"Assertion" and Intentionality
Forthcoming in *Philosophical Studies*
Jason Stanley

My philosophical preoccupation has been, and continues to be, the problem of intentionality the problem of saying what it is to represent the world in both speech and thought. The problem expands, since one can never fully disentangle questions about the nature of representation from questions about the nature of what is represented. We can describe and think about the world only with the materials we find in it.
-Robert Stalnaker[1]

Beliefs and desires guide us in action. In accounting for the nature of these states, it is natural to make appeal to somewhat mysterious representational notions. For example, the content of the belief that Hesperus is Phosphorus *prima facie* involves two different ways of thinking about the planet Venus. One method of fitting ways of thinking into a naturalistic world view is by associating them with psychological types, perhaps as expressions in a language of thought. But this holds our practice of attitude ascription hostage to a thorny empirical hypothesis about the structure of the mind/brain.

Robert Stalnaker's philosophical project is to give a naturalistic explanation of how rational agents can have representational states, one that is at the very least consistent with the hypothesis that representation does not involve a language of thought. Stalnaker's paper "Assertion" is a central element in this project, written to address problems with Stalnaker's account of intentionality. It is now thirty years after the publication of "Assertion", and so an appropriate time to evaluate the success of the task.

I begin by showing how Stalnaker's "causal-pragmatic" account of the problem of intentionality is supposed to motivate his theory of the content of representational states. I then show how the analyses in Stalnaker's classic paper "Assertion" fit into this project, as responses to certain problems with his view of content. I subsequently raise two kinds of worries. First, I will raise concerns about the relation between the causal-pragmatic account of intentionality and his favored account of content. Secondly, I will discuss concerns about the adequacy of the mechanisms described in "Assertion" for resolving problems with the account of content.

---

[1] http://www.pyke-eye.com/view/phil_II_19.html

1. From a Theory of Intentionality to the Nature of Content

Stalnaker's purpose is to provide a naturalistic solution to the problem of intentionality –
to explain how physical beings can have intentional states such as beliefs and desires.
Stalnaker calls his account a causal-pragmatic theory of intentionality. Stalnaker's causal-
pragmatic account is a version of what is often called *informational semantics*. The goal
of an informational semantics is to provide a reduction of content bearing states to purely
naturalistic notions by analyzing them as involving law-like, causal relations between
agents and the environment. So, the belief relation is identified with a naturalistic relation
R, and particular states of believing that p are identified with the naturalistic states of
being in the state Rn, for some n. In order for Rn to be the naturalistic state corresponding
to the belief that p, Rn must have the functional role of the belief that p, and Rn must,
under certain conditions C, carry the information that p. As Barry Loewer (1987, p. 288)
writes, in order for an informational semantic (IS) proposal to be successful, it must meet
the following desiderata:

> *(IS)* is intended as providing a *reduction* of facts about what A believes to
> physical facts, similar to, e.g., "Water is H20". In view of this, *(IS)* must satisfy
> two adequacy conditions if it is to be a correct reduction. (1) *(IS)* and its instances
> must be law-like and true. The test of (IS)'s correctness is that the beliefs that it
> attributes must tally with those attributed by folk psychology. Its law-likeness
> would presumably come from its systematically associating belief states with their
> contents. (2) The states [Rn], the conditions C, and the notion of information
> employed must all be specifiable without appeal to semantic or intentional
> notions. This means, for example, that the C, are characterized without recourse to
> other beliefs, to meanings, and so forth.

A satisfactory naturalistic account of intentionality should explain what it is for a creature
to have a belief or desire with a particular content, without exploiting intentional notions.
In order for the account to be naturalistic, neither the explanation of the relation of belief
or the relation of desire, nor the characterization of the content of the belief or desire, can
employ ineliminable intentional vocabulary.

      Stalnaker proposes that intentional states like belief and desire have two aspects to
them. First, they have a forward looking aspect, which tells us what an agent with the

belief that p or the desire that p is disposed to do – i.e. which tells us the functional role
of that state:

> Belief and desire…are correlative dispositional states of a potentially rational
> agent. To desire that p is to be disposed to act in ways that would tend to bring it
> about that p in a world in which one's beliefs, whatever they are, were true. To
> believe that p is to be disposed to act in ways that would tend to satisfy one's
> desires, whatever they are, in a world in which p (together with one's other
> beliefs) were true. (Stalnaker, 1984, p. 15)

Secondly, they have a backward looking aspect, which tells us what is indicated to be the
case by the fact that an agent has the belief that p:

> What I want to suggest is that belief is a version of the propositional relation I
> called indication. We believe that p just because we are in a state that, under
> optimal conditions, we are in only if p, and under optimal conditions, we are in
> that state because p, or something that entails p. (Ibid. p. 18)

Thus, his causal-pragmatic account of mental states entails that:

> Beliefs have determinate content because of their presumed causal connections
> with the world. Beliefs are *beliefs* rather than some other representational state,
> because of their connection, through desire, with action. (Ibid. p. 19)

It should be clear that the causal-pragmatic account is a model of an informational
semantics.

As Loewer points out, the notion of information employed in an informational
semantics must be characterized without recourse to intentional notions. Stalnaker's
favored account of content is just such a notion of information. According to Stalnaker,
the informational content of a state is a function from metaphysically possible worlds to
truth-values. The function is not itself identified by intentional vocabulary. For example,
the function expressed by "Mark Twain is a writer" is specified by the semantic values of
the constituent elements of the sentences, which are crucially for Stalnaker, objects and
properties in the world, rather than ways of thinking. Central to Stalnaker's philosophy is
adherence to the direct reference theory of names – their only semantic content is their

referent. Since "Mark Twain" refers to the same object in the world as "Samuel Clemens", "Mark Twain is a writer" expresses the same function from metaphysically possible worlds to truth-values as "Samuel Clemens is a writer". This is a 'coarse-grained' account of content; many utterances that appear to express different propositions will end up expressing the same proposition on this picture. For example, any two utterances of necessary truths will end up expressing the same proposition – e.g. "Water is H_2O" and "Water is water" – or, for that matter, any two utterances of mathematical truths.

According to Stalnaker, the causal-pragmatic account of intentionality is supposed to *entail* the coarse-grained, possible-worlds account of content. He gives two reasons. First, it is supposed to follow from the fact that belief is an indication relation that, if p and q are necessarily equivalent propositions, the belief that p indicates the same fact as the belief that q (Ibid. p. 24). Presumably, Stalnaker is here presupposing that facts are individuated in a coarse-grained manner. Secondly, it is supposed to follow from the pragmatic aspect of belief and desire that all that is relevant for content is distinguishing between different metaphysical possibilities. How those metaphysical possibilities are represented is not relevant:

> It is essential to rational activities such as deliberation and investigation that the participants represent alternative possibilities, and it is essential to the role of beliefs and desires in the explanation of action that the contents of these attitudes distinguish between the alternative possibilities. The particular ways in which alternative possibilities are represented, or the particular means by which distinctions between them are made, are not essential to such activities and explanations, even if it is essential that the possibilities be represented, and the distinctions be made, in some way or other. (Ibid. p. 23)

Belief is a state that guides an agent in deliberation and inquiry. But what is essential to deliberation and inquiry is distinguishing between alternate possibilities. A theory of inquiry and deliberation does not need take account of the way in which agents represent alternate possibilities. As Stalnaker writes:

> According to the kind of account I have outlined, the form in which beliefs and desires are represented is not essential to their content. Two different agents might

have the same beliefs even if the forms in which the beliefs are represented are radically different. The conceptual separation between form and content is, I think, the central feature which distinguishes the conception of thought implicit in the pragmatic picture from the one implicit in the linguistic picture. (Ibid.)

So specifying the functional role of the belief that p does not require us to distinguish more finely than metaphysical possibility. Differences that cut finer than metaphysical possibilities are just matters of how possibilities are represented; they are not differences between possibilities. So what guides us in inquiry and deliberation should not be sensitive to differences that cut finer than metaphysical possibilities. Since what guides us in inquiry and deliberation just provides the functional role of beliefs and desires, the contents of these states should cut no finer than the possible worlds account tells us they should.

Stalnaker's causal-pragmatic account of the problem of intentionality is elegant. The causal aspect of the account tells us that belief is an indication relation. Indication relations are perfectly naturalistic (e.g. under optimal conditions, the rings on a tree indicate its age). But for a particular state of indicating that p to serve as the belief that p, it must also plausibly play the functional role of the belief that p. Fortunately, for the state of indicating that p to play the functional role of the belief that p, we do not need a notion of content that cuts more finely than the coarse-grained, possible-worlds account, since that is the only notion that is required in an account of deliberation and inquiry. Finally, the coarse-grained possible-worlds account of content is a notion of content that can be characterized without recourse to beliefs or meanings or any other intentional notion. So the naturalistic desiderata are satisfied.

