# Understanding Assertion

### Scott Soames

#### INTRODUCTION

In his groundbreaking 1978 article, "Assertion," Robert Stalnaker presents an elegant model of discourse designed to solve philosophical problems arising, in part, from his identification of propositions with functions from possible world-states to truth values, and his restriction of the epistemically possible to the meta-physically possible. Among these problems are those posed by Kripkean examples of the necessary aposteriori. Being necessary, all such examples are seen by Stalnaker as semantically expressing the same trivial, universally known, apriori truth. Nevertheless, assertive utterances of them often result in the assertion of propositions that are both highly informative and knowable only aposteriori. A central task of "Assertion" is to explain how this can be so.

Stalnaker's model is based on the insight that considerably more goes into determining what is said by an assertive utterance than the meaning of the sentence uttered. Additional assertion-determining factors include (i) objective features of the context of utterance, such as the speaker, audience, time, place and world-state of the context, (ii) general conversational rules, including those against asserting what is already known or presupposed to be true (or to be false), and (iii) salient beliefs and assumptions known to be shared by conversational participants. These latter encompass beliefs and assumptions about who is speaking to whom, what words are being uttered and what they mean, what is happening in and around the speech situation, the topic of conversation, what has already been established or taken for granted, and what remains on the conversational agenda. Imagine, for example, a speaker who utters She is late in response to the entrance of a woman who comes into a meeting after it has already begun, attracting everyone's attention. The speaker relies on the fact that everyone will recognize him to be saying, of the woman who just entered, that she is late for the meeting. In this case, obvious facts about the conversational context interact with the meaning of the sentence uttered to determine the proposition asserted.

Other cases are more indirect, and may involve reinterpretation of what the speaker might, at first, appear to have said. For example, after listening to the

#### Understanding Assertion

remarks of a well-known campus orator, Mary might turn to her companion and say "Norman really is God's fountain pen, isn't he?" knowing full well that her companion won't take her to be asserting the proposition semantically expressed by the sentence she utters-one which predicates of Norman the property of being a certain kind of artifact (used to write by depositing ink on paper) possessed by God. Since that proposition is egregiously false, Mary realizes that her conversational partner will recognize that she is not committing herself to it, and so will look for an alternative interpretation of her remark. Depending on what else is taken for granted in the conversation, Mary might be taken to have asserted (i) that God really is using Norman to communicate his thoughts and desires, (ii) that (God aside) Norman really does have the truth on the matters about which he is speaking, or (iii) that Norman is a blow-hard who takes himself to be an authority, even though he really isn't (irony). This variation is not a sign of indexicality. Although Mary's utterance would result in assertions of different propositions in different contexts, this is not a matter of semantics. It is a matter of the way in which the meaning of the sentence uttered interacts with both principles governing discourse and the background beliefs and assumptions of conversational participants.

Similar principles govern the reinterpretation of utterances of trivially obvious, literal truths. For example, a candidate awaiting the results of an election after the polls have closed might respond to an early bit of unfavorable news by saying "What will be will be." In such a case, the speaker understands that his audience will not take him to have asserted a trivial tautology, but will instead interpret him to have said something significant—typically something to the effect that since the outcome is out of his control, he is prepared to accept whatever the result proves to be.

These examples are simple and relatively uncontroversial. However, as Stalnaker correctly recognizes, they are part of a larger, more systematic picture. A central message of "Assertion" is that there is often a substantial gap between the propositions assigned to sentences by a correct semantic theory and those asserted by utterances of these sentences in different contexts. Because of this, continued progress in solving problems in semantics and philosophical logic depends on our coming to have a better understanding of the ways in which semantic and nonsemantic factors interact in determining what is asserted and conveyed by utterances, and in guiding rational discourse and inquiry. This is the central insight, and seminal contribution, of Stalnaker's inquiry—and one with which I fully agree.

Nevertheless, I am skeptical about some of the burdens taken up in the article—in particular, the attempt to render familiar examples of the necessary aposteriori compatible with the restriction of epistemically possible to metaphysically possible world-states, and the identification of the objects of assertion with functions from such states to truth values. Thus, my first task, after reconstructing Stalnaker's discourse model, will be to demonstrate why it cannot be used to

reconcile the necessary aposteriori with his antecedent conceptions of possibility and of propositions. Instead, I will argue, a proper understanding of the necessary aposteriori requires the restriction of epistemic possibility to metaphysical possibility to be abandoned, and the existence of epistemically possible worldstates that are metaphysically impossible to be recognized. Having reached this point, I will next isolate inherent limitations of the discourse model itself by showing that *de re* belief and the nontransparency of meaning lead to insoluble problems—even after the model has been improved by substituting the broader class of epistemically possible world-states for the narrower class of metaphysically possible world-states. Finally, I will draw lessons from these problems, distinguish aspects of Stalnaker's discourse model that need to be revised from those that should be retained, and suggest how further progress can be made in understanding the ways in which semantics and pragmatics interact in assertion.

#### STALNAKER'S MODEL OF DISCOURSE

According to the model, conversations take place against a set of background assumptions shared by the conversational participants which rule out certain possible world-states as not obtaining, or "being actual."<sup>1</sup> As the conversation proceeds, and assertions are made and accepted, new propositions are admitted into the set of shared background assumptions, and the set of world-states that remain compatible with what has been assumed or established shrinks. This set is called the *context set* (at any given point in the conversation). The aim of further discourse is to further narrow down this set of possibilities, within which the actual state of the world—the maximally complete property that the universe really instantiates—is assumed to be located. When one asserts p, the function of one's assertion is to shrink the context set by eliminating from it all world-states in which p is not true.

Stalnaker postulates three rules governing assertion.<sup>2</sup>

R1. A proposition asserted should always be true in some but not all of the possible world-states in the context set.

224

FN:1

<sup>&</sup>lt;sup>1</sup> Here, and throughout, I use the term *possible world-state* instead of the more familiar *possible world* to reflect my view, shared with Stalnaker, that the items under discussion are not alternate concrete universes, but ways the world could have been, i.e. maximally complete properties the universe could have instantiated. To say of such a state that it "is actual" is to say that the universe instantiates it. I also limit myself to contexts of utterance that Stalnaker calls *nondefective* (those in which the possible world-states that the speaker takes to be compatible with everything believed and assumed by conversational participants at a given point in the conversation are the same as those that the hearers take to have this property), plus contexts that he calls *close enough* (those in which differences between speakers and hearers on this point don't arise, or affect the course of the conversation.)

<sup>&</sup>lt;sup>2</sup> Stalnaker (1999, 88).

Understanding Assertion

- R2. Any assertive utterance should express a proposition, relative to each possible world-state in the context set, and that proposition should have a truth value in each possible world-state in the context set.
- R3. The same proposition should be expressed relative to each possible worldstate in the context set.

The rationale for R1 is that a proposition true in all world-states of the context set would be uninformative, and so would fail to perform the essential function of assertion, which is to narrow down the range of world-states that conversational participants take to be candidates for being the way the world actually is. By the same token, a proposition false in all world-states in the context set would contradict what has already been conversationally established. Since it would eliminate the entire context set, it would also fail to narrow down the range in which the actual world-state is to be located. Of course, this rule, like the others, allows for some flexibility in how it applies. If someone seems to say something that violates it, one may sometimes conclude that no violation has really taken place because the context set isn't quite what one originally thought, or because the speaker didn't really assert, or mean, what he at first seemed to assert or mean. This is not to say that violations never occur, but it is to say that common knowledge of the rule can sometimes be exploited for conversational purposes—as when a speaker deliberately says something the literal interpretation of which would violate the rule, knowing full well that he will be reinterpreted in a certain obvious way so as to be seen as conforming to it.3

Stalnaker's rationale for R2 is that if an utterance violates it, then for some world-state w in the context set, the assertive utterance won't determine whether it should remain in the set, or be eliminated. If the sentence uttered does not express a proposition at w, or if it does express a proposition, but one for which no truth value—truth or untruth—is defined at w, then no verdict on whether w stays or goes will, Stalnaker thinks, be forthcoming. This is to be avoided.<sup>4</sup>

In explaining the rationale for R3 Stalnaker employs his notion of *the propositional concept associated with an assertion*. A propositional concept is very much like one of David Kaplan's characters. For Stalnaker, it is a function from worldstates, considered as possible contexts of utterance, to propositions—where propositions are taken to be nothing more than assignments of truth values to worldstates, considered as circumstances of evaluation. The propositional concept associated with an utterance of a sentence S at a certain moment m in a conversation is a function that maps each world-state w in the context set at m onto a proposition—which is simply an assignment of truth values to all world-states in the context set. This assignment of truth values is (implicitly) identified with the proposition that would be expressed by S at m, if the actual context of utterance were to turn out to be w.

