Epistemic Modals in Context

Andy Egan John Hawthorne Brian Weatherson

2005

A very simple contextualist treatment of a sentence containing an epistemic modal, e.g. *a might be F*, is that it is true iff for all the contextually salient community knows, *a* is *F*. It is widely agreed that the simple theory will not work in some cases, but the counterexamples produced so far seem amenable to a more complicated contextualist theory. We argue, however, that no contextualist theory can capture the evaluations speakers naturally make of sentences containing epistemic modals. If we want to respect these evaluations, our best option is a *relativist* theory of epistemic modals. On a relativist theory, an utterance of *a might be F* can be true relative to one context of evaluation and false relative to another. We argue that such a theory does better than any rival approach at capturing all the behaviour of epistemic modals.

In the 1970s David Lewis argued for a contextualist treatment of modals (Lewis 1976, 1979). Although Lewis was primarily interested in modals connected with freedom and metaphysical possibility, his arguments for contextualism could easily be taken to support contextualism about epistemic modals. In the 1990s Keith DeRose argued for just that position (DeRose 1991, 1998).

In all contextualist treatments, the method by which the contextual variables get their values is not completely specified. For contextualist treatments of metaphysical modality, the important value is the class of salient worlds. For contextualist treatments of epistemic modality, the important value is which epistemic agents are salient. In this paper, we start by investigating how these values might be generated, and conclude that it is hard to come up with a plausible story about how they are generated. There are too many puzzle cases for a simple contextualist theory to be true, and a complicated contextualist story is apt to be implausibly ad hoc.

We then look at what happens if we replace contextualism with relativism. On contextualist theories the truth of an utterance type is relative to the context in which it is tokened. On relativist theories, the truth of an utterance token is relative to the context in which it is evaluated. Many of the puzzles for contextualism turn out to have natural, even elegant, solutions given relativism. We conclude by comparing two versions of relativism.

We begin with a puzzle about the role of epistemic modals in speech reports.

1 A Puzzle

The celebrity reporter looked discomforted, perhaps because there were so few celebrities in Cleveland.

"Myles", asked the anchor, "where are all the celebrities? Where is Professor Granger?"

"We don't know," replied Myles. "She might be in Prague. She was planning to travel there, and no one here knows whether she ended up there or whether she changed her plans at the last minute."

This amused Professor Granger, who always enjoyed seeing how badly wrong CNN reporters could be about her location. She wasn't sure exactly where in the South Pacific she was, but she was certain it wasn't Prague. On the other hand, it wasn't clear what Myles had gotten wrong. His first and third sentences surely seemed true: after all, he and the others certainly *didn't* know where Professor Granger was, and she *had* been planning to travel to Prague before quietly changing her destination to Bora Bora.

The sentence causing all the trouble seemed to be the second: "She might be in Prague." As she wiggled her toes in the warm sand and listened to the gentle rustling of the palm fronds in the salty breeze, at least one thing seemed clear: she definitely wasn't in Prague – so how could it be true that she might be? But the more she thought about it, the less certain she became. She mused as follows: when I say something like *x might be F*, I normally regard myself to be speaking truly if neither I nor any of my mates know that x is not *F*. And it's hard to believe that what goes for me does not go for this CNN reporter. I might be special in many ways, but I'm not semantically special. So it looks like Myles can truly say that I might be in Prague just in case neither *he* nor any of *his* mates knows that I am not. And I'm sure none of them knows that, because I've taken great pains to make them think that I am, in fact, in Prague – and reporters always fall for such deceptions.

But something about this reasoning rather confused Professor Granger, for she was sure Myles had gotten something *wrong*. No matter how nice that theoretical reasoning looked, the fact was that she definitely wasn't in Prague, and he said that she might be. Trying to put her finger on just where the mistake was, she ran through the following little argument.[1]

(1) When he says, "She might be in Prague" Myles says that I might be in Prague.

(2) When he says, "She might be in Prague" Myles speaks truly iff neither he nor any of his mates know that I'm not in Prague.

(3) Neither Myles nor any of his mates know that I'm not in Prague.

(4) If Myles speaks truly when he says that I might be in Prague, then I might be in Prague.

(5) I know I'm not in Prague.