2. The Role of "Assertion"

Unfortunately, there are serious worries with Stalnaker's coarse-grained, possible-worlds account of content. As we have seen, for example, it entails that necessarily equivalent propositions are the same. So there is only one mathematical truth, and we all know it, since we all know that 0=0. Though a number of philosophers have defended the thesis that knowledge is closed under *known* entailment, the possible worlds account of content

5

entails that both knowledge and belief are closed under *entailment*, which is a thesis shockingly at odds with common sense in the most mundane of cases. Stalnaker has much of interest to say both about the problem of deduction, and the particular case of mathematics (Ibid. Chapter 5, Stalnaker (1999b, c)). Though I am not in the end convinced by his many discussions of the matter, I will not pursue these worries here. For even setting the problem of deduction and the case of mathematics aside, there are serious problems for the possible-worlds view, namely the ones which the framework of "Assertion" is devised to solve.

Stalnaker accepts the direct reference view of names – the view that there is nothing to the meaning of a name but its bearer. The case for the direct reference view of names was made by intuitions and thought-experiments, rather than by appeal to theoretical commitments. Yet Stalnaker (1999d, p. 213) thinks the intuitions that support the direct reference view of names "do not just reflect an accidental fact about the way we happen to talk about speech and thought", but instead "show something about the nature of intentionality." Stalnaker argues (1999d) that what they reflect is the externalism that is part and parcel of the information-theoretic account of intentionality.

As we have seen, according to the information-theoretic account of intentionality, states carry content in virtue of relations to the environment. It is because of the fact that under normal or optimal conditions, they obtain only if the relevant state of the environment obtains that their obtaining indicates that the environment is in a certain state. It is therefore no surprise that our intuitions about content are sensitive to the actual objects and properties in our actual environment upon which the obtaining of our information bearing states counterfactually depend. The fact that we think the truth of our "Hesperus" utterances depends upon the object Hesperus in counterfactual situations in which another object is the first star seen in the evening is a reflection of the deeper truth that the content of "Hesperus" is determined by its causal relation to the actual object Hesperus. The intuitions that direct reference theorists discovered are reflections of the naturalistic information-theoretic account of content that is at the root of Stalnaker's view of intentionality. The central purpose of "Assertion" is (as Scott Soames, 2005a, p. 85 also notes) to square commitments to Stalnaker's orthodox possible worlds account of content with semantic conclusions from the direct reference program.

Consider "Hesperus is Phosphorus" and "Hesperus is Hesperus". Suppose we accept the direct reference view of names, the view that there is nothing to the meaning of a name but its bearer. Suppose we accept further that the two sentences are not context-sensitive. It follows that utterances of these sentences semantically express the same proposition. But there is a persuasive argument that utterances of "Hesperus is Phosphorus" express a different proposition than utterances of "Hesperus is Hesperus". After all, "Hesperus is Phosphorus" is potentially informative. In contrast, "Hesperus is Hesperus" is not potentially informative. If an utterance of a sentence is informative in virtue of the proposition it semantically expresses being informative, then one can conclude that "Hesperus is Phosphorus" and "Hesperus is Hesperus" sometimes can be used to semantically express different propositions. If we add the further premise that the two sentences are not context-sensitive, and therefore express the same propositions in every context, we can conclude that "Hesperus is Phosphorus" and "Hesperus is Hesperus" never express the same proposition. *Mutatis Mutandis* with necessarily true theoretical identity statements, such as "Water is H_2O" versus "Water is water".

Stalnaker's view of content entails that utterances of any two necessarily equivalent propositions express the same proposition. Assuming direct reference about names, and assuming that (1) and (2) do not contain any context-sensitive elements, this entails that (1) and (2) express the same proposition:

(1) Hesperus is seen in the morning sky.
(2) Phosphorus is seen in the morning sky.

Yet it certainly seems that one can believe what (1) expresses without believing what (2) expresses. If so, then (1) and (2) express different propositions. These are all worries with Stalnaker's framework that are independent of the problem of deduction.

In defense of the coarse-grained account, Stalnaker writes:

> …the thesis that necessarily equivalent propositions are identical – the main substantive consequence of the possible worlds analysis of propositional content – is a thesis that is tied to, and motivated by, the causal-pragmatic explanation of intentionality…[this shows] that they, and the possible worlds analysis of proposition that goes with them, have a deeper philosophical motivation than has

sometimes been supposed. If this definition had been proposed simply as a technical apparatus meant to systematize brute intuitions about the structure and identity conditions for objects of belief, then the examples of necessary truths and other nontrivial equivalences would show that the definition had missed the mark. The proper response would be to replace the technical apparatus with one that could make finer discriminations between the contents of attitudes and expressions. But since we have an argument to show that the identity conditions are right, as well as examples that seem to show that they are wrong, the proper response is not so clear. Stalnaker (1984, p. 24)

In Chapter 2 of Inquiry, Stalnaker makes the case that the causal-pragmatic account of intentionality is more promising than its most obvious rivals.[2] The causal-pragmatic - account of intentionality leads to the coarse-grained, possible worlds account of content. Since Stalnaker thinks the causal-pragmatic account of intentionality is the most promising naturalistic approach, we should not take apparent counterexamples to the possible worlds account of content at face-value. Since the success of the most plausible naturalistic account of intentionality is at stake, we should instead seek alternative accounts of the counterexamples to the possible worlds account. Here is where we find the motivation for the analyses in Stalnaker's classic paper "Assertion"; they are provided to explain away some of the apparent counterexamples to the possible worlds account of content. It is because Stalnaker has a prior justification for taking content to be sets of metaphysically possible worlds – a justification stemming from the causal-pragmatic account of mental states – that it is legitimate for him to seek an alternative account of apparent counterexamples.

Let us now turn from the motivation for "Assertion" to the strategies developed in it. The strategy of the paper is to argue that, while utterances that express necessarily equivalent propositions have the same semantic content, they nevertheless can used to assert different propositions. This is a strategy that has in the ensuing years since the publication of "Assertion" become familiar from the work of advocates of direct reference theories of proper names who work within the structured 'Russellian

---

[2] Stalnaker considers and rejects both the account of intentionality defended in Hartry Field's (2001a), according to which the intentionality of mental states is to be explained in terms of the intentionality of expressions in the language of thought, as well as the interpretationist program of Donald Davidson. Field (2001b) persuasively replies to some of Stalnaker's critiques of him.

proposition' tradition, such as Nathan Salmon and Scott Soames.[3] But it is worthwhile to examine its distinctive implementation in the framework of "Assertion".

Perhaps the crucial element of the framework in "Assertion" is the notion of a *context-set*. The context-set for a speaker S is "the set of possible worlds recognized by the speaker to be the 'live options' relevant to the conversation" (Stalnaker, 1999a, p. 85). A proposition is *presupposed* by a speaker if and only if it is true in all the worlds in her context-set. The propositions an agent presupposes are those "whose truth she takes for granted as part of the background of the conversation (Ibid., p. 84). So, S's context set will be the set of worlds consistent with all of her presuppositions about the conversation. Finally, a context is non-defective if and only if all of its participants have the same presuppositions, and hence the same context-sets.[4]

The central notions in "Assertion" are characterized in such a way as to reflect the pragmatic view of the attitudes outlined in the first chapter of <u>Inquiry</u>. For example, a context-set is a set of possible worlds, because "[t]o engage in conversation is, essentially, to distinguish among alternative possible ways that things may be", and "[t]he purpose of expressing propositions is to make such distinctions" (Ibid., p. 85).

With this background in place, Stalnaker proposes three pragmatic principles guiding interpretation:

1. A proposition asserted is always true in some but not all of the possible worlds in the context set.
2. Any assertive utterance should express a proposition, relative to each possible world in the context set, and that proposition should have a truth-value in each possible world in the context set.
3. The same proposition is expressed relative to each possible world in the context set.

On the possible-worlds conception of a proposition, the first principle is a way of stating the two requirements that assertions must be informative and consistent with the

---

[3] See e.g. pp. 78-9 of Salmon's <u>Frege's Puzzle</u> (Cambridge: MIT Press, 1986) and Soames (2002, Chapter 3). It is worth mentioning that Salmon's solution to these problems also involves other elements, outlined in Chapter 8 of <u>Frege's Puzzle</u>.
[4] The notion of a non-defective context is obviously somewhat of an idealization (see Stalnaker's discussion of 'close enough to non-defective', Ibid.).

assumptions of the conversational participants. The second principle is a way of stating that the semantic presuppositions of any assertion must be satisfied in each world in the context set.[5] The third principle needs a bit more introduction, as it plays a central role in the explanatory apparatus of "Assertion".

On Stalnaker's view, the context set is not just supposed to reflect the live *non-linguistic* options. It is also supposed to reflect live options about the interpretation of words that may be used in the discourse. Context-sensitive discourse provides the clearest model of possible ignorance of this sort. Suppose, to use Stalnaker's example, that I utter the sentence "You are a fool". There are two kinds of ignorance possible. One is that you might not know whether or not you are a fool. The other is that you might not know who I am referring to with my use of "you"; perhaps I am speaking to the person standing next to you. If you suffer from the latter kind of ignorance, then in one world in your context set, the proposition expressed by "you are a fool" will differ from the proposition expressed in another world in your context set. In one world in your context set, my utterance of "you are a fool" will express a proposition about you; in another world in your context set, it will express a proposition about the person standing next to you. This is a case in which the same proposition is *not* expressed relative to each possible world in your context set. Since I don't know who you meant to refer to by your use of "you", I do not know which proposition you expressed, which is reflected by the fact that, relative to different worlds in my context-set, different propositions were expressed by your utterance of "you are a fool".