<sup>3</sup> Ibid., p. 89. <sup>4</sup> Ibid., pp.89–90.

FN:3

Propositional concepts can be given pictorial representations, as is indicated by Stalnaker's matrix D.

D represents the propositional concept associated with the use of S at a moment m in which the context set consists of the world-states i, j, and k. D tells us (i) that if i is the state the world is actually in at m, then the proposition (semantically) expressed by the speaker's utterance of S is the proposition that assigns truth to every world-state of the context set, (ii) that if j is the state the world is actually in at m, then the proposition (semantically) expressed assigns truth to a the proposition (semantically) expressed assigns truth to k and falsity to i and j, and (iii) that if k is the state the world is actually in at m, then the proposition (semantically) expressed assigns falsity to i and truth to the other two world-states.<sup>5</sup> Stalnaker uses D to give the following rationale for R3.

To see why the principle must hold, look at the matrix for the propositional concept D. Suppose the context set consists of i, j, and k, and that the speaker's utterance determines D. What would he be asking his audience to do? Something like this: If we are in the world i, leave the context set the same; if we are in the world j, throw out worlds i and j, and if we are in world k, throw out just world i. *But of course the audience does not know which of those worlds we are in, and if it did the assertion would be pointless.* So the statement, made in that context, expresses an intention that is essentially ambiguous. Notice that the problem is not that the speaker's utterance has failed to determine a unique proposition. Assuming that one of the worlds i, j, or k, is in fact the actual world, then that world will fix the proposition unambiguously *The problem is that since it is unknown which proposition it is that is expressed, the expression of it cannot do the job that it is supposed to do.*<sup>6</sup>

The idea is that if R3 is violated, the conversational participants won't know which proposition is (semantically) expressed by the sentence uttered, because they won't know which world-state "is actual." But if the proposition asserted is always the one (semantically) expressed by the sentence uttered (in the context), then the conversational participants won't know what is asserted, and so will be at a loss as to how to update the context set and proceed with the conversation. This is the rationale for R3.

226

FN:5

<sup>&</sup>lt;sup>5</sup> Stalnaker (1978) is not fully clear about what status the propositions "expressed by" S at the world-states of the context set are supposed to have. Although they are not always the propositions that would be asserted, if those world-states were to obtain, they are often propositions semantically expressed in those eventualities—hence the parenthetical "semantically" above. However, there are exceptions to this—to which I will return—which prevent any such general identification.

<sup>&</sup>lt;sup>6</sup> Ibid., pp. 90–1, my emphasis.

Judith Jarvis Thomson chap11.tex V1 - May 30, 2006 10:53 A.M. Page 227

227

#### Understanding Assertion

#### PRESUPPOSITIONS OF THE MODEL

Before going further it is worth pointing out certain presuppositions of the model, and dealing with obvious worries that might arise. The model presupposes that speakers have a great deal of knowledge—(i)–(iii)—about possible world-states.

- (i) For every world-state w, conversational participants at a time t in the conversation know whether w is compatible with everything believed, established, or assumed in the conversation at t, and hence whether w is in the context set at t.
- (ii) For any sentence S that might be uttered at t, any world-state w in the context set at t, and any proposition p, if an utterance of S at t would express p, were it to turn out that w were actual (i.e. were w to turn out to be the world-state that is actually instantiated), then conversational participants know that this is so.
- (iii) For any proposition p and world-state w, conversational participants know the truth value of p in w, i.e. they know what the truth value of p would be were w to be actual.

It is natural to wonder whether speaker-hearers really have all this knowledge of world-states.

One worry concerns what might be called the size of world-states. Each worldstate encodes a massive amount of information about the universe-far too much for our minds to encompass. But if that is so, how are we able to know anything significant about such entities? The answer is that the knowledge of world-states required by the model is not very extensive. Although each world-state encodes a massive amount of information, only a tiny fragment of it will be relevant in any given conversational setting. Because of this, we can ignore differences among world-states that are irrelevant to our conversational purposes. For example, if our conversation has been exclusively about the 2004 American League Championship Series between the Boston Red Sox and the New York Yankees, we can form equivalence classes of world-states that agree on their accounts of the series, while differing arbitrarily on extraneous matters. In representing the context set and propositional concept for an utterance u, we can then take each 'w' as standing for one of these equivalence classes. Knowing of each equivalence class (and thereby of each member in it) that it is compatible with everything assumed, established, or believed at the time of u can then be assimilated to knowing, of the account of the series on which all members of the class agree, that it is compatible with all this background information. If this is correct, then the worry about size disappears. Although nothing I have said guarantees that speaker-hearers can, in general, be relied upon to have knowledge of types (i)-(iii), presupposed by the model, the sheer quantity of information encoded in world-states is not itself an obvious barrier.

The next thing to notice is that the model presupposes systematic *de re* knowledge of world-states. For each relevant world-state w, sentence S, and proposition p, speakers in a conversation C are said to know (i) that w is (or is not) compatible with the background assumptions of C, (ii) that an utterance of S would (or would not) express p, if w were to turn out to be actual, and (iii) that p would be true (false), if w were to obtain. In each case, an occurrence of the variable 'w' appears inside the content clause of the knowledge ascription, while being bound by a quantifier outside the clause. Since this is the mark of *de re* knowledge-ascriptions, the model attributes far-reaching *de re* knowledge of world-states to conversational participants. How should we think of this?

Consider the following example: A is speaking to B at a conference on the philosophy of language. A points at a man across the room and says "He teaches at UCLA." Suppose, for whatever reason, that the following world-states are members of the context set.

- w1: A is pointing at David Kaplan, and Kaplan teaches (exclusively) at UCLA, and . . .
- w<sub>2</sub>: A is pointing at David Kaplan, and Kaplan teaches (exclusively) at USC, and . . .

Stalnaker's discourse model presupposes that conversational participants know of  $w_1$  and  $w_2$  that if they "are actual" (i.e. if either one obtains or is instantiated), then A's utterance will express the proposition  $p_k$  that David Kaplan teaches at UCLA. In addition, the model presupposes that conversational participants know of  $w_1$  that if it "is actual," then  $p_k$  will be true, while knowing of  $w_2$  that if it "is actual," then  $p_k$  will be true, while knowing of  $w_1$  and  $w_2$  are those indicated above, this *de re* knowledge of world-states amounts, essentially, to knowledge of the following propositions:

- a. that if A is pointing at David Kaplan, then A's utterance expresses the proposition that David Kaplan teaches at UCLA.
- b. that the proposition that David Kaplan teaches at UCLA is true, if Kaplan teaches (exclusively) at UCLA, while it is false, if Kaplan teaches (exclusively) at USC.

On this account, the *de re* knowledge of world-states presupposed by the model is pretty easy to come by. In the case of (a), it involves knowledge of David Kaplan, and of the meaning of the sentence uttered. In the case of (b), it is apriori knowledge that every speaker acquainted with Kaplan can be expected to have. Seen in this light, the presuppositions of the model may seem to be readily satisfiable.

There is, however, cause for concern. Typically, the *de re* knowledge of worldstates presupposed by the model will, as in the previous example, bottom out in

#### Understanding Assertion

ordinary de re knowledge of individuals (natural kinds, or other constituents of the world). In our example, knowledge of the world-states w<sub>1</sub> and w<sub>2</sub>, and of the proposition  $p_k$  that the latter would be expressed if either of the former were "actual," as well as knowledge of the truth values the proposition would have in those eventualities, is really nothing more than knowledge of David Kaplan (a) that if the speaker is pointing at him, then the speaker's utterance will express  $p_k$ , and (b) that if he teaches at UCLA, then  $p_k$  will be true, whereas if he teaches at USC,  $p_k$  will be false. However, *de re* attitudes of this sort are notorious for resisting the neat logical transitions presupposed by Stalnaker's model. For example, it is well known-from the discussion of puzzling Pierre in Kripke (1979), as well as from the discussions of other examples-that one can know of one and the same individual i that he is F and that he is G, without knowing (or being in a position to know) of i that he is both F and G. Similarly, one can know of i that he is F and that if he is F, then he is G, without knowing (or being in a position to know) of i that he is G; and one can know of i that he is F, while also knowing that S expresses the proposition that he is F, without knowing (or being in a position to know) that S expresses a truth. Cases like this pose a threat to Stalnaker's model of discourse.

