 'On the projection problem for presuppositions', in D. Flickinger *et al*. (eds), *Proceedings of the Second West Coast Conference on Formal Linguistics*, Stanford University, 114–25 (reprinted in S. Davis (ed.) (1991), *Pragmatics*, Oxford University Press, New York, 397–405).

Heim, Irene (1988), *The Semantics of Definite and Indefinite Noun Phrases*, Garland Press, New York (reprinted from 1982 University of Massachusetts, Amherst, Ph.D. thesis.)

Higginbotham, James (1989), 'Reference and control'. *Rivista di Linguistica*, **1**, 2: 301–26.

Hintikka, Jaakko (1969), 'Semantics for propositional attitudes', in J. W. Davis *et al*. (eds), *Philosophical Logic*, Dordrecht: Reidel, 21–45.

Kamp, Hans (1981), 'A theory of truth and semantic representation', in J. Groenendijk, T. Janssen & M. Stokhof (eds), *Formal Methods in the Study of Language*, Mathematical Centre, Amsterdam, 277–322.

Kaplan, David (1969), 'Quantifying in', in D. Davidson & J. Hintikka (eds), *Words and Objections*, Reidel, Dordrecht, 178–214.

Karttunen, Lauri (1973a), 'Presuppositions of compound sentences', *Linguistic Inquiry*, **4**: 169–93.

Karttunen, Lauri (1973b), 'The last word', mimeograph, University of Texas, Austin.

Karttunen, Lauri (1974), 'Presupposition and linguistic context', *Theoretical Linguistics*, **1**, 181–94.

Karttunen, Lauri & Stanley Peters (1979), 'Conventional implicature', in C.-K. Oh and D. A. Dinneen (eds), *Syntax and Semantics 11: Presupposition*, Academic Press, New York, 1–56.

Klein, Ewan (1975), 'Two sorts of factive predicate', *Pragmatic Microfiche*, **1**, 1: B5–C14.

Kratzer, Angelika (1981), 'The notional category of modality', in H. J. Eikmeyer & H. Rieser (eds), *Words, Worlds, and Contexts*, de Gruyter, Berlin.

Kripke, Saul (1990), Talk at workshop on anaphora, Princeton University.

Lewis, David (1973), *Counterfactuals*, Harvard University Press, Cambridge, MA.

Lewis, David (1979), 'Attitudes de dicto and de se', *Philosophical Review*, **88**, 513–43.

Lewis, David (1986), *On the Plurality of Worlds*, Blackwell, Oxford.

Quine, W. V. O. (1956), 'Quantifiers and propositional attitudes', *Journal of Philosophy*, **53**.

Roberts, Craige (1989), 'Modal subordination and pronominal anaphora in discourse', *Linguistics and Philosophy*, **12**: 683–721.

Roberts, Craige (1991), 'Domain restriction in dynamic semantics', MS, Ohio State University.

Sandt, Rob van der (1989), 'Presupposition

and discourse structure', in R. Bartsch, J. van Benthem & P. van Emde Boas (eds), *Semantics and Contextual Expression*, Foris, Dordrecht.

Sandt, Rob van der (1990), 'Anaphora and accommodation', in workshop reader for workshop on Presupposition, Lexical Meaning and Discourse Processes, University of Nijmegen.

Soames, Scott (1979), 'A projection problem for speaker presuppositions', *Linguistic Inquiry*, **10**, 623–66.

Soames, Scott (1982), 'How presuppositions are inherited: a solution to the projection problem', *Linguistic Inquiry*, **13**: 483–545.

Soames, Scott (1989), 'Presupposition', in D. Gabbay & F. Guenthner (eds), *Handbook of Philosophical Logic*, vol. IV: 553–616. Reidel, Dordrecht.

Stalnaker, Robert (1968), 'A theory of conditionals', in N. Resher (ed.), *Studies in Logical Theory*, Blackwell, Oxford.

Stalnaker, Robert (1973), 'Presuppositions', *Journal of Philosophical Logic*, **2**: 447–57.

Stalnaker, Robert (1974), 'Pragmatic presuppositions', in M. Munitz & D. Unger (eds), *Semantics and Philosophy*, New York University Press, New York, 197–213.

Stalnaker, Robert (1975), 'Indicative conditionals', *Philosophia*, **5**, 3: 269–86.

Stalnaker, Robert (1979), 'Assertion', in P. Cole (ed.), *Syntax and Semantics 9: Pragmatics*, Academic Press, New York, 315–32.

Stalnaker, Robert (1984), *Inquiry*, MIT Press, Cambridge, MA.

Stalnaker, Robert (1988), 'Belief attribution and context', in R. Grimm & D. Merrill (eds), *Contents of Thought*, University of Arizona Press, Tucson, pp. 140–56.

Stechow, Arnim von (1981), 'Presupposition and context', in F. Günthner and U. Mönnich (eds), *Philosophical Logics*.

Stechow, Arnim von (1984), 'Structured propositions and essential indexicals', in F. Landman & F. Veltman (eds), *Varieties of Formal Semantics*, Foris, Dordrecht, 385–403.

Stowell, Tim (1982), 'The tense of infinitives', *Linguistic Inquiry*, **13**: 561–70.

Zeevat, Hendrik Willem (1991), *Aspects of Discourse Semantics and Unification Grammar*, University of Amsterdam doctoral thesis.