Clearly, efficacy in communication is best brought about by adherence to Principle 3. If Principle 3 is violated by an assertive utterance of a sentence S in X's context set, as it was in the previous example, then X will not know what proposition was expressed by the utterance of S. But efficient communication presupposes that our interlocutors do know what propositions are expressed by our utterances. As a conversational agent, I am responsible for preserving efficiency in communication, by

---

[5] By "semantic presupposition" I am here referring to those presuppositions the failure of which gives rise to truth-value gaps. There may also very well be presupposition requirements dictated by conventional rules of the language, the violation of which lead to falsity rather than truth-value gaps. The second principle does not speak to these kinds of presuppositions.

only uttering sentences that express propositions my interlocutors know that they express, and as my interlocutor, you can reasonably expect that I adhere to this norm. In short, Principle 3 is a pragmatic principle governing conversation, akin to Grice's maxims.

With these principles in place, we can explain Stalnaker's account of some of the problematic counterexamples to the possible-worlds characterization of propositions. Consider "Hesperus is Phosphorus" and "Hesperus is Hesperus". As we have seen, there is a persuasive argument that these two sentences express different propositions. There is a similarly persuasive argument that "Hesperus is a planet" and "Phosphorus is a planet" express different propositions. Both arguments presuppose that epistemic properties and relations hold of the propositions semantically expressed by a sentence. For example, the first argument we discussed, for the conclusion that "Hesperus is Phosphorus" and "Hesperus is Hesperus" express different propositions, presupposed that an utterance is informative in virtue of the informativity of the proposition it semantically expresses. Similarly, the second argument presupposed that what one believes is what is semantically expressed by utterances of (1) and (2). Stalnaker uses the apparatus in "Assertion" to reject these assumptions. More specifically, he uses the framework to construct contingent propositions, which he then argues are what is communicated (though not semantically expressed) by utterances of "Hesperus is Phosphorus". It is these contingent propositions that are what are informative and what are believed, rather than what is semantically expressed by "Hesperus is Phosphorus".

Suppose you and I overhear a woman in Penn Station speaking loudly about counterpart theory.[6] I say to you, "That's either Karen Bennett or Delia Graff Fara". What information do you grasp by the utterance of my sentence? The semantic content of my utterance is a singular proposition. Furthermore, it is either necessarily true or necessarily false. But I cannot be intending to communicate the proposition semantically expressed. After all, lacking knowledge of the reference of "that", I do not know which proposition is semantically expressed by my utterance. Suppose there are three worlds in the context set, i, j, and k. Suppose i is a world in which the woman is Karen Bennett, j is a world in which it is Delia Graff Fara, and k is a world in which it is someone else, say

---

[6] This is Stalnaker's example, from Stalnaker (1999a, 91). I have updated it to fit the times.

Lil' Kim. Clearly, what I am trying to say is that we are in world i or j, and not in world k. This is not the proposition semantically expressed by my utterance in any of the worlds in the context-set, i, j, or k. Rather, it is the proposition that is true in a world w in the context set if and only if the sentence I utter is true when it is considered as uttered in w. This is the so-called *diagonal proposition*.

In this case, I utter the sentence "That's either Karen Bennett or Delia Graff Fara", intending to communicate not its semantic content, but rather a proposition constructed from the live options in the context about its semantic content. That is, I utter a sentence intending to communicate the diagonal proposition, which is a proposition defined using the context-set. You employ diagonalization in interpreting me, which is the process of constructing the diagonal proposition using the context-set (in this case, representing the set of live options for the semantic content of the sentence I uttered), and assigning it to every world in the context set as an interpretation of my utterance.

Stalnaker's framework makes some intuitive sense for mutual ignorance about the referent of context-sensitive expressions. But the chief application of the framework in "Assertion" is to expressions that Stalnaker does not regard as context-sensitive. For Stalnaker (as for Carnap before him), being informative must always be explained in terms of the contingency of some proposition:

> …any way of conceiving of necessary truths as having content is at the same way a way of conceiving them as contingent – as one way things could have been among others. This is, I think, because we do think of content and information in terms of alternative possibilities. Whether the source of my information is my senses, authority, or a faculty of Intellectual intuition with access to a Platonic realm of abstract entities, its deliverances are not news unless they might have been different. (Stalnaker, 1984, p.85)

Stalnaker aims to use the framework to explain the phenomenon of the "necessary a posteriori" – cases in which an utterance seems both to be an expression of a necessary truth, but also an informative claim. The explanation for the informativity is that the

diagonal proposition is contingent. The explanation for the necessity is that the proposition semantically expressed is necessary.[7]

How can one explain the potential informativity of "Hesperus is Phosphorus", on the assumption that it always semantically expresses something trivial? It is not enough to note that an utterance of "Hesperus is Phosphorus" violates Principle 1, which tells us that what is said must be true in some but not all of the worlds in the context set. Violations of Principle 1 are not enough to explain why reinterpreting the utterance via diagonalization is informative. If "Hesperus is Phosphorus" expresses the necessarily true proposition in every world in the context set, then the diagonal proposition will also be the uninformative necessarily true proposition. In order for the diagonal proposition to be non-trivial, "Hesperus is Phosphorus" must express something other than the necessarily true proposition in at least one world in the context set.

The way that Stalnaker ensures the result that an utterance of "Hesperus is Phosphorus" conveys something other than the necessarily true proposition in at least one world in the context set is by stipulating that in any context in which this is informatively uttered, there is ignorance about what proposition is expressed by the utterance. In *some* sense of "knowledge of reference", such a context involves lack of knowledge of the reference of either "Hesperus" or "Phosphorus". This is not ignorance about the contextual factors that determine the reference of context-sensitive elements in the sentence used. It is rather ignorance of *the language being spoken*. In at least one world in the context set, at least one of "Hesperus" and "Phosphorus" is assigned something other than the planet Venus. This represents that it is a live option that this name has a

---

[7] Stalnaker's theory thus belongs in the tradition of views that take modal operators to operate on something different than epistemic operators such as "it is a posteriori/a priori that". What is informative about an assertion is often one thing, for Stalnaker, and what is necessary is another. When the two diverge, Stalnaker takes what is said to be informative, or "news", and what is necessary to be "the semantic content". Evans (1985a), Forbes (1989, Chapter 5), and Stanley (1997, 2002) also think what is informative is often or always different from the object of modal predicates, and take what is said or what is believed to be the notion linked to the epistemic notions. The views I took in Stanley (1997, 2002) are explicitly inspired by Stalnaker's discussion in "Assertion" (see Stanley, 1997, pp.146-8, and pp. 155-6), though following the terminology in Dummett (1991, p. 48), I call the object of modal operators, *the ingredient sense*.

reference other than its actual one. Since reference is meaning, on the direct reference semantic theory Stalnaker assumes, it is a live option that one of the names has a meaning other than its actual one.

When "Hesperus is Phosphorus" is uttered, it expresses the necessarily true proposition in some worlds in the context set. But in at least one world, it expresses the necessarily false proposition (a world in which, say, "Hesperus" denotes Mars). So, when "Hesperus is Phosphorus" is informatively uttered, it results in a violation of Principle 3; the same proposition is not expressed with respect to every world in the context set. Since Principle 3 is violated, the hearer employs the diagonalization operator as a repair strategy. Relative to each world in the context set, the diagonal proposition is what is conveyed. The diagonal proposition is a proposition that is true in just those worlds in which "Hesperus is Phosphorus" is true, considered as uttered in that world. As a result of accepting the diagonal proposition, the hearer ends up eliminating from the context set those worlds in which "Hesperus" denotes something different than "Phosphorus".[8]

Stalnaker's account of the potential informativity of necessary a posteriori statements presupposes that they are informative because they have the effect of correcting our semantic beliefs. An utterance of "Hesperus is Phosphorus" –or an utterance of any other instance of the necessary a posteriori, such as "Water is H_2O" – can only felicitously be made in a context in which there is uncertainty about the meaning of at least one of the terms in the sentence being uttered. The effect of uttering the instance of the necessary a posteriori is to eliminate this uncertainty. If there were no uncertainty about the meanings of any of the terms in the sentence uttered, then the diagonal proposition would not be informative, and diagonalization would not work as a repair strategy.