This threat will be examined in due course. For now, we simply note it. Before moving on, we need to understand another aspect of the *de re* knowledge of world-states presupposed by the model—*de re* knowledge of the actual worldstate. According to Stalnaker, the point of rational inquiry and conversation is to reduce, as much as possible, the space of possible world-states within which the actual world-state is believed, or known, to be located. On this picture, the idealized goal of these activities is to eliminate all possible world-states but one, which can then be correctly identified as actual. Why is this desirable? Well, it is natural to think, an agent who correctly identifies the actual world-state @ is thereby in a position to know everything. This will be so if (i) coming to know of the world state @ that it is actual inevitably involves coming to know, for each genuine truth p, that p is true in @,and (ii) knowing that p is true in @ involves knowing, or being in a position to know, p. However, this reasoning is incorrect.

First consider (ii). A fundamental presupposition of the model is that for any world-state w in the context set, and any proposition p that might be asserted in the conversation, speakers and hearers know the truth value of p in w. In many conversations—those without false presuppositions or assertions—the actual world-state @ will be a member of the context set. It follows that, in these conversations, there will be many propositions p that agents know to be true in @, without knowing (or having any way of coming to know) p. Thus, knowing that p is true in the world-state @, which actually obtains, is not sufficient for knowing p.

This brings us to (i). It is tempting to think that the reason an agent can know, of @, that p is true in it without knowing p is that, in cases like this, the agent does not know, of @, that it really obtains, or is instantiated. Were the agent to know this, the thought continues, the agent could not know that p is true in

@ without thereby knowing p, as well. In fact, however, this is highly dubious. Imagine an agent who, at a certain point, says or thinks to himself "this worldstate, the one I find myself in now (which I know to be such and such, and so and so) is the one that really obtains, or is instantiated." It certainly seems that such an agent demonstratively refers to the actual world-state @, with which he is acquainted, and truly says, of @, that it obtains, or is instantiated. If he is sincere, the agent should qualify as knowing, of @, that it is the actual world-state. Most likely, he and his fellow conversationalists have known this all along. But then, if identifying the actual world-state is coming to know, of that state, that it "is actual," then identifying the actual world-state cannot be the idealized goal of rational inquiry, or conversation.

This conclusion could be resisted, if it could be shown that the familiarity of ordinary agents with the way things are in the universe is inevitably too limited and fragmentary to provide them with the sort of acquaintance with @ needed to acquire *de re* knowledge of it at all. Although the idea that such knowledge is unattainable is not without force, it is also not easy to accept. For one thing, accepting it would render that staple of indexical semantics, the actuality operator, essentially useless. Since the proposition expressed by [Actually S] is a proposition which says, of the world-state  $C_w$  of the context, that p is true in  $C_w$  (where p is the proposition expressed by S in C), an inability to have *de re* knowledge of @ would prevent speakers from ever knowing the propositions expressed by utterances of sentences containing the actuality operator, thereby depriving such sentences of any normal use. Since these sentences do seem to have such a use, there is reason to believe that *de re* knowledge of @ is possible.

Such knowledge is also defensible on other grounds. Although de re knowledge of individuals normally requires some sort of contact with them, it does not require extensive or systematic knowledge of the totality of facts involving them. For example, even though my knowledge of my city, my country, my planet, my solar system, and my universe is an infinitesimal fraction of all there is to be known about these things, I am surely able to acquire some de re knowledge of them. If de re knowledge of states of individuals (including states of the universe) is similar in this respect to de re knowledge of individuals themselves, then it too is compatible with extreme limitations on the extent and systematicity of such knowledge. Thus, the limited and fragmentary nature of our knowledge of the actual world-state, @, presents no obvious bar to our having some de re knowledge of it. Finally, it should be noted that a proponent of Stalnaker's model of discourse is in no position to deny this. Since the model routinely attributes de re knowledge of world-states to speakers on the basis of a much slenderer acquaintance with those states than any of us have with @, the proponent of the model ought to accept the idea that conversational participants do have de re knowledge of @. Once this is accepted, there is, as I have argued, no plausible grounds for denying that we know of @ that it is actual, or instantiated.

Judith Jarvis Thomson chap11.tex V1 - May 30, 2006 10:53 A.M. Page 231

231

#### Understanding Assertion

This brings us back to the goal of rational inquiry and discourse presupposed by the model. We have seen that the goal cannot be that of identifying the actual world-state, in the sense of coming to know, of the actual world-state @, that it obtains, or is instantiated. What, then, should we take the goal to be? The answer that the proponent of the model ought give is, I think, that the goal is to "identify the actual world-state" in the sense of arriving at maximally complete, descriptive knowledge of the form, the state of the world that actually obtains, or is instantiated, is one in which p, q, r, ... where 'p, q, r, ...' are filled in with a comprehensive list of the facts of @. There are two things to notice about this answer. First, it is compatible with Stalnaker's discussion of the model-since approaching the goal involves learning, or coming to accept, more and more truths, which has the effect of shrinking the set in which the world-state that actually obtains is, and is assumed to be, located. Second, on this way of understanding the goal, world-states have no priority over propositions. The goal of identifying the actual world, in the sense in which we have now come to understand it, is simply that of learning as many (relevant) truths as we can. It is hard to quarrel with that.

#### A LESSON ABOUT THE NECESSARY APOSTERIORI

We are almost ready to tackle Stalnaker's attempt to use his discourse model to explain Kripkean examples of the necessary aposteriori. Before we do, however, it is worth pausing to tease out an important consequence of the model regarding how the necessary aposteriori should **not** be understood. The consequence involves sentences containing the actuality operator that are often taken to be paradigmatic instances of this category of truths. Although any sentence of this sort will do, we will focus on those constructed from contingently codesignative descriptions—*the x: x is F* and *the x: x is G*—that are rigidified using *actually.* This gives us two descriptions—*the x: actually x is F* and *the x: actually x is G*—which designate the same object o in every possible world-state in which o exists, and designate nothing in any world-state in which o doesn't exist. These are used to construct (1).

(1). If [the x: actually x is F] exists, then [the x: actually x is F] = [the x: actually x is G]

(1) is necessary, since it is true by falsity of antecedent in any world-state in which o doesn't exist, and true by truth of the consequent in any world-state in which o does exist.

Is (1) knowable aposteriori? Well, one might come to know it is by first coming to know the contingent truth

(2). If [the x: x is F] exists, then [the x: x is F] = [the x: is G]

and inferring (1) from (2). Since (2) can be known only aposteriori, anyone who comes to know (1) by this route knows it aposteriori. However, since all apriori

truths can also be known aposteriori, there is nothing significant about this. In order to show that (1) is a genuine instance of the necessary aposteriori, one must show that it cannot be known apriori, and so is knowable only aposteriori. However, if the lessons we have drawn from Stalnaker's model about our knowledge of possible world-states are correct, then this cannot be shown.

Consider a scenario in which we imagine a possible state of the world to ourselves, or perhaps a class of such states. We say to ourselves, *Let w be a possible world-state in which o is the unique thing which is F, and o is the unique thing which is G, and*.... *and*.... We go on enumerating the aspects of w for awhile, and then ask *Is w a world-state with respect to which* (2) *is true?* We answer that, of course, it is. On Stalnaker's model this counts as knowing of w that (2) is true with respect it—which is knowing that which is expressed by (3a).

(3a). In w: if [the x: x is F exists], then [the x: x is F] = [the x: x is G]

Since (3a) is apriori-equivalent to (3b) and (3c), knowing the former, on the basis of our apriori imagining, provides a sufficient basis for coming to know the latter in the same way.

(3b). If, in w, [the x: is F] exists, then, in w, [the x: x is F] = [the x: is G]

(3c). If [the x: in w, x is F] exists, then [the x: in w, x is F] = [the x: in w, x is G]

Now let's suppose something else, namely that the state of the world w we have been imagining is, unknown to us, its actual state. In other words, the state the universe actually is in has precisely the characteristics we were imagining, even though we didn't realize this at the time. If this is so, then in knowing (3a), and hence, (3c), apriori, we knew (4) apriori as well.<sup>7</sup>

(4). If [the x: in @, Fx] exists, then [the x: in @, x is F] = [the x: in @, x is G]

But then, since (1) expresses the very same thing as (4), it too is knowable apriori, and so is not an instance of the necessary aposteriori. A similar conclusion holds for every purported instance of the necessary aposteriori that makes essential use of the actuality operator. This is significant, since for a number of philosophers, particularly those who attempt to explain the necessary aposteriori by appeal to so-called two-dimensionalist semantics, such sentences have provided the template for understanding a necessary aposteriori truths.<sup>8</sup>

FN:7

FN:8

<sup>&</sup>lt;sup>7</sup> We may, of course, have been imagining a class of world-states satisfying our stipulations, of which @ is a member, rather than imagining @ by itself. However, if, as Stalnaker's model presupposes, knowing of this class that (2) is true with respect to its members counts as knowing of each member that (2) is true with respect to it, then the argument is not affected. Note, the argument does not depend on the model's problematic identification of propositions with functions from world-states to truth values.