(6) It's not the case that I know I'm not in Prague if I might be in Prague.

[1] Some of Professor Granger's thoughts sound a little odd being in the present tense, but as we shall see, there are complications concerning the interaction of tense with epistemic modals, so for now it is easier for us to avoid those interactions.

There must be a problem here somewhere, she thought – for (1) – (6) are jointly inconsistent. (Quick proof: (2) and (3) entail that Myles speaks truly when he says, "She might be in Prague". From that and (1) it follows he speaks truly when he says Professor Granger might be in Prague. From that and (4) it follows that Professor Granger might be in Prague. And that combined with (5) is obviously inconsistent with (6).) But wherein lies the fault? Unless some fairly radical kind of scepticism is true, Professor Granger can know by observing her South Pacific idyll that she's not in Prague – so (5) looks secure. And it seems pretty clear that neither Myles nor any of his mates know that she's not in Prague, since they all have very good reason to think that she is – so it looks like (3) is also OK. But the other four premises are all up for grabs.

Which exactly is the culprit is a difficult matter to settle. While the semantic theory underlying the reasoning in (1)-(6) is mistaken in its details, something like it is very plausible. The modal 'might' here is, most theorists agree, an *epistemic* modal. So its truth-value should depend on what someone knows. But who is this *someone*? If it is Myles, or the people around him, then the statement "she might be in Prague" is true, and it is unclear where to block the paradox. If it is Professor Granger, or the people around her, then the statement is false, but now it is unclear why a competent speaker would ever use this kind of epistemic modal. Assuming the *someone* is Professor Granger, and assuming Professor Granger knows where she is, then "Granger might be in Prague" will be true iff "Granger is in Prague" is true. But this seems to be a mistake. Saying "Granger *might* be in Prague" is a way to weaken one's commitments, which it could not be if the two sentences have the same truth conditions under plausible assumptions. So neither option looks particularly promising.

To make the problem even more pressing, consider what happens if a friend of Professor Granger's who knows she is in the South Pacific overhears Myles's comment. Call this third party Charles. It is prima facie very implausible that when Myles says that Professor Granger might be in Prague he means to rule out that *Charles* knows that she is not. After all, Charles is not part of the conversation, and Myles need not even know that he exists. So if Myles knows what he is saying, what he is saying could be true even if Charles knows Professor Granger is not in Prague. But if Charles knows this, Charles cannot regard Myles's statement as true, else he will conclude that Professor Granger might be in Prague, and he knows she is not. So things are very complicated indeed.

In reasoning as we have been, we have been assuming that the following inferences are valid.

(7) A competent English speaker says *It might be that S*; and
(8) *S*, on that occasion of use, means that *p*; entail
(9) That speaker says that it might be that *p*.

Further, (9) plus

(10) That speaker speaks truly; entail
(11) It might be that *p*.

If Charles accepts the validity of both of these inferences, then he is under considerable pressure to deny that Myles speaks truly. And it would be quite natural for him to do so – for instance, by interrupting Myles to say that "That's wrong. Granger couldn't be in Prague, since he left on the midnight flight to Tahiti." But it's very hard to find a plausible semantic theory that backs up this intervention, although such reactions are extremely common. (To solidify intuitions, here is another example: I overhear you say that a certain horse might have won a particular race. I happen to know that the horse is lame. I think: you are wrong to think that it might have won.)[2]

Our solutions to this puzzle consist in proposed semantic theories for epistemic modals. We start with contextualist solutions, look briefly at invariantist solutions, and conclude with relativist solutions. Although we will look primarily at the costs and benefits of these theories with respect to intuitions about epistemic modals, it is worth remembering that they differ radically in their presuppositions about what kind of theory a semantic theory should be. Solving the puzzles to do with epistemic modals may require settling some of the deepest issues in philosophy of language

2 Contextualist Solutions

In his (1991), Keith DeRose offers the following proposal:

> S's assertion "It is possible that P" is true if and only if (1) no member of the relevant community knows that P is false, and (2) there is no relevant way by which members of the relevant community can come to know that P is false. (593-4)

DeRose intends 'possible' here to be an epistemic modal, and the proposal is meant to cover all epistemic modals, including those using 'might'.[3] We will not discuss here the

[2]Note that it also seems implausible to say that this is an instance of *metalinguistic* negation, as discussed in Horn (1989). When Charles interrupts Myles to object, the objection isn't that the particular form of words that Myles has chosen is inappropriate. The form of words is fine, and Myles' utterance would be completely unobjectionable if Charles's epistemic state were slightly different. What's wrong is that Myles has used a perfectly acceptable form of words to say something that's false (at least by Charles' lights—more on this later). We also think it's implausible to understand the 'might' claims in question here as claims of objective chance or objective danger.