It seems prima facie odd to hold that the ignorance that we have that makes instances of the necessary a posteriori informative is *semantic* ignorance. As Frege

---

[8] Why not treat the elimination of the deviant semantics world as a repair strategy? That is, why doesn't the pragmatic repair strategy involve narrowing down the context set to worlds in which the semantics is the same? Then we are back to the problem with explaining informativity violations with the first Principle alone. Then, the proposition expressed by "Hesperus is Phosphorus" would be the necessarily true one in every world, and the assertion will have been pointless (a violation of the first Principle).

famously pointed out, when someone tells us that Hesperus is Phosphorus, we seem to learn something about the world, not our language.[9] Stalnaker has several kinds of defenses against this response. First, he points out that learning something about the language can involve learning something about the world.[10] Secondly, he produces examples of the necessary a posteriori that do clearly seem to be informative because they resolve semantic ignorance. For example, in "Assertion" (Stalnaker, 1999a, p. 92), he employs the example "An ophthalmologist is an eye doctor." It does seem right that this is only assertible in contexts in which participants have semantic ignorance. It also seems that the point of uttering it is to resolve this semantic ignorance. Finally, diagonalization seems an adequate account of how the hearer is supposed to use the utterance to resolve the semantic ignorance. For example, the intended audience of an assertion of "An ophthalmologist is an eye doctor" is supposed to eliminate those worlds in which "ophthalmologist" does not mean eye doctor from their set of live options.

Stalnaker's view of what is asserted by an utterance of "Hesperus is Phosphorus" allows him to reply to one of the arguments above for the conclusion that an utterance of "Hesperus is Phosphorus" expresses a different proposition than an utterance of "Hesperus is Hesperus". Recall that according to the argument, the two propositions must be different, because it is possible that what is said by an utterance of "Hesperus is Phosphorus" is informative, whereas it is not possible that what is said by an utterance of "Hesperus is Hesperus" is informative. Stalnaker's response is that when an utterance of "Hesperus is Phosphorus" is informative, this is not because *the proposition it semantically expresses* is informative. Rather, it is informative because *what it asserts* is informative, which is distinct what it actually semantically expresses. That is, when an utterance of "Hesperus is Phosphorus" is informative, it is because the diagonal proposition it conveys in that context is informative. So Stalnaker denies the assumption

---

[9] A different concern is that it doesn't seem that the proposition conveyed by an utterance of "Hesperus is Phosphorus" or "Water is H_2O" is a metalinguistic proposition, that is, a proposition *about* words.
[10] "But while [the information that Hesperus is Phosphorus] is in this sense semantic, it is also astronomical. One who learns that Hesperus is Phosphorus learns something about the way the solar system is arranged." (Stalnaker, 1999d, p. 236); "…astronomical facts and semantic facts…are interconnected, and perhaps cannot be separated…" (Stalnaker, 2003a, p. 200).

in the argument that the potential informativity of an utterance of "Hesperus is Phosphorus" has a bearing on the proposition it semantically expresses – that is, the proposition that is assigned to it by the actual context-independent semantic rules of the language.

Stalnaker has a parallel response available for the second version of the argument, involving "believes". Someone could believe what is said by an utterance of "Hesperus is Phosphorus" without believing what is said by an utterance of "Hesperus is Hesperus". But this does not entail that the two utterances can semantically express distinct propositions. As I have explained, Stalnaker takes what is said by these utterances to be the diagonal proposition conveyed in the relevant context by the utterance of "Hesperus is Phosphorus", rather than what is semantically expressed.

The framework developed in "Assertion" has additional flexibility not brought out in that paper. First, according to Stalnaker, the apparatus of diagonalization is not just applicable to the case of assertion. It is also what provides candidates for the objects of belief in the semantics of belief ascription (Stalnaker (1999f,g)). In other words, when we make a belief ascription involving a sentence S that semantically expresses a necessary proposition, but is intuitively a posteriori (that is, potentially informative), we are ascribing to the agent belief in the diagonal proposition asserted by the utterance of S relative to a certain context. This context is not the *conversational* context, because I might be ascribing to someone belief in a proposition that the conversational participants know to be false. Rather, it is what Stalnaker calls the "derived context", which may be "different, often disjoint, from the basic context" (Stalnaker, 1999f, p. 157). So, for example, if Sue utters in a conversation with Frank, "John believes that Hesperus is Phosphorus", what is at issue is the diagonal proposition constructed from a set of worlds determined by Sue and Frank's beliefs about what John believes.[11]

There is a second way the framework in "Assertion" can be brought to bear on various semantic puzzles, but it is one Stalnaker does not mention. Consider again (1) and

---

[11] There are definitely complications in the application of the framework to belief-ascriptions. For example, if John speaks a language neither Frank nor Sue does, it's rather complicated to see how the derived context is even set up. Stalnaker's most extended discussion of this issue (Stalnaker, 1999g, pp. 127-8) is quite compressed, and I do not clearly grasp his solution.

(2). Both of these sentences semantically express contingent propositions, so Stalnaker can account for the fact that utterances of each of them are potentially informative, without appealing to diagonalization. Yet (1) and (2) appear to express propositions that are differentially informative, in that one can believe that Hesperus is seen in the morning sky, without believing that Phosphorus is seen in the morning sky. It is hard to see how the apparatus of diagonalization can be brought to bear in this case. But in fact the framework Stalnaker lays out in "Assertion" can provide an account of the differences between (1) and (2).

Suppose we are in a context in which it is a live option that "Hesperus" and "Phosphorus" do not co-refer – in other words, a context in which one can legitimately assertively utter "Hesperus is Phosphorus", and thereby assert the diagonal proposition. There will be at least one world in that context set in which one of either "Hesperus" or "Phosphorus" receives a semantic value other than its actual semantic value, Venus. Suppose that it is known that "Phosphorus" refers to Venus, but it is a live option that "Hesperus" refers to a different planet. Then, assuming that it's also a live option that "Hesperus" refers to Venus, any utterance of a sentence containing "Hesperus" will result in a violation of Principle 3 – different propositions will be expressed by that utterance relative to different worlds in the context set (since there are different live options for the semantic value of "Hesperus"). So, by uttering any sentence containing "Hesperus", what one will assert is the diagonal proposition, since diagonalization will be used as a repair strategy. It therefore follows that, relative to contexts in which it is not known that "Hesperus" and "Phosphorus" co-refer (and there are different live options for the semantic value of at least one of these terms), utterances of (1) and (2) will convey different propositions. At least one utterance will convey the diagonal proposition. So in fact Stalnaker's framework can give an account of the difference in cognitive value between assertions of (1) and assertions of (2).

A number of philosophers, including those with less coarse-grained accounts of content, are committed to accounts of content that entail that when S and S' differ only in that S' contains a co-referring name in the same syntactic position as S, then S and S' semantically express the same proposition, relative to the same context. All such philosophers think that it is *pragmatics* that explains the felt differences. But Stalnaker's

theory has a number of advantages over alternative theories. First, some of the alternative proposals (e.g. Salmon (1986, Chapter 8), Braun (1998)) appeal prominently to *guises* or *ways of grasping* objects and propositions. In contrast, Stalnaker does not appeal to such entities anywhere in his defense of coarse-grained contents. For him, they are either not part of the "materials we find" in the natural world we have available to us to describe representational states, or they must be explained in terms of aspects of the linguistic picture of intentionality that he rejects. Secondly, many of the other alternatives appeal (at least in part) to error-theories, according to which the felt differences result from confusing the proposition pragmatically conveyed with the one semantically expressed.[12] In contrast, Stalnaker's theory is not an error theory of any kind. The employment of diagonalization in response to hearing an utterance is not an error, but the rational strategy to take in light of encountering an utterance that violates a general conversation norm. Relatedly, because Stalnaker thinks that the diagonal proposition figures in the semantics of propositional attitude ascription, he does not have to give an error theory in this domain, either (though there are serious worries about whether the resulting semantics for propositional attitude ascriptions is compositional).

Stalnaker's account also does not require a drastic revision of a broadly Gricean conception of the semantics-pragmatics distinction.[13] Certainly, the process by which a diagonal proposition is asserted is not exactly like an implicature. Unlike the case of implicature, the semantic content of the sentence uttered is not in any sense what is said, or even what the speaker "makes as if" to say. However, the semantic content nevertheless is intimately involved in determining what is said, in a way familiar from Grice. An utterance is made, which flouts general conversational norms. Recognizing this, the audience seeks a proposition other than the semantic content that the speaker

---

[12] As Soames (2002, p. 228) writes in summarizing his approach to explaining speakers' intuitive judgments about differences in meaning between cases such as S and S', "…those intuitive judgments are based on a confusion of what a sentence means…with what speakers use it to assert and convey in particular contexts…speakers are prone to confuse the semantic content of a sentence, its meaning in the language, with what a speaker uses it to assert in a particular context…"

[13] In contrast, some recent accounts of the challenges facing direct reference theories have appealed to non-Gricean pragmatic processes, such as free pragmatic enrichment (e.g. Soames, 2005b).

plausibly is trying to communicate by her utterance. The proposition asserted in such a case does not float free of its semantic content. As with the explanation of most implicatures in the Gricean framework, the semantic content is one of the sources of the violation of the conversational norms (e.g. an utterance of "Hesperus is Phosphorus" violates the principle that assertions must be informative and consistent with the context set). The proposition conveyed (the diagonal proposition) is also constructed in part as a *function* of the actual semantic content, in cases in which the actual world (with its actual semantics) is one of the worlds in the context set. So while the process is not strictly a Gricean one, in which the semantic content is also what is said, which is what triggers Gricean violations, it is nevertheless non-mysterious. It is explicable for the same reasons that Grice's explanation of the phenomenon of implicature is explicable.