<sup>&</sup>lt;sup>8</sup> See Davies and Humberstone (1980) and Soames (2005), plus the references cited there.

Judith Jarvis Thomson chap11.tex V1 - May 30, 2006 10:53 A.M. Page 233

#### Understanding Assertion

233

## STALNAKER'S ACCOUNT OF THE NECESSARY APOSTERIORI

We return to the three rules governing assertion in Stalnaker's model of discourse.

- R1. A proposition asserted should always be true in some but not all of the possible world-states in the context set.
- R2. Any assertive utterance should express a proposition, relative to each possible world-state in the context set, and that proposition should have a truth value in each possible world-state in the context set.
- R3. The same proposition should be expressed relative to each possible worldstate in the context set.

Having motivated these rules, Stalnaker uses them to explain assertive utterances of Kripkean examples of the necessary aposteriori.

As with the other principles, one may respond to apparent violations [of R3] in different ways. One could take an apparent violation as evidence that the speaker's context set was smaller than it was thought to be, and eliminate possible worlds relative to which the utterance receives a divergent interpretation. Or, one could reinterpret the utterance so that it expresses the same proposition in each possible world. Consider an example: hearing a woman talking in the next room, I tell you, That is either Zsa Zsa Gabor or Elizabeth Anscombe. Assuming that both demonstrative pronouns and proper names are rigid designators-terms that refer to the same individual in all possible worlds-this sentence comes out expressing either a necessary truth or a necessary falsehood, depending on whether it is one of the two mentioned women or someone else who is in the next room. Let i be the world in which it is Miss Gabor, j the world in which it is Professor Anscombe, and k a world in which it is someone else, say Tricia Nixon Cox. Now if we try to bring the initial context set into conformity with the third principle [R3] by shrinking it, say by throwing out world k, we will bring it into conflict with the first principle [R1] by making the assertion trivial. But if we look at what is actually going on in the example, if we ask what possible states of affairs the speaker would be trying to exclude from the context set if he made that statement, we can work backward to the proposition expressed. A moment's reflection shows that what the speaker is saying is that the actual world is either i or j, and not k. What he means to communicate is that the diagonal proposition of the matrix E exhibited below, the proposition expressed by  $\Lambda E$ , is true.<sup>9</sup>

E	i	j	k	↑E	i	j	k
i	Т	Ť	Т	i	Т	Ť	F
j	Т	Т	Т	j	Т	Т	F
k	F	F	F	k	Т	Т	F

<sup>9</sup> Stalnaker (1999, 91).

In this example, the propositional concept E associated with the sentence S uttered by the speaker tells us two things: (i) we don't know which proposition is (semantically) expressed by S in the actual context, because which proposition is expressed depends on which world-state actually obtains, and we don't know which state does obtain; (ii) none of the possible propositions expressed would serve any useful purpose in the conversation. To assert a necessary truth is to assert something which is of no use in narrowing down the location of the actual world-state within the context set; and asserting a necessary falsehood is even worse. Thus, E violates R3, and any attempt to avoid this violation by excluding one or more of the world-states will violate R1. So, if we are to avoid violation entirely, and to regard the speaker's utterance as useful and informative, we must take it as asserting some proposition other than the proposition it (semantically) expresses at i, j, or k. Which proposition?

Since whatever the actual world-state turns out to be, the speaker will be committed to the utterance of S expressing a truth in the context, that is what we should take to be asserted. The proposition asserted is the proposition that is true (false) at a world-state w (of the context set) just in case the proposition (semantically) expressed by S in w is true (false) at w—it is the assignment of truth values that arises from E by looking along the diagonal and selecting the truth value that appears in row w of column w, for each w. Stalnaker calls this the *diagonal proposition*. Since, in this example, the diagonal proposition is neither true in all world-states of the context set nor false in all those states, it can do the job that asserted proposition that is really asserted by the speaker's utterance—no matter which member of the context set turns out actually to obtain. This is what  $\uparrow E$ represents, where ' $\uparrow$ ' (pronounced DAGGER) is an operator that maps a propositional concept C1 onto the propositional concept C2 that arises from C1 by taking each of the rows of C2 to be the diagonal proposition determined by C1.

This is the prototype for Stalnaker's treatment of the necessary aposteriori, which—extrapolating and generalizing his explicit remarks—we may take as suggesting T1.

T1. Although no necessary propositions are knowable only aposteriori, a sentence S, as used in a particular conversation C, is an example of the necessary aposteriori iff the proposition (semantically) expressed by S at the world-state that really obtains in the speaker's context is necessary, but the diagonal proposition asserted by a use of S in C is contingent, and hence knowable only aposteriori.

Given T1, plus Stalnaker's discussion of the example—*That is either Zsa Zsa Gabor or Elizabeth Anscombe*—motivating it, one might get the mistaken impression that he thought that all genuine examples of the necessary aposteriori are indexical, in the sense of semantically expressing different propositions in

#### Understanding Assertion

different contexts of utterance. However, he didn't believe this. How, then, were instances of the necessary aposteriori involving names and natural kind terms to be treated? He addresses this point in the following passage.

I suggest that a common way of bringing utterances into conformity with the third principle [R3] is to interpret them to express the diagonal proposition, or to perform on them the operation represented by the two-dimensional operator DAGGER. There are lots of examples. Consider: *Hesperus is identical with Phosphorus, it is now three o'clock, an ophthalmologist is an eye doctor.* In each case, to construct a context which conforms to the first principle [R1], a context in which the proposition expressed is neither trivial nor assumed false, one must include possible worlds in which the sentence, interpreted in the standard way, expresses different propositions. But in any plausible context in which one of these sentences might reasonably be used, it is clear that the diagonal proposition is the one that the speaker means to communicate. The two-dimensional operator DAGGER may represent a common operation used to interpret, or reinterpret, assertions and other speech acts so as to bring them into conformity with the third principle [R3] constraining acts of assertion.<sup>10</sup>

Let us focus on (5a) and (5b).

- 5a. Hesperus is identical with Phosphorus.
- 5b. An ophthalmologist is an eye doctor.

Since (5a,b) don't contain indexicals, their meanings, i.e. their Kaplan-style characters, will be constant functions. Each expresses the same (necessary) proposition in every context of utterance. If the propositional concepts associated with them in these conversations were simply their meanings, then the application of the dagger operation would have no effect, and Stalnaker's explanation of their informative use wouldn't get off the ground. Thus, in these cases, he must not have been taking the needed propositional concepts to be the meanings (characters) of the sentences uttered.<sup>11</sup>

Instead, it is natural to interpret him as taking the propositional concept associated with an utterance of S in a conversation to be that which speaker-hearers (jointly) believe the meaning of S to be. In cases in which they know all the relevant semantic facts, this will simply be the meaning of S. In cases in which they are ignorant of, or confused about, some of these facts, the propositional concept associated with S may be something less than the actual meaning of S. For example, in the case of (5a), the propositional concept may by given by the formula *x* is identical with *y*—with different possibilities regarding the constant

<sup>10</sup> Stalnaker (1999, 92).•

FN:10

FN:11

• Q1

<sup>&</sup>lt;sup>11</sup> Unlike later two-dimensionalists, Stalnaker never subscribed to the general thesis that names and natural kind terms are indexical, rigidified descriptions. For discussion, see the introduction to Stalnaker (1999), esp. pp. 14–19, and also Stalnaker (2001), esp. pp. 199–200 (of Stalnaker 2003).

functions from world-states to objects which are candidates for the meanings of the names to be substituted for the 'x' and 'y' being reflected in different worldstates of the context set. The case of (5b) is similar, except that the different possibilities for filling in the content of 'O' in the relevant formula—An O is an eye *doctor*—are meanings of general terms, rather than meanings of proper names. On this interpretation, the context set for an utterance of (5a) will contain some world-states in which one or both of the names Hesperus and Phosphorus stand for something other than what they both actually stand for, and the context set for an utterance of (5b) will contain some world-states in which ophthalmologist means something other than what it actually means. Presumably, the justification for this way of looking at things is the idea that (5a) and (5b) will be used only if (some) conversational participants are ignorant about what these words actually mean, or stand for, with the result that world-states in which the words mean, or stand for, something different from what they actually mean, or stand for, will be among the genuine possibilities left open by the conversation prior to the utterances. But then, the thought continues, different propositions will be expressed when the sentences are "interpreted in the standard way," at these world-states, considered as contexts. This, I think, is how Stalnaker intended to generalize his explanation beyond genuinely indexical sentences.<sup>12</sup>