[3]We take the puzzle to be a puzzle about sentences containing epistemic modal operators, however they are identified. We are sympathetic with DeRose's (1998) position that many sentences containing 'might' and 'possible' are unambiguously epistemic, but do not wish to argue for that here. Rather, we simply take for granted that a class of sentences containing epistemic modal operators has been antecedently identified.

There are two differences between 'possible' and 'might'. The first seems fairly superficial. Sentences where *might* explicitly takes a sentence, rather than a predicate, as its argument are awkward at best, and may be ungrammatical. *It is possible that Professor Granger is in Prague* is much more natural than *It might be the case that Professor Granger is in Prague*, but there is no felt asymmetry between *Professor Granger is possibly in Prague* and *Professor Granger might be in Prague*. We will mostly ignore these issues here, and follow philosophical orthodoxy in treating epistemic modals as being primarily sentence modifiers rather than predicate modifiers. The syntactic features of epistemic modals are obviously important, but we're fairly confident that the assumption that epistemic modals primarily operate on sentences does not bear any theoretical load here, and could be replaced if necessary.

issues that arise under clause (2) of DeRose's account, since we'll have quite enough to consider just looking at whether clause (1) or anything like it is correct.[4]

In our discussion below, we consider three promising versions of contextualist theory. What makes the theories contextualist is that they all say that Myles spoke truly when he said "She might be in Prague", but hold that if Professor Granger had repeated his words she would have said something false.[5] And the reason for the variation in truth-value is just that Myles and Professor Granger are in different contexts, which supply different relevant communities. Where the three theories differ is in which constraints they place on how context can supply the community in question.

The first is the kind of theory that DeRose originally proposed. On this theory, there is a side constraint that the relevant community always includes the speaker: whenever S truly utters *a might be F*, S does not know that *a* is not *F*. We'll call this *the speaker-inclusion constraint*, or sometimes just *speaker-inclusion*. There is some quite compelling evidence for speaker-inclusion. Consider, for example, the following sort of case: Whenever Jack eats pepperoni pizza, he forgets that he has ten fingers, and thinks "I might only have eight fingers." Jill (who knows full well that Jack has *ten* fingers) spots Jack sitting all alone finishing off a pepperoni pizza, and says, "He might have eight fingers." Jill has said something false. And what she's said is false because it's not compatible with what *she* knows that Jack has eight fingers. But if the relevant community could ever exclude the speaker, one would think it could do so here. After all, Jack is clearly contextually salient: he's the referent of 'he,' the fingers in question are on his hand, and no one else is around.[6] Now, a single case does not prove a universal[7] – but the case does seem to provide good *prima facie* vidence for DeRose's constraint.

The other difference will be relevant to some arguments that follow. 'Might' can interact with tense operators in a way that 'possible' does not. *It might have rained* could either mean MIGHT (WAS it rains) or WAS (MIGHT it rains), while *It possibly rained* unambiguously means POSSIBLY (WAS it rains). It is often hard in English to tell just which meaning is meant when a sentence contains both tense operators and epistemic modals, but in Spanish these are expressed differently: Puede haber llovido; Podría haber llovido.

[4] There are three kinds of cases where something like DeRose's clause (2) could be relevant.

First, Jack and Jill are in a conversation, and Jack knows p while Jill knows $p \rightarrow \neg Fa$. In this case intuitively neither could truly say *a might be F* even though neither knows *a* is not *F*.

Second, there are infinitely many mathematicians discussing Fermat's Last Theorem. The first knows just that it has no solutions for n=3, the second just that it has no solutions for n=4, and so on. Intuitions are (unsurprisingly) weaker here, but we think none of them could say *Fermat's Last Theorem might have solutions*, because the group's knowledge rules this out.