It is a noteworthy fact about Stalnaker's explanation of the puzzles occasioned by co-referring names that it is difficult to separate from the possible-worlds conception of content that gives rise to them. Its implementation requires something like a possible-worlds conception of content. To construct a proposition that is the diagonal proposition, one needs an update operation and a set of point-wise alternatives as a representation of context, which yields a set of such points as a value. It is difficult to see how to construct a diagonal proposition in (for example) a framework that treats propositions as structured n-tuples of objects and properties. In short, Stalnaker's pragmatic explanations seem to rely on the possible-worlds conception of proposition he advocates. The lack of an equally plausible pragmatic explanation of the informativity differences between "Hesperus is Phosphorus" and "Hesperus is Hesperus" on other frameworks can then be used by Stalnaker as a point in favor of the possible-worlds conception.

Here are two concerns about Stalnaker's theoretical account of the counterexamples to the coarse-grained, possible worlds account, both due to Scott Soames. First, Stalnaker's account presupposes that a sentence like "Hesperus is Phosphorus" can only be informatively uttered in a context in which there is ignorance about the meaning of "Hesperus" or the meaning of "Phosphorus". But it seems that this presupposition is false. As Soames (2005, p. 93) writes:

No matter what elaboration might be offered, it is not true that ["Hesperus is identical to Phosphorus"] would be used only in a conversation in which it is not presupposed that conversational participants know what 'Hesperus' and 'Phosphorus' mean and refer to, in the sense most relevant to Stalnaker's model. Certainly, each of the conversational participants 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, these speakers know of the referent of each name that it is the referent of that name.

For example, it seems that someone can know that "Hesperus" refers to Hesperus, and that "Phosphorus" refers to Phosphorus, without knowing that Hesperus is identical to Phosphorus. But then Stalnaker's account cannot be correct.

However, Stalnaker has an easy reply to this objection (see Stalnaker (2006)). In a context in which "Hesperus is Phosphorus" can be informatively uttered, utterances of "'Hesperus' refers to Hesperus", or "'Hesperus' refers to that object" [pointing in the evening to Venus] communicate truths. But at least for one of these, the truth that it communicates is the contingent diagonal proposition of that sentence, relative to that context. Furthermore, this follows straightforwardly from Stalnaker's general description of the cases. In a context in which "Hesperus is Phosphorus" can be informatively uttered, there are different candidate references for either "Hesperus" or "Phosphorus" in the context-set. So, for at least one of "'Hesperus' refers to Hesperus" and "'Phosphorus' refers to Phosphorus", there will be a violation of Principle 3. This violation will generate the pragmatic repair strategy of diagonalization. So in some sense it is true that it will be presupposed that (say) "'Hesperus' refers to Hesperus", or "'Hesperus' refers to that object" [pointing in the evening to Venus]. But what will be presupposed will be the diagonal proposition of that utterance.[14]

_____

[14] This response also undermines a recent criticism of Jeff Speaks (2006, pp. 448-449) against Stalnaker's account of the problem of deduction. Speaks assumes that if a speaker understands the statement of Fermat's last theorem, then "we may suppose that he believes the meta-linguistic proposition expressed by "'No whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power" means that no whole number raised to a power greater than two is equal to the sum of two other whole numbers, each raised to that power."". However, Stalnaker clearly would not grant that speakers would grasp the metalinguistic proposition in question – at most they would grasp the diagonal proposition conveyed by that

Here is a second worry Soames raises about Stalnaker's account, this time about its consequences for de re belief (Soames, 2005a, pp. 96-99). In a context in which I utter, pointing to a glass paperweight perceptually salient to all of us, "That is made of glass", the proposition expressed is a necessary truth (assuming various widely accepted metaphysical theses). If so, then in order for it to be informative, it must be that there are worlds in the context-set in which my use of "that" refers to something other than the glass paperweight in front of us (this allows an informative diagonal proposition to be constructed). But it seems that an utterance of "That paperweight is made of glass", when the speaker and audience are visually presented with a glass paperweight, should be a paradigm singular, de re belief. So it seems that, pace Stalnaker's account, the same paperweight should be the referent of my use of "that paperweight" in every world in the (derived) context set, on pain of denying that we share a de re belief.

Stalnaker (1999f, pp. 163ff.) defends what he calls a "pragmatic" account of de re belief attribution.[15] According to it, whether or not an ascription of the form "Jones believes that p" is a de re belief ascription about a certain object o depends upon whether Jones is in a belief state that varies as a function of facts about o. One might interpret Stalnaker's characterization of de re belief as entailing that Jones's belief that that paperweight is made of glass counts as a de re belief. If we vary facts about that paperweight, then the content of Jones's belief state will vary accordingly. For example, if we put a smudge on the paperweight, then the counterpart objects that are possible semantic values of "that paperweight" in the context set will also have smudges on them. This means that the set of possible worlds that is the object of his belief changes as a function of facts about that paperweight. Therefore, one might think that Jones's belief counts as a de re belief about that paperweight, according to the pragmatic account. However, though this account would entail that Soames's case is an instance of a de re belief, the resulting account of de re belief ascription threatens to over-generalize quite badly. For example, consider the belief ascription "Eric believes that Sarah is smart". The

---

metalinguistic proposition relative to a particular context of use. See section 4 below for a discussion of Stalnaker's view of grasp of meaning.

[15] Presumably in this paper Stalnaker is somewhat of a skeptic about de re belief *simpliciter*.

21

worry is that this counts as a de re belief ascription about Eric's brain, since he is in a state that varies as a function of facts about his brain.[16]

Whether or not Stalnaker can correctly account for Soames's case as an instance of a genuine de re belief ascription, he can nevertheless make sense out of the intuition that we can make *other* de re belief ascriptions about that paperweight. For example, our belief that the paperweight is on the table counts as a de re belief ascription on this characterization (since the semantic content of "that paperweight is on the table" is a contingent proposition). Perhaps this is enough to ameliorate the concern that his account has some unintuitive consequences.[17]

There are other unintuitive consequences of Stalnaker's theoretical account, of the sort that are familiar from Frege's classic discussion in "On Sense and Meaning". As we have seen, the *model* for Stalnaker's analysis of the necessary a posteriori is an example such as "an ophthalmologist is an eye doctor", or "A fortnight is fourteen days." These are cases in which Stalnaker's analysis seems correct. But even Stalnaker admits that intuitively these *are* cases of sentences that are uttered to resolve semantic ignorance, whereas "Water is H_2O" and "Hesperus is Phosphorus" intuitively are not:

> In some simple cases, it should be uncontroversial that [the information that distinguishes between necessary truths (which on a straightforward interpretation all have the same informational content) be semantic information – information] is right: if O'Leary fails to believe that all woodchucks are groundhogs, or that a fortnight is a period of fourteen days, then it is clear that the information O'Leary lacks, and the information that we are saying he lacks when we deny that he has those beliefs, is information about the semantic values of certain words. But in most cases, the information in question does not seem, intuitively, to be information about expressions. Plausible semantic theories tell us, for example, that it is necessarily true that Hesperus is Phosphorus, and that measles is caused by a virus. If I point at Oliver North, and say "*That* is Oliver North", the proposition I express is necessarily true. And of course all mathematical statements are necessarily true or necessarily false. But not of these statements seems to be about language; the information they convey seems to be about astronomical and medical facts, facts about who is being pointed at, or, in the case

---

[16] Thanks to Eric Swanson here for the example.
[17] Thanks to David Manley, Sarah Moss, and Eric Swanson for extensive discussion of Soames's objection. Sarah Moss pointed out unclarities in my original discussion, David Manley suggested the interpretation of Stalnaker on de re belief that entails that Soames's case is in fact a case of de re belief, and Eric Swanson raised the concern that Manley's construal results in an account of de re belief that over-generalizes.

of arithmetical and geometric statements, facts about numbers and the abstract structure of certain spaces. One does not need to know any names to know, or be ignorant of, the fact that Hesperus is Phosphorus, and we may share mathematical beliefs with those who do not share our language. (Stalnaker, 1999e, p. 235-6).

However, raising this objection against Stalnaker's account misplaces the dialectic. The causal-pragmatic account of intentionality entails the coarse-grained, possible-worlds account of content. There is therefore a foundational philosophical motivation for reinterpreting the apparent counterexamples to the possible-worlds framework. Though the reinterpretation of the examples may seem unintuitive, this should not be surprising. The causal-pragmatic theory is a naturalistic explanation of intentionality. Our intuitions about the distinction between linguistic and non-linguistic information might have to be revised in the face of pressure from naturalistic inquiry. Indeed, if the only consequence of such an explanation is the sacrifice of these rather recherché intuitions, one might think that it is in the character of a discovery rather than a cost.