At this point, however, we run into a problem. Although there may be some sentences and conversations that fit the picture, some do not. For example, it is not true that (5a) would be used only in a conversation in which (some) conversational participants are ignorant of what Hesperus and Phosphorus stand for, in the sense most relevant to Stalnaker's model. Each participant may know perfectly well that 'Hesperus' refers to this object [pointing in the evening to Venus] and that 'Phosphorus' refers to that object [pointing in the morning to Venus]. They may even have done the pointing themselves. Clearly, such speakers know of the referent of each name that it is the referent of that name. Hence the (contingent) propositions expressed by 'Hesperus' refers to x and 'Phosphorus' refers to x relative to an assignment of Venus to 'x' should be among those that have already been assumed or established in the conversation. But then, metaphysically possible world-states in which the names mean and refer to different things will already have been eliminated from the context set as incompatible with what has been assumed or established. Since (5a) can, nevertheless, be used in these circumstances perfectly intelligibly, Stalnaker's explanation cannot successfully be applied to this case. This is an instance of the general problem noted earlier. When de re attitudes are involved, speakers cannot always determine the compatibility relations presupposed by the model.

236

<sup>&</sup>lt;sup>12</sup> On this interpretation, propositional concepts map each world-state w in the context set onto the proposition that speaker-hearers believe would be semantically expressed if w were to obtain. When they know all relevant semantic facts, these are the propositions that really would be expressed if w obtained.

Judith Jarvis Thomson chap11.tex V1 - May 30, 2006 10:53 A.M. Page 237

237

#### Understanding Assertion

#### FAILURE OF THE MODEL

A related problem is posed by a different example. Imagine you are sitting across from me in my office, you point to a paperweight in plain view on my desk, and ask What is that paperweight made of?, and I respond It is made of wood. Although you don't know, prior to my utterance, what the paperweight is made of, we both assume that, whatever it is made of, it is an essential property of that paperweight that it be made of that stuff. Since, in fact, the paperweight is made out of wood, my remark is an example of the necessary aposteriori. How would this conversation be represented in Stalnaker's model of discourse? Prior to the utterance there would be different possible world-states in the context set that were compatible with everything assumed or established in the conversation up to that point. We may take these to include a context/world-state i in which the thing that, in i, is the one and only one paperweight on the desk is made of wood, a context/worldstate j in which the paperweight on the desk in j is made of something else, e.g. plastic, and a context/world-state k in which a paper weight in front of us in k is made out of something else again-say, metal. In short, in Stalnaker's model, the propositional concept PW would be associated with my utterance.

PW	i	j	k
i	Т	Ť	Т
j	F	F	F
k	F	F	F

The rules R1–R3 for assertion would then yield two conclusions: (i) that on hearing my utterance you had no way of knowing which proposition was (semantically) expressed by my sentence, because which proposition was expressed depended on which world-state—i, j, or k—actually obtained, and you didn't know, in advance of accepting my remark, which world-state did obtain; and (ii) that none of the propositions that might have been expressed would have served a useful purpose. To have asserted a necessary truth would have been to have asserted something uninformative, and of no use in narrowing down location of the actual world-state in the context set; and to have asserted a necessary falsehood would have been a nonstarter. So, if you were to regard my utterance as successful, you had to take it as asserting some proposition other than any of the candidates for being the one it (semantically) expressed.

Which proposition might that have been? Since you knew that whatever the world-state of the context turned out to be, I would be committed to my remark being true, the proposition you must have taken me to have asserted is a proposition that is true (false) at a world-state of the context set iff the proposition expressed by my sentence at that world-state is true (false) at that world-state. This is the diagonal proposition associated with PW. Since it is neither true at

all world-states in the context set, nor false at them all, asserting it does the job that assertions are intended to do. Implicitly recognizing this, we both rightly understood the diagonal proposition to be the proposition I asserted.

That is the explanation provided by Stalnaker's model. There are two things wrong with it. First, it is wrong to suppose that you had any relevant doubt about what proposition was (semantically) expressed by my utterance of It is made of wood in response to your question, What [pointing at the paperweight] is that made of? The proposition I expressed is one that predicates being made of wood of that very paperweight-the one we both were looking at, and saw clearly sitting on the edge of my desk. You knew that it was the object you had asked about, and about which I had given an answer. Since you also knew what wood is, you knew precisely which property was predicated of which object by my remark. Surely, then, you did know the proposition my sentence expressed. In short, there was a proposition p such that you and I both knew that my utterance expressed p, even though you didn't know, in advance of accepting my remark, whether or not p was true, and so didn't know whether or not p was necessary. Of course, given his identification of propositions with functions from metaphysically possible world-states to truth values, Stalnaker can't say this, since the fact that p is necessary would require him to say (i) that you knew p all along, and (ii) that you knew that my utterance expressed a trivial truth, simply by virtue of understanding it. Since this is absurd, he is forced to the patently counterintuitive conclusion that upon hearing my utterance, you didn't know that it expressed p (where p is the proposition it actually did express).

The second thing wrong with Stalnaker's explanation is that the world-states j and k in the context set must either be (a) ones that are not really metaphysically possible, or (b) ones that are not compatible with all the shared assumptions of the conversational participants prior to my utterance—both of which are contrary to the dictates of the model. What are the world-states i, j, and k? They are total possibilities regarding how the world might be in which one and only one paperweight is sitting on my desk, seen by us, and the subject of our discourse. The paperweight satisfying these conditions in i is made of wood, whereas the paperweights satisfying them in j and k are made of plastic in one case and metal in the other. What paperweights satisfy these conditions in j and k? If j and k are really metaphysically possible, as Stalnaker insists, then the paperweights in j and k can't be the paperweight that is really on my desk. Since that paperweight is made of wood in every genuinely possible world-state in which it exists, it is not made of plastic in j or metal in k. It follows that j and k must be worldstates in which some other paperweight is between us on the desk, seen by us, and the subject of our conversation. But how can that be? Surely, one thing that was part of the shared conversational background prior to my remark was the knowledge that this very paperweight [imagine me demonstrating it again] was between us on the desk, seen by us both, and the subject of our conversation. To deny this would be tantamount to denying that we ever know, of anything

#### Understanding Assertion

we perceive or talk about, that it has one property or another. Even if we put the question of knowledge aside, surely we both believed these things about this very paperweight, which is all the model requires. But if we did have this *de re* knowledge, or these *de re* beliefs, then the discourse model's requirement that the world-states in the context set be compatible with everything assumed and established in the conversation must have eliminated all metaphysically possible world-states in which other paperweights, not made out of wood, were the one and only paperweight under discussion. But then, there is no room for the diagonalization required by Stalnaker's explanation.

This is the fundamental problem. Unless some persuasive defense can be found for excluding obvious, shared *de re* belief and knowledge from the conversational model, Stalnaker's explanation cannot succeed. I will argue that no such defense can be given. First, however, I will improve the model by liberalizing the notion of possibility it employs. Having strengthened the model so that it can accommodate cases involving essential properties of objects, like my paperweight, I will return to the problems posed by *de re* knowledge and belief, and investigate why they are intractable.

#### IMPROVEMENTS AND PROBLEMS

The model can be improved by dropping Stalnaker's antecedent philosophical commitment to restricting epistemic possibility to metaphysical possibility. To drop this commitment is to recognize world-states that are metaphysically impossible but epistemically possible, i.e., maximally complete properties that the universe couldn't really have had, but which we cannot know apriori that it doesn't have (on analogy with properties that ordinary objects couldn't have had, but which we cannot know apriori that they don't have). When we allow context sets to include such world-states, the propositional concept associated with my utterance about the paperweight turns out to be different from the one we earlier took it to be. On this way of looking at things, i, j, and k are different epistemic possibilities involving the very same object, o—where o is the paperweight that we actually see on my desk, are talking about, and know that we are talking about. In world-state i, o is made of wood; in j, o is made of plastic; and in k, o is made of metal. The resulting matrix is PW\*.

PW*	i	j	k
i	Т	F	F
j	Т	F	F
k	Т	F	F

Since the same proposition is expressed with respect to each epistemologically possible world-state, and since it is neither trivially true nor trivially false, no diagonalization is needed.