Third, if *S* was very recently told that *a* is not *F*, but simply forgot this, then intuitively she speaks falsely if she says *a might be F*.

Fourth, if S has the materials for easily coming to know P from her current knowledge, but has not performed the relevant inference, then we might be inclined (depending on how easy the inferential steps were to see and so on) to say that she is wrong to utter 'It might be that not P'.

Rather than try and resolve the issues these cases raise, we will stick to cases where the only thing that could make *a might be F* false is that someone knows that *a* is not *F*.

[5] She would also have violated some pragmatic principles by knowingly using a third-person pronoun to refer to herself, but we take it those principles are defeasible, and violation of them does not threaten the truth-aptness of her utterance.

[6] Notice that intuitions do not change if we alter the case in such a way that Jack has a strange disorder that makes it very hard for him to come to know how many fingers he has. Thus clause (2) of Derose's analysis cannot do the work of the relevant side constraint.

[7] And see the case of Tom and Sally in the maze below for some countervailing evidence.

One implication of DeRose's theory is that (1) is false, at least when Professor Granger says it. For when Professor Granger reports that Myles says "She might be in Prague," she is reporting a claim he makes about his epistemic community – that her being in Prague is compatible with the things that they know. But when she says (in the second clause) that this means he is saying that she might be in Prague, she speaks falsely. For in her mouth the phrase "that I might be in Prague" denotes the proposition that it's compatible with the knowledge of an epistemic community that includes Professor Granger (as the speaker) that Professor Granger is in Prague. And that is not a proposition that Myles assented to. So DeRose's theory implies that the very intuitive (1) is false when uttered by Granger.

(1) When he says, "She might be in Prague" Myles says that I might be in Prague.

It is worth emphasizing how counterintuitive this consequence of speaker-inclusion is. If the speaker-inclusion constraint holds universally then in general speech involving epistemic modals cannot be reported disquotationally. But notice how natural it is, when telling the story of Jack and Jill, to describe the situation (as we ourselves did in an earlier draft of this paper) as being one where "Whenever Jack eats pepperoni pizza, he forgets that he has ten fingers, and thinks he might only have eight." Indeed, it is an important generalization about how we use language that speakers usually do not hesitate to disquote in reporting speeches using epistemic modals. So much so that exceptions to this general principle are striking – as when the tenses of the original speech and the report do not match up, and the tense difference matters to the plausibility of the attribution.

One might try to explain away the data just presented by maintaining a laxity for 'says that' reports. A chemist might say 'The bottle is empty' meaning it is empty of air, while milkman might utter the same sentence, meaning in my context that it is empty of milk. Nevertheless, the milkman might be slightly ambivalent about denying:

> When the chemist says 'The bottle is empty', she says that the bottle is empty.

And this is no doubt because the overt 'says that' construction frequently deploys adjectives and verbs in a rather quotational way. After all, the chemist could get away with the following speech in ordinary discourse: "I know the milkman said that the bottle is empty. But he didn't mean what I meant when I said that the bottle is empty. When he said that the bottle was empty he meant that it was empty of milk."[8] Thus the conventions of philosophers for using 'say that' involve regimenting ordinary use in a certain direction.[9] But the disquotational facts that we are interested in cannot be explained away simply by invoking these peculiarities of 'says that' constructions, for the same disquotational ease surrounds the relevant belief reports. In the case just considered, while we might argue about whether it was acceptable for the chemist to say, in

[8] Notice that this use prohibits the inference from: The speaker said that the bottle was empty, to, The speaker expressed the proposition/said something that meant that the bottle was empty.

[9] We are grateful for correspondence with John MacFarlane here.

her conversational context, "The milkman said that the bottle was empty", it is manifestly unacceptable for her to say "The milkman believes that the bottle is empty". This contrasts with the case of 'might': If someone asked Professor Granger where Myles thought she was, she could quite properly have replied with (12).

(12) He thinks that/believes that I might be in Prague.

Indeed, we in general tend find the following inference pattern – a belief-theoretic version of (7) to (9) above – compelling:

1. A competent English speaker sincerely asserts *It might be that S*
2. S, in that context of use, means that *p*.; therefore,
3. That speaker believes that it might be that *p*

Our puzzle cannot, then, be traced simply to a laxity in the