3. Does the Causal-Pragmatic Account of Intentionality entail the possible-worlds account of content?

It's worth emphasizing how radical Stalnaker's project is. Rudolf Carnap (1949) argued for a possible-worlds account of content, and a possible-worlds account of informativity, including the informativity of utterances of sentences like "Hesperus is Phosphorus". But by "necessity", Carnap certainly did not mean metaphysical necessity – he meant some kind of conventional, linguistic notion of necessity. Carnap also thought the project of extending this account of content to the semantics of propositional attitude ascriptions was hopeless. As a result, he gave a sententialist account of propositional attitude ascriptions. In contrast, not only does Stalnaker want to defend an account of content that explains it in terms of metaphysically possible worlds, he also wants to extend the account to the semantics of propositional attitude ascriptions. The natural worry is that the project is too ambitious to be correct.

Stalnaker's general argument can be summarized as follows. First, the causal-pragmatic account of intentionality is the most promising naturalistic approach to the problem of intentionality. Secondly, the causal-pragmatic account of intentionality entails

23

the possible-worlds account of content. So, those sympathetic to a naturalistic account of intentionality should adopt a possible-worlds account of content. We thus discover, from our commitment to naturalism, that in fact whatever is going on when one utters a sentence like "Hesperus is Phosphorus", or "That is Oliver North", is analogous to what is going on when we utter "An ophthalmologist is an eye doctor" or "A fortnight has fourteen days".

However, there are problems with Stalnaker's discussion. First, Stalnaker gives little reason to think that the notion of "optimal conditions" in his characterization of indication can be spelled out in non-intentional terms. But:

> …for most beliefs, if they possess optimal conditions at all, these conditions involve other beliefs. Whether or not a person's belief-state reliably covaries with a state of affairs depends on what other beliefs that person has. For example, a person who fails to believe that fossils are derived from once-living organisms, or who believes that the earth is 6,000 years old, will not reliably form beliefs about the age of a fossil. If there are optimal conditions for forming beliefs concerning the age of fossils, those conditions will involve having certain beliefs and not having certain other beliefs. (Loewer, 1997, pp. 114-5).

In general, the project of spelling out "optimal conditions" in non-intentional terms has not successfully been carried out. These are worries with the information theoretic project, and therefore with Stalnaker's motivation for the project of "Assertion".[18]

Of greater concern, however, is that there are serious problems with Stalnaker's view that the pragmatic picture entails the coarse-grained possible-worlds account of belief. According to the pragmatic picture, the functional role of the belief that p is given by its role in guiding an agent's actions. What is essential to deliberation and inquiry is that the contents of an agent's beliefs distinguish between alternative possibilities, but "[t]he particular ways in which alternative possibilities are represented, or the particular means by which distinctions between them are made, are not essential to such activities and explanations." However, by "possibility", Stalnaker means *metaphysical* possibility. Because of this, the distinction Stalnaker draws here between a *possibility* and a *way of representing a possibility* is tendentious.

---

[18] Speaks (2006, pp. 442ff.) raises additional concerns about the role the appeal to optimal conditions plays in Stalnaker's causal-pragmatic account.

In a neutral sense of possibility, Stalnaker is no doubt right that what is essential about a belief and a desire to an agent's deliberations is just that the belief distinguish between alternate possibilities; this is the sense of possibility in which, in a proof in mathematics, we might say 'there are two possibilities A and B', and then proceed to reduce one to absurdity. What is tendentious is that in Stalnaker's intentionality-free, *metaphysical* sense of possibility all that is essential to the content of a state for the deliberations of an agent in that state is that the content distinguishes among possibilities. As Hartry Field (2001b, p. 85-86) writes:

> …even when one's task is to explain *non-linguistic* behavior, there is often a *prima facie* need to attribute a belief in one proposition and a disbelief in an equivalent proposition (i.e., there is a prima facie need for a conception of content more fine-grained than S[talnaker]-content). For instance, if I offer someone who doesn't know much mathematics $1000 for an example of a plane map that requires more than four colors to color (according to the usual coloring conventions), he will behave very differently than he would if I had offered $1000 for a trisection of a Euclidean 60 degree angle by straight edge and compass; to explain this I need to attribute different beliefs and desires to him in the different cases, and it is prima facie difficult to see how I can do this in a relevant way if the desire to do one impossible task is identified with the desire to do any other impossible task.

In a neutral sense of possibility, it is possible that there is a plane map that requires more than four colors to color. But this neutral sense of possibility is useless for Stalnaker's project, because it is a notion of possibility that we need to intentional notions to explicate. As Stalnaker (1984, pp. 24-5) writes:

> Could we escape the problem of equivalence by individuating propositions, not by genuine possibilities, but by epistemic possibilities – what the agent takes to be possible? This would avoid imposing implausible identity conditions on propositions, but unfortunately it would also introduce intentional notions into the explanation, compromising the strategy for solving the problem of intentionality.

In short, Stalnaker is incorrect that the pragmatic of the attitudes entails the coarse-grained, possible-worlds account of content. If we think of possibility as genuine metaphysical possibility, then different ways of representing the same genuine possibility are relevant – and indeed essential – for deliberation and inquiry. The pragmatic picture

does not after all lead to any "conceptual separation between form and content". At best, it leads to a conception of proposition according to which propositions are functions from worlds to truth-values, but ones that are individuated in intentionally laden terms.[19]

Field's point that we must make distinctions between non-genuine possibilities in accounting for deliberation and inquiry straightforwardly carries over to the problem space of "Assertion". If I am asked to purchase a book I believe is written by Mark Twain, I am rational to pay more for it than if I believe it is written by Samuel Clemens, and do not know that Mark Twain is Samuel Clemens. But on Stalnaker's view of content, the possibility that this book is written by Mark Twain is the very same possibility as the possibility that this book is written by Samuel Clemens. So the pragmatic picture of the attitudes, which tells us to individuate contents as finely as is required for explaining deliberation and inquiry, should lead us to a more fine-grained picture of content than the coarse-grained, possible-worlds account.[20]

Of course, Stalnaker could reply by arguing that he can distinguish the two envisaged possibilities. One is a possible situation in which there is a famous writer named "Mark Twain" who wrote this book, and the other is one in which there is a

---

[19] There are two models here. First, it could be that propositions are functions from epistemically possible worlds to truth-values (this is the possibility envisaged by Stalnaker above). Secondly, it could be that propositions are functions from metaphysically possible worlds to values, where the functions are individuated in epistemic terms, as in the work of David Chalmers (2002) and Frank Jackson (1997).

[20] Perhaps Stalnaker could here appeal to a thesis about the metaphysics of worlds – that, for example, possible worlds are conceptually determined by the activities of rational agents. After all, Stalnaker (1984, p. 166) does claim that "Possible worlds, as I understand them, are abstractions from the dispositions of rational agents". Given this conception of possible world, there is no gap between the alternatives distinguished by the contents of the attitudes of rational agents, and the possibilities themselves. However, it is very difficult to see how Stalnaker can say this. First, as Jeffrey King (2007, p. 451) has emphasized, this would leave us with distinctions that cut more finely than metaphysical possibilities, for the reason Field gives above. Secondly, and more pressingly for present purposes, the thesis that metaphysically possible worlds are determined by the distinctions made by the contents of the attitudes of rational agents is flatly inconsistent with the informational semantic program that is the basis of Stalnaker's work. It is for this reason that, in his response to King's discussion, Stalnaker (2007, p. 482) goes out of his way to "emphatically retract" this earlier "careless remark" of his, and states "I think it is wrong that [possible worlds] depend for their existence on the kind of rational activities that we use them to help explain."

different, undistinguished writer named "Samuel Clemens" who wrote this book. But this misstates the dialectic. Stalnaker's meta-linguistic recasting of the example is intended to be a reinterpretation of an apparent counterexample. The reinterpretation is supposed to be motivated by the fact that the theory that leads to the counterexamples, namely the coarse-grained, possible-worlds account of content, is entailed by the causal-pragmatic account of intentionality. This is Stalnaker's "argument to show that the identity conditions are right". Since the causal-pragmatic account of intentionality is plausible, we should seek to reinterpret the apparent counterexamples in meta-linguistic terms. But, as we have just seen, there are worries with the causal side of the causal-pragmatic account of intentionality. More worrisomely, the pragmatic side of the causal-pragmatic account does not in fact lead to the coarse-grained, possible-worlds account of content. So there is no "argument to show that the identity conditions are right" that should lead us to reinterpret the apparent counterexamples in meta-linguistic terms.

4. Is the framework in "Assertion" adequate?

According to Stalnaker, before it was discovered that Hesperus is Phosphorus, there was ignorance either of the meaning of the name "Hesperus" or the meaning of the name "Phosphorus". This is because *what* was discovered when it was discovered that Hesperus is Phosphorus was that a certain live option for the meaning of one of these names was incorrect. Since that deviant meaning hypothesis was a live option, it was epistemically possible. Since it was epistemically possible that (say) "Hesperus" had a meaning other than it actually does, the meaning of "Hesperus" was not known. It is in this sense that Stalnaker's account of the potential informativity of "Hesperus is Phosphorus" *essentially appeals* to lack of knowledge of meaning.