This improvement encourages a certain thought. Perhaps Stalnaker's model of inquiry can be divorced from the philosophically contentious motivations that partially inspired it. The idea is to give up the identification of epistemic possibility with metaphysical possibility, to give up the goal of explaining away the necessary aposteriori, and to give up the analysis of propositions as functions from metaphysically possible world-states to truth values. We retain the idea that utterances are associated with propositional concepts or matrices, plus the general model of discourse that makes use of these matrices. We also retain the idea that the point of a discourse is to narrow the set of the possibilities—now thought of as including both epistemic and metaphysical possibilities—within which the actual world-state is presumed to be located. As before, an assertion is supposed to shrink the set of possibilities compatible with everything that has previously been assumed or established in the conversation. On this new picture, the conversational rules R1-R3 remain intact.

Scott Soames

The model can be illustrated using the following example. I say, He is John Hawthorne (demonstrating a man sitting at the end of the table) in a conversation in which it is common knowledge that this man—the one I am talking about-is either John Hawthorne or Ted Sider. The utterance takes place in a context in which everyone knows a few facts about John and Ted already, but not everyone knows what they look like. Perhaps everyone has talked to each of them on the phone, or read the work of each, or corresponded with each, or some combination of the three, even though many would not recognize John or Ted by sight. Let us stipulate that everyone already knows of John that his name is 'John', that he is a Rutgers professor, and that he is not Ted—similarly for everyone's antecedent knowledge of Ted. Moreover, this shared knowledge is known to be shared, and so the propositions known are part of the presupposed conversational background. In this situation I utter the sentence, He is John Hawthorne, demonstrating John, who is sitting at the end of the table. The sentence uttered contains a name, which, like the demonstrative *he*, is a rigid designator with respect to all possible world-states, epistemic and metaphysical alike.

What are the epistemic possibilities prior to my utterance? It might seem that the two most obvious possibilities—j and t—could be described as follows: in j there is a unique person sitting at the end of the table and that person is John, and in t there is a unique person sitting there and that person is Ted. This gives us the following matrix.

	j	t
j	Т	Т
t	F	F

R1-R3 dictate that we perform the diagonalization operation, which gives us an asserted proposition that is true just in case John is sitting at the end of the table, and false otherwise. That is a good result, since it, or something quite like

#### Understanding Assertion

it, would normally be regarded as having been asserted by such an utterance. If you were to report my remark by saying *Scott said that John Hawthorne was sitting there* (gesturing to the place at the end of the table), I think most people would judge what you said to be true.

Nevertheless, the way we reached this result is problematic. World-states j and t are supposed to be epistemic possibilities compatible with everything taken for granted in the conversation prior to my remark. But, as in the earlier example about the paperweight, I left out of the specifications of j and t certain things known by all conversational participants. I ignored the fact that it was known (prior to my remark) that he [imagine me pointing again at John] was sitting there and also the fact that it was known (prior to my utterance) that since there weren't two people sitting there, and since John and Ted are different people, if John was sitting there, then Ted wasn't. When these things are added to the conversational background, t becomes incompatible with what is known or assumed by conversational participants, and so is excluded from the context set.

Why? First, since it is known (prior to my utterance) that he [pointing at John] is sitting there, it follows that he, John, is an x, such that it is known that x is sitting there. This is just to say that the singular proposition p which says of John that he is sitting there is known to be true by the conversational participants, and so must be true with respect to t, if t is to be compatible with everything commonly known or assumed. Second, since it is known (prior to my remark) that if John is sitting there, Ted isn't, it again follows that John is an x such that it is known that if x is sitting there then Ted isn't. But it has already been stipulated that the proposition q that Ted is sitting there is true in t. Hence, t can be compatible with everything which is known or assumed in the conditional proposition the antecedent of which is p and the consequent of which is the negation of q—is consistent. Since this trio is inconsistent, t must be excluded from the context set, in which case our revised, Stalnaker-style explanation of what I asserted fails in a way similar to the way the original explanation of my assertion about the paperweight failed.

How, then, is it that my utterance of *He is John Hawthorne* was informative? Since I discuss this sort of issue in considerable detail in chapters 3 and 4 of Soames (2002), I will deal with it only briefly here. We know that prior to the utterance my audience already believed of John that he was John.<sup>13</sup> So the new belief acquired by virtue of accepting my utterance wasn't that one. What might it have been? One such belief was surely that he, the person sitting there, was John Hawthorne. Everyone in the audience could see—without any appeal to propositional concepts or diagonalization—that I was attributing the property of being John Hawthorne to the guy sitting there, at whom I was pointing. So naturally

 $<sup>^{13}</sup>$  The account initially given in Soames (2002) of the relationship between the semantic content of a sentence S in a context C and what is asserted by uttering S in C is modified and extended in Soames (2004).

I was committed to that being true. Moreover, if someone in the audience were to describe what I said to a third party who hadn't been present, he might say, *At first several of us didn't know who was sitting at the end of the table, but then Scott said that John Hawthorne was the one sitting there.* In ordinary life, such a report would be taken to be completely correct. If it is correct, then not only did I convey this informative proposition, I actually said (i.e., asserted) it. This is evidence that what I asserted went a little beyond the strict semantic content of the sentence I uttered in the context. In this respect, I agree with Stalnaker; in cases like this the speaker does assert a proposition which is not the proposition semantically expressed by the sentence he utters. But the mechanism by which this occurs is a rather ordinary one, and typically doesn't involve any forced twodimensionalist diagonalization.

#### THE NATURE OF THE PROBLEM

If what I have just argued is correct, then Stalnaker's elegant model of discourse can be improved, but not saved, by liberalizing it to allow for epistemically possible world-states, over and above those that are metaphysically possible. The fundamental, mistaken assumption embedded in the model leading to its failure is that conversational participants can do two things: (i) identify, at the time of each utterance, precisely which possible world-states are compatible with everything previously assumed or established in the conversation; and (ii) determine which of these possible states are compatible with propositions expressed by the sentence we utter under different assumptions about which possible world-state actually obtains. In reality, we can't always do these things, no matter whether the possible world-states in question are metaphysical or epistemic. We can't do them because the relationship between sentences and the propositions they express is nontransparent in an important way. There are pairs of sentences S1 and S2, and contexts C, such that in C

- (a) S1 expresses a proposition p1, S2 expresses p2, and speaker-hearers understand both sentences, while knowing that to accept S1 is to believe p1 and to accept S2 is to believe p2,
- (b) p1 bears some intimate "logical" relation to p2, e.g. p1 is the negation of p2, or p1 is identical with p2, or p1 is a conditional and p2 is its antecedent,

even though

(c) speaker-hearers have no way of knowing that the relation mentioned in(b) holds between the proposition believed in virtue of accepting S1 and the proposition believed in virtue of accepting S2.

Because of this, there are cases in which speaker-hearers believe p1, and yet are in no position to recognize that in believing p2 they are believing something

#### Understanding Assertion

inconsistent with this, which, in terms of the model, rules out all epistemically possible world-states. In other cases, in which p1 and p2 are consistent, but some different relation holds between them, the fact that speaker-hearers believe both p1 and p2 may rule out some but not all possible world-states, without their being able to recognize which. Because of this non-transparency in the relationship between sentences, the propositions we believe (assert) in virtue of accepting (uttering) them, and the world-states in which these propositions are true, our beliefs and assertions cannot always interact with one another in the way the model presupposes. Because of this, the model fails.

The assumptions that lead to this result are modest. In order to reach our conclusion, one may, but need not, endorse the contentious, but I believe correct, doctrine that the semantic contents of names and indexicals (relative to contexts) are their referents, or the similarly contentious, but correct, doctrine that the semantic contents of natural kind terms are the kinds they designate. One reason these semantic assumptions are not needed is that we can generate corresponding problems for the model using pairs of synonymous expressions of other sorts—for example catsup/ketchup and dwelling/abode—where in each case a speaker can understand both expressions without realizing that they are synonymous.<sup>14</sup> Another reason that contentious semantic assumptions are not necessary is that what generates problems for the model are not so much semantic facts about the sentences involved, as cognitive facts about speakers who use them. When an agent looks directly at the paperweight on my desk, and sincerely utters That [pointing at the paperweight] is the paperweight I am talking about, he is correctly described as believing, of the paperweight, that he is talking about it, where the proposition believed is also expressed by x is the paperweight I am talking about, relative to an assignment of the object itself to 'x'. It is believing this proposition that creates trouble for the model, whether or not we identify it with the semantic content of the sentence uttered. Similar points hold for examples in which the sentence uttered contains a proper name or natural kind term. In all these cases, the propositions that prove problematic for the model are among those that conversational participants come to believe and assume at later stages of the conversation. Since these assumptions determine the context set for later utterances, the only hope of saving the model is to exclude beliefs of this sort from playing this role. However, there seems to be no reasonable way of doing this.