       In fact, for any two names or natural kind terms in our language "a" and "b", if we do not know whether a is b, then we are ignorant of the meaning of either (both?) "a" or "b". Stalnaker is quite explicit about this commitment:

> …no one who is ignorant of the fact that Hesperus is Phosphorus can be fully informed about the informational content of statements containing the name 'Hesperus' and 'Phosphorus'. (Stalnaker, 1999e, p. 236)

Stalnaker's point here is that if one is ignorant of the fact that "Hesperus is Phosphorus" expresses a truth (that is, if one is ignorant of the diagonal proposition), one does not know what proposition is expressed by sentences containing "Hesperus" and "Phosphorus". But one knows what proposition is expressed by a statement if and only if one knows the meaning of its constituent parts. So the view is that no one who is ignorant of the fact that "Hesperus" co-refers with "Phosphorus" knows the meaning of both "Hesperus" and "Phosphorus".

There are other theorists who are committed to the view that sentences like "Hesperus is Phosphorus" or "Mark Twain is Samuel Clemens" semantically express the same propositions as trivial identity claims, as would be made by utterances of "Hesperus is Hesperus". For example, direct reference theorists about names and natural kind terms who advocate structured proposition accounts of content are in a similar position. However, many of these other theorists do not give accounts of the potential informativity of "Hesperus is Phosphorus" that make essential appeal to ignorance of meaning. Stalnaker's view is that when we find "Hesperus is Phosphorus" informative, it is because we do not know what proposition it expresses. In contrast, according to Scott Soames (2002, pp. 70-71):

> …we have no reason to expect that whenever two sentences semantically express the same proposition, competent speakers who understand the sentences will recognize that they express the same proposition, and thus mean the same thing. There are many examples in the literature, involving not just proper names and natural kind terms, but expressions of all different sorts, which show that it is quite possible for a competent speaker to understand a pair of sentences that mean the same thing without realizing that they do.

On Soames's view, even in a situation in which "Hesperus is Phosphorus" can be informatively uttered, we know what "Hesperus" and "Phosphorus" refer to, and we grasp the proposition it expresses. It is just that we do not recognize that this proposition is the same as the proposition expressed by "Hesperus is Hesperus". As a result, Soames's account of what is communicated by an utterance of "Hesperus is Phosphorus" does not appeal to ignorance of meaning.

Stalnaker's position here is reminiscent of one of Frege's arguments for the existence of sense as reconstructed by Michael Dummett:

> In invoking the notion of information to support his contention that the sense of a name cannot consist merely in its having the reference which it does have, Frege is tacitly connecting the notion of sense with that of knowledge; and this is the justification for our representing Frege's views by saying that sense is an ingredient in meaning, where meaning is that which a man knows when he understands a word. For the argument, spelled out in full, runs thus: If the sense of a name consisted in its having a certain reference, then anyone who understood the name would thereby know what object it stood for, and who understood two names which had the same reference would know that they stood for the same object, and hence would know the truth of the statement of identity connecting them, which could therefore not be informative for him. The underlying assumption is the compelling principle that, if someone knows the senses of two words, and the two words have the same sense, he must know that they have the same sense: hence, if the sense of a name consists merely in its reference, anyone who understands two names having the same referent must know that they have the same referent. (Dummett, 1973, p. 95)

Stalnaker accepts that the sense (meaning) of a name is its referent. It seems that he is also committed to the principle Dummett calls 'the transparency of sense' – that if two terms have the same meaning, and X knows the meaning of both terms, then X must know that the terms have the same meaning – after all, this is logically equivalent with Stalnaker's claim, cited above, that if X does not know, of two terms that have the same meaning, that the two terms have the same meaning, then X does not know the meanings of at least one of these terms. Like Dummett's Frege, Stalnaker concludes that if the meaning of a name is its referent, then anyone who does not accept the sentence "Hesperus is Phosphorus" does not know the meaning of either "Hesperus" or "Phosphorus". It is just that whereas Frege takes this to be a *reductio* of the position that the meaning of a name is its referent, Stalnaker takes it to be a desirable consequence, around which he builds his account of informativity.[21]

---

[21] Stalnaker's view that we (often) lack knowledge of reference is in tension with some of the considerations that lead to the direct reference theory that he advocates. For example, according one version of the semantic argument against description theories of names and natural kind terms, a speaker can know the meaning of a name or a natural kind term while possessing very little information about its referent:

A natural objection to Stalnaker's account is that it seems inadequate to the phenomenology of the situation. There are certain cases in which it does seem that we appeal to diagonalization in communication and understanding. But it does not seem that, when we utter non-trivial identity statements such as "Hesperus is Phosphorus", we do anything of the sort. We are certainly unaware of having the intention that our audience re-interprets our utterance meta-linguistically, nor do we think of the context as containing different potential meaning assignments to our words. Rather, it seems that we utter such statements with the intention that our audience gains some information about the non-linguistic world.[22]

There is a coherent perspective from which this point of phenomenology is not an objection. According to Quine (1976a, p. 132), "the lore of our fathers…is a pale gray lore, black with fact and white with convention". Stalnaker follows Quine in rejecting the distinction between semantic and worldly facts. Stalnaker would certainly use his Quinean side as a weapon against the phenomenological datum.

Equally importantly, Stalnaker does not in the end think of the apparatus described in "Assertion" as a description akin to a Gricean rational reconstruction. Rather, he thinks of it as a description of how *we as theorists* knowledgeable about the fact that Hesperus is Phosphorus, could describe what is going on in a situation in which "Hesperus is Phosphorus" is informatively uttered. As Stalnaker writes, "We can describe and think about the world only with the materials we find in it." By "we", Stalnaker means the theorist trying to give an account of intentionality. As theorists, we know that "Hesperus" and "Phosphorus" co-refer. We cannot therefore use the content of "Hesperus is Phosphorus" in our context, in which it could only be uninformatively uttered, to describe what is going on in a context in which it could be informatively uttered. The task

> The picture associated with [a certain version of the description] theory is that only by giving some unique properties can you know who someone is and thus know what the reference of your name is. Well, I won't go into the question of knowing who someone is. It's really very puzzling. I think you do know who Cicero is if you just can answer that he's a famous Roman orator. (Kripke, 1980, p. 83)

So Kripke holds that it is fairly easy to "know what the reference" of a name is.
[22] Thanks to Jeff King for this point.

is to give an adequate description of what is going on in such a context, with the resources available to us in our context – not provide a Gricean rational reconstruction of a process that is plausibly occurring in that context. On this conception of the project, it is akin to the project of the modal actualist, who seeks to describe, with just the resources available in the actual world, alternative ways things could have been.

Of course, this alternative perspective is not acceptable to those of us who do recognize a distinction between semantic and worldly facts. One way of arguing for the alternative perspective in the light of natural skepticism is by justifying it by appeal to the causal-pragmatic account. But as we have seen, insofar as it is supposed to be justified by appeal to the causal-pragmatic account, it is not. Secondly, if we think of the apparatus, as Stalnaker seems to now, as a description of what the *theorist* does, rather than as a plausible description of what is happening between conversational agents, it loses many of its advantages over alternative accounts. The appeal of the apparatus, after all, was supposed to be that it was a plausible description of the phenomena in terms of general norms motivated by the goals and purposes of conversation. This appeal vanishes when the apparatus is thought of as an account given from the theorist's perspective.

These are perhaps minor quibbles about the analysis of some of the cases that the apparatus in "Assertion" was designed to treat. My central concern, however, is not that Stalnaker fails to give an adequate account of some puzzles involving co-referring names. My worry is rather that the apparatus he provides is not adequate to explain what is happening in the full range of cases involving co-referring terms.

There are two kinds of discoveries. The first kind of discovery occurs when we learn something that eliminates an epistemic possibility. This kind of discovery resolves our ignorance. The second kind of discovery is when we learn that something we firmly believed in fact is false. This kind of discovery cures us of error. So far, we have been evaluating Stalnaker's account of the first kind of discovery. I have raised some familiar concerns about the meta-linguistic character of Stalnaker's solution. I turn now to Stalnaker's account of the second kind of discovery, which is more problematic.

Suppose that prior to having discovered that Hesperus is Phosphorus, we weren't just agnostic on whether Hesperus was Phosphorus, but rather explicitly denied that Hesperus is Phosphorus. On Stalnaker's account, then, when we discover that Hesperus is

Phosphorus, we cannot merely be eliminating worlds from our context set worlds in which "Hesperus" and "Phosphorus" do not co-refer. If we were, we would be left with no beliefs, since "Hesperus" and "Phosphorus" do not co-refer in any worlds in our context set. Something else must be occurring.

I am not completely certain what Stalnaker would say about this case. But the very fact that Stalnaker has to provide a completely different account of this discovery is an objection to his solution. Whether we believe that Hesperus is not Phosphorus, or are rather merely agnostic on the matter, should be independent of the informational content of the discovery that Hesperus is Phosphorus. But on Stalnaker's account, it is not. The informational content of the discovery that Hesperus is Phosphorus is the diagonal proposition, if we are agnostic about the question of whether Hesperus is Phosphorus. If however we formerly believed that Hesperus is not Phosphorus, then the informational content is something different. This is a problematic result.

It is also unclear what the informational content of utterances of "Hesperus is F" and "Phosphorus is F" would be, in a language community in which everyone believed that Hesperus was not Phosphorus. Diagonalization does not enter the picture, because there is no room for uncertainty about meaning. Everyone is certain that "Hesperus" has a certain denotation and "Phosphorus" has a certain denotation, and that they are distinct.

Perhaps what Stalnaker would say is that in this community, some terms are never used in accord with their actual meaning. At least one of the terms "Hesperus" and "Phosphorus" is always used with a meaning different than its actual meaning. This would deliver the result that utterances of "Hesperus is F" convey different propositions than utterances of "Phosphorus is F", and it would do so without appealing to any of the mechanisms described in "Assertion".

However, the resulting account is very problematic. As we have seen, the principle virtue of Stalnaker's explanation of informativity puzzles is that they seem to proceed from independently motivated general pragmatic norms, where what is communicated is always in part a function of the semantic content, both in the sense that the semantic content gives rise to the violation of the conversational norm, and in the sense that what is conveyed is built in part out of the semantic content. But nothing like this is going on in the envisaged account. What utterances of (say) "Hesperus is F" would

end up conveying is a proposition that is in no sense dependent upon the semantic content of "Hesperus". Rather, they are determined just as a function of the beliefs of the speaker. The actual semantic content turns out to be an idle wheel in the explanation of what is conveyed.

There are worse concerns with the envisaged account. Suppose that pre-discovery, everyone believed that Hesperus is never anywhere near Phosphorus (after all, one appears in the morning, and another at night). So according to the envisaged account, utterances of "Hesperus is never anywhere near Phosphorus" would convey truths, since (say) "Hesperus" would have a deviant semantic value, one that referred to object (Mars?) that was never near Venus. Looking back at such utterances, we would like to evaluate what was said by such utterances as false. Yet if what is asserted by an utterance of "Hesperus is never anywhere near Phosphorus" relative to pre-discovery times is a true contingent proposition, then what is said by the utterance is true, rather than false. It is hard to see how Stalnaker could avoid this conclusion if he treated what is conveyed by sentences containing "Hesperus" as in part a function of its deviant semantic value.

In sum, it is quite unclear how Stalnaker would account for the case in which it is widely believed that a is not b, when "a" and "b" in fact co-refer. Since Stalnaker thinks that if "e" and "d" have the same semantic value, no one ignorant of the fact that they do can be competent with both "e" and "d", the same problematic arises for expressions of every category.

5. Conclusion

Gottlob Frege originally raised the set of puzzles that have collectively become known as "Frege's Puzzle" – the difficulty of accounting for the fact that co-referring names can have different informational contributions. Many philosophers who reject the letter of Frege's solution – that "Hesperus" and "Phosphorus" semantically express different meanings- nevertheless accept that something like ways of thinking, perhaps properly naturalized as expressions in the language of thought, play a central role in an account of these puzzles. In contrast, Stalnaker seeks an account of Frege's puzzle that does not appeal to ways of thinking. His task is made more difficult by the importance to his views

of the direct reference semantic program, which prevents him from treating names as synonymous with non-rigid definite descriptions.

Stalnaker's theory of intentionality involves rejecting an assumption of Frege's discussion, that the discovery that Hesperus is Phosphorus was exclusively astronomical, rather semantic. According to him, the discovery that Hesperus is Phosphorus, or the discovery that Water is H_2O, like the "discovery" that an ophthalmologist is an eye doctor, or the "discovery" that a fortnight is fourteen days. I have argued that the fact that we do not regard the latter cases as plausible discoveries is a cost of Stalnaker's program that is not justified by the causal-pragmatic account of intentionality. Furthermore, the apparatus described in "Assertion" is only applicable to a limited range of instances of Frege's puzzle, and it is unclear what Stalnaker could say about very similar cases.

However, Stalnaker presents us with a systematic account of the problem of intentionality. It is not easy to see how to separate parts of the project and evaluate them individually. The attractiveness of the parts lends credulity to the whole. Though I cannot in the end accept the concomitant replacement of mere epistemic possibility with semantic ignorance, it is not surprising that there are many who can.[23]

Bibliography

Braun, David (1998) "Understanding Belief Reports", *Philosophical Review* **105**: 555-595.
Carnap, Rudolf (1949) Meaning and Necessity (Chicago: University of Chicago Press).
Chalmers, David (2002) "On Sense and Intension", Philosophical Perspectives **16**.
Dummett, Michael (1973) Frege: Philosophy of Language (London: Duckworth).
Dummett, Michael (1991) The Logical Basis of Metaphysics (Cambridge: Harvard University Press).
Evans, Gareth (1985) Collected Papers (Oxford: Clarendon Press, 1985)

Evans, Gareth (1985a) "Reference and Contingency", in Evans (1985): 178-213.

Field, Hartry (2001) Truth and the Absence of Fact (Oxford: Oxford University Press: 2001): 30-67),

Field, Hartry (2001a) "Mental Representation", reprinted as Chapter 2 of Field (2001): 30-67.

Field, Hartry (2001b) "Stalnaker on Intentionality", reprinted as Chapter 3 of Field (2001).

Forbes, Graeme (1989) Languages of Possibility (Oxford: Basil Blackwell).

Jackson, Frank (1997) From Metaphysics to Ethics: A Defense of Conceptual Analysis (Oxford: Oxford University Press).

King, Jeffrey (2007) "What in the world are the ways things might have been"", Philosophical Studies 133: 443-453

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

Loewer, Barry (1987) "From Information to Intentionality", *Synthese* **70**: 287-317.

Loewer, Barry (1997) "A Guide to Naturalizing Semantics", in B. Hale and C. Wright, ed., A Companion to the Philosophy of Language (Oxford: Blackwell Press): 108-126.

Putnam, Hilary (1975) Mind, Language, and Reality: Philosophical Papers, Volume 2 (New York: Cambridge University Press).

Quine, W.V. (1976) The Ways of Paradox and Other Essays (Cambridge: Harvard University Press).

Quine, W.V. (1976a) "Carnap on Logical Truth", in Quine (1976a): 107-132.

Salmon, Nathan (1986) Frege's Puzzle (Cambridge: MIT Press)

Soames, Scott (2002) Beyond Rigidity (Oxford: Oxford University Press).

Soames, Scott (2005a) Reference and Description: The Case against Two-Dimensionalism (Princeton: Princeton University Press)

Soames, Scott (2005b) "Naming and Asserting" in Szabo (2005): 356-382.

Speaks, Jeff (2006) "Is Mental Content Prior to Linguistic Meaning?" *Nous* **40.3**: 428-467.

Stalnaker, Robert (1984) Inquiry (Cambridge: MIT Press)

Stalnaker, Robert (1999) Context and Content (Oxford: Oxford University Press).

Stalnaker, Robert (1999a) "Assertion" in Stalnaker (1999): 78-95.

Stalnaker, Robert (1999b) "The Problem of Logical Omniscience I", in Stalnaker (1999): 241-54.

Stalnaker, Robert (1999c) "The Problem of Logical Omniscience II", in Stalnaker (1999): 255-73.

Stalnaker, Robert (1999d) "Twin Earth Revisited", in Stalnaker (1999): 210-221.

Stalnaker, Robert (1999e) "Mental Content and Linguistic Form", in Stalnaker (1999): 225-240.

Stalnaker, Robert (1999f) "Belief Attribution and Context", in Stalnaker (1999): 150-66.

Stalnaker, Robert (1999g) "Semantics for Belief", in Stalnaker (1999): 117-129.

Stalnaker, Robert (2003) Ways a World Might Be (Oxford: Clarendon Press).

Stalnaker, Robert (2003a) "On Considering a Possible World as Actual", in Stalnaker (2003): 188-200.

Stalnaker, Robert (2006): Replies, in Content and Modality: Themes from the Philosophy of Robert Stalnaker (Oxford: Oxford University Press).

Stalnaker, Robert (2007) "Responses", *Philosophical Studies* **133**: 481-491.

Stanley, Jason (1997) "Rigidity and Content" in <u>Language, Thought, and Logic: Essays in Honor of Michael Dummett</u>, R. Heck, ed. (Oxford: Oxford University Press, 1997): 131-56.
Stanley, Jason (2002) "Modality and What is Said", in *Philosophical Perspectives* **16**, J. Tomberlin, ed. (Blackwell Press, 2002): 321-344.
Szabo, Zoltan (ed.) (2005) <u>Semantics versus Pragmatics</u> (Oxford: Oxford University Press).