#### THE UBIQUITY OF THE DE RE

The fundamental reason that *de re* belief can't be excluded from the model is that the model itself presupposes such belief. As I have stressed, it presupposes *de re* knowledge (or belief) of world-states, which, we have seen, is founded in

<sup>14</sup> See Salmon, (1990) and Rieber (1992).

*de re* knowledge (or belief) of individuals or kinds. This knowledge (or belief) of world-states is of three sorts:

- (i) knowledge (belief) of world-states that they are, or are not, compatible with propositions previously assumed or established in the conversation,
- (ii) knowledge (belief) of world-states w in the context set, sentences S, and propositions p, that if w obtains then S expresses p, and
- (iii) knowledge of the truth value of p in w, for each w in the context set.

In each of these cases, the *de re* knowledge (belief) of world-states required by the model is inextricably linked to ordinary *de re* knowledge (belief) of individuals (or kinds). Hence there is no excluding the latter.

Regarding (i), the propositions previously assumed or established in the conversation will standardly include singular propositions-knowledge of which amounts to de re knowledge of individuals (or kinds)-about the speaker and other conversational participants, the salient items in the context of utterance, and the various words in use plus their meanings. In many cases, the propositions previously assumed or established will also include singular propositions about the individuals (or kinds) which are topics of the conversation. It is not unusual for these singular propositions to be more readily available as commonly held assumptions of conversational participants than many of their purely descriptive counterparts. For example, there surely are cases in which conversational participants discussing an individual i each knows of i that they all know that i is the individual being talked about, and each know that it has been assumed or established that i has one or another property P—even though the descriptive information about i possessed by conversational participants varies so much from one participant to next that there may be few, if any, (purely qualitative) descriptions D that uniquely identify i which are known by each participant to be associated by all of them with any of the terms used in the conversation. In such cases, the descriptive differences between the parties will wash away, and the most salient proposition about i known to be commonly assumed or established in the conversation may well be a singular proposition that predicates P of i. For reasons like these, there seems to be no way for a viable model of discourse to exclude singular propositions from the set of propositions commonly assumed or established in a conversation. The consequence of this for the de re knowledge of world-states of type (i) presupposed by Stalnaker's discourse model is easy to see. If the propositions assumed or established include a proposition that says of o that it "is F", then to know of an arbitrary world-state w whether it is compatible with what has been assumed or established in the conversation (and hence to know whether w is in the context set), one must know of w whether it is a world-state with respect to which o "is F". It is not enough to know that w is a world-state in which whatever object satisfies a certain description (in the world-state in which the conversation actually takes place) "is F"; one must know of o itself that it "is F", or that it "is not F", in w.

Judith Jarvis Thomson chap11.tex V1 - May 30, 2006 10:53 A.M. Page 245

245

#### Understanding Assertion

A similar point holds for the combination of (ii) and (iii), which generates the propositional concepts, or matrices, on which Stalnaker's model is based. To generate these matrices, speaker-hearers must know, of each sentence S and pair of world-states w1 and w2 in the context set, whether the proposition that an utterance of S would express, were w1 to obtain, would be a true, or a false, description of w2. With this in mind, suppose, as Stalnaker does in several of his own examples, that S contains a rigid designator  $\alpha$  which rigidly designates o1 if w1 obtains, and o2 if w2 obtains. Suppose further that S says of whatever is designated by  $\alpha$  that it "is F". Then, in order for the background knowledge and beliefs of speaker-hearers to generate the propositional concept employed by Stalnaker, speaker-hearers must know of both o1 and o2 whether they "are F" in w1 and w2. This is ordinary *de re* knowledge and belief of those objects. Since the discourse model presupposes knowledge and belief of this sort, it cannot relegate it to the sidelines.

#### LESSONS

For this reason, Stalnaker's model of assertion fails. However, it is important not to overreact. Although certain aspects of the model, and the uses to which it was put, must be abandoned, other features of it can be retained. Among the former are Stalnaker's revisionary account of the necessary aposteriori, his restriction of possible world-states to the metaphysically possible, and his identification of propositions with functions from such states to truth values. Among the latter are versions of his rules for assertion, R1–R3, suitably reformulated to avoid the model's problematic features.

The correct account of the necessary aposteriori is illustrated by the homely example of the paperweight on my desk, which I show you. You see it, pick it up and feel it, but can't tell what it is made of. You imagine that it might be made of plastic, or metal, or wood. In imagining this, you are imagining the very object itself having the property of being made of plastic, being made of metal, or being made of wood. In so doing, you are also imagining different properties the universe might have-the property of containing this very object being made of plastic, the property of containing this very object being made of metal, and so on. You don't know which, if any, of these properties the universe really does have. Since you can't find this out by apriori reasoning alone, these properties, or more complete versions of them, are conceivable ways the world might be that are epistemically possible. When you finally learn that the paperweight is, in fact, made out of wood, you realize that it couldn't have existed without being made out of wood, and so you realize that certain epistemically possible ways the world might be are not ways that it could genuinely have been, and so are metaphysically impossible. Since these ways are just world-states, this elementary Kripke-style example of the necessary aposteriori shows that certain epistemically possible world-states are metaphysically impossible.

#### Scott Soames

This example relies on a potentially contentious metaphysical doctrine-the essentiality of constitution. However, there is nothing special about the particular essential property chosen. Other essential properties or relations (e.g. the relation of non-identity) would serve equally well. The important thing is simply that there be such properties (and relations). Given that there are, we may reason that just as there are properties (relations) that ordinary objects could possibly have had (stood in) and other properties (relations) they couldn't possibly have had (or stood in), so there are certain maximally complete properties that the universe could have had-metaphysically possible states of the world-and other maximally complete properties that the universe could not have had — metaphysically impossible states of the world. Just as some of the properties that ordinary objects couldn't have had are properties that one can coherently conceive them as having, and that one cannot know apriori that they don't have, so some maximally complete properties that the universe could not have had are properties that one can coherently conceive it as having, and that one cannot know apriori that it doesn't have. Given this, one can explain the informativeness of certain necessary truths as resulting (in part) from the fact that learning them allows one to rule out certain impossible, but nevertheless coherently conceivable, states of the world. Moreover, one can explain the function played by empirical evidence in providing the justification needed for knowledge of necessary aposteriori propositions. Empirical evidence is required to rule out certain metaphysically impossible world-states which cannot be known apriori not to be instantiated, with respect to which these propositions are false. Thus, by expanding the range of epistemically possible states of the world to include some that are metaphysically impossible, one can accommodate Kripkean examples of the necessary aposteriori. This-rather than any Stalnaker-style diagonalization-is the correct account of the necessary aposteriori.15

Finally, there is the matter of salvaging what we can from Stalnaker's rules governing conversation and inquiry. Stalnaker's R1 is equivalent to  $R1_{NC}$ .

R1<sub>NC</sub>. A proposition asserted should never be a necessary consequence of, or necessarily inconsistent with, the set of propositions already assumed or established in the conversation. (A proposition p is a necessary consequence of a set S of propositions iff there is no metaphysically possible world-state w in which the members of S are all true, while p is not; p is necessarily inconsistent with S iff there is no metaphysically possible world-state in which p and the members of S are jointly true.)

Since necessary consequences that are not apriori consequences are not, in general, discernable to agents,  $R1_{NC}$  won't do as a conversational maxim. Although  $R1_{AC}$  is an improvement over  $R1_{NC}$ , it won't quite do, either.

<sup>15</sup> For more on the distinction between metaphysically and epistemically possible world-states, and its role in explaining the necessary aposteriori, see Soames (2005, 198–209).

#### Understanding Assertion

R1<sub>AC</sub>. A proposition asserted should never be an apriori consequence of, or apriori-inconsistent with, the set of propositions already assumed or established in the conversation. (A proposition p is an apriori consequence of a set S of propositions iff there is no epistemically possible world-state w in which the members of S are all true, while p is not; p is apriori-inconsistent with S iff there is no epistemically possible world-state in which p and the members of S are jointly true.)

An important problem with  $R1_{AC}$  is that some apriori consequences are highly unobvious, requiring intricate and arduous apriori reasoning to reach. For this reason, the assertion of such a consequence of propositions already assumed or established may be highly informative, and effective in furthering the purposes of the conversation. Thus, it makes sense to replace  $R1_{AC}$  with  $R1_{OAC}$ .

R1<sub>OAC</sub>. A proposition asserted should never be an obvious apriori consequence of, or obviously apriori-inconsistent with, the set of propositions already assumed or established in the conversation.

At this point, however, the effects of the nontransparency of belief (and assertion) must be faced. Imagine a case in which conversational participants accept a pair of utterances—*This A is B* and *That A is C* (separated by a brief span of time)—accompanied by a pair of demonstrations, each clearly indicating a given object o in full view of each of the parties, without it being recognized that the same object is demonstrated on both occasions. In such a case, the propositions assumed or established in the conversation shortly after the remarks have been accepted will include a singular proposition p that says of o that it "is both A and B", and a similar proposition q that says of o that is "is both A and C". With p and q in the conversational background, it would be a violation of the intent of R1<sub>OAC</sub> for any of the conversational participants to assert that nothing "is both A and B", or that nothing "is both A and C", since to do so would be to assert something obviously inconsistent with what has already been established. Similarly, it would not do to assert-without further ado-that something "is both A and B", or that something "is both A and C", since these are obvious apriori consequence of p, and q, respectively. However, it might be very informative, and not in the least inappropriate, for someone to assert of o that it "is A, B, and C", even though the proposition thereby asserted is equivalent to the conjunction of p and q, or to assert that something "is A, B, and C", even though the truth of that proposition is a trivial consequence of the conjunction of p and q. Similar remarks hold for certain assertions obviously inconsistent with the conjunction of p and q-assertions of o that it "is not A, B, and C", or that nothing "is A, B and C". Although such assertions would render the set of propositions accepted by conversational participants inconsistent, in cases in which such inconsistency is undetectable, there is no culpable violation of conversational rules.

These considerations suggest reformulation  $R1_{OAC}$  so as not to focus exclusively on the proposition asserted, the propositions assumed or established, and the logical relations between them. When singular propositions are involved, the ways in which propositions are presented or entertained are as important as which propositions are asserted, believed, or accepted. This can be accommodated by reformulating  $R1_{OAC}$  along the following lines.

R<sub>O</sub>. An assertive utterance U should allow the conversational participants to correctly identify the proposition asserted, but U should never be such that it is obvious to conversational participants that the proposition asserted by U is a consequence of, or is inconsistent with, the propositions already assumed or established in the conversation.

Since a violation of  $R_O$  will occur only when the speaker's utterance presents the proposition asserted in a way that allows conversational participants to see that it is an obvious consequence of, or obviously inconsistent with, the propositions already assumed or established in the conversation,  $R_O$  accommodates the non-transparency of belief and assertion that caused problems for  $R1_{OAC}$ . Thus,  $R_O$ , or something like it, offers the best hope we have for salvaging what was correct about Stalnaker's rule R1.

In requiring that conversational participants be able to identify the proposition asserted,  $R_O$  also incorporates what was correct about R3. The insight embedded in the latter was, essentially, that one should be able to recognize what is asserted without settling the open questions remaining in the conversation. The reason  $R_O$  is able to accommodate this point is that the sense in which it requires the proposition asserted to be identified, or recognized, is quite weak. What is required is simply that for some p, the conversational participants be able to correctly recognize that U is an assertion of p. It is not required that U present p to conversational participants in a way that allows them to determine the truth value of p in all genuine metaphysically possible world-states compatible with propositions already assumed or established, or even to accurately assess whether the truth of all those propositions would guarantee the truth, or the falsity, of p. It is only because Stalnaker built these unreasonable requirements of cumulative, global transparency into the discourse model that R3 was made to seem to incorporate a truth more far-reaching than it really does.

Similar remarks apply to R2. Is it important that U express a proposition, and that speaker-hearers be able to recognize that it does, without answering all the questions still open in the conversation? Of course it is, but this is already implicitly incorporated into  $R_0$ . No reference to the different world-states in the context set is needed, since when the nontransparency of knowledge, belief and assertion is accommodated, and propositions are no longer identified with functions from world-states to truth values, these states fall away and the rules of discourse can best be stated directly in terms of structured propositions and assertive utterances. In many cases, there may be no set of world-states

#### Understanding Assertion

compatible with everything assumed or established, because that which has been assumed or established contains nontransparent inconsistencies. Even when there are no such inconsistencies, compatibility relations may be obscured by pockets of nontransparency among the propositions already accepted. To treat the set of world-states that are compatible with all these propositions as if it were a central component of the discourse model guiding the computations of speaker-hearers is to assume a global and cumulative transparency in our beliefs, assertions, and knowledge that is simply not a part of our cognitive or conversational lives.<sup>16</sup>

If I am right about all this, then Stalnaker's model of discourse must be drastically modified. Its central idea—that what is asserted in literal, nonmetaphoric speech often differs substantially from the semantic content of the sentence assertively uttered—is both true and important. However, the basic structure of the model—involving metaphysically possible world-states, propositions as functions from such states to truth values, propositional concepts, and speakerhearer calculations involving these items—cannot accommodate many of the facts about language use for which any acceptable theory of discourse must be responsible. I have indicated that an important core of truth can be salvaged from Stalnaker's rules, R1–R3, governing assertion. However, this is only the beginning. There is much more to be said about the ways in which the propositions commonly assumed in the conversational background, together with the meaning and semantic content of the sentence uttered, contribute to the proposition, or propositions, asserted by the utterance.

One important factor which I have not been able to talk about here differs sharply from anything in the Stalnaker model. This is the phenomenon of routine pragmatic enrichment of the semantic content of the sentence uttered, explored in Soames (2002 and 2004). In the Stalnaker model, the assertion of a proposition other than the one semantically expressed by the sentence S that is assertively uttered is always forced by a conflict between the conversational background, general conversational rules governing assertive utterances, and what speaker-hearers take the meaning of S (and the context) to be. By contrast, I believe that the semantic content of S can be intimately related to the proposition asserted, without there being any general but defeasible presumption that the aim of a literal, assertive utterance of S is the assertion of the proposition that S semantically expresses (in the context).

How fruitful this idea will prove to be remains to be seen. However, whatever success may be in store for us in the future will come on top of the progress made by the pioneering work done by Robert Stalnaker. No one has done more than he to open up this important field of investigation.

<sup>&</sup>lt;sup>16</sup> The second clause of R2, requiring propositions asserted to have truth values, also goes by the board. Here there is no significant truth to be salvaged, since there is no conversational rule of the sort Stalnaker imagines against asserting propositions which cannot be assigned a truth value, as opposed to those that are simply false or untrue. For discussion see Soames (1989, 583–9).

#### REFERENCES

- Davies, M., and Humberstone L. (1980), "Two Notions of Necessity," *Philosophical Studies* 38 (1980), 1–30
- Kripke, S. (1979), "A Puzzle about Belief," in A. Margalit (ed.), *Meaning and Use* (Dordrecht: Reidel), 239-83
- Rieber, S. (1992), "Understanding Synonyms Without Knowing That They Are Synonymous," *Analysis* 52, 224–28
- Salmon, N. (1990), "A Millian Heir Rejects the Wages of Sinn," in C. A. Anderson and J. Owens (eds.), Propositional Attitudes: The Role of Content in Logic, Language, and Mind (Stanford, Calif.: CSLI), 215–47
- Soames, S. (1989), "Presupposition," in D. Gabbay and F. Guenthner (eds.), *Handbook* of *Philosophical Logic*, Vol. 4: *Topics in the Philosophy of Language* (Dordrecht: Reidel), 553–616

\_\_\_\_(2002), Beyond Rigidity (New York: Oxford University Press)

- (2004), "Naming and Asserting," in Z. Szabo, *Semantics vs. Pragmatics*, (Oxford: Oxford University Press), 356–82
- \_\_\_\_(2005), Reference and Description: The Case Against Two-Dimensionalism (Princeton and Oxford: Princeton University Press)
- Stalnaker, R. (1978), "Assertion," Syntax and Semantics 9: 315-32 reprinted in Stalnaker (1999: 78-95)
- \_\_\_\_(1999), Context and Content (Oxford: Oxford University Press)
- (2001), "On considering a Possible World as Actual," Proceedings of the Aristotelian Society, suppl. vol. LXXV: 141–74; reprinted in Stalnaker (2003: 188–200)
- (2003), Ways a World Might Be (Oxford: Oxford University Press)

Judith Jarvis Thomson chap11.tex V1 - May 30, 2006 10:53 A.M.

# Queries in Chapter 11

¢

Q1. This is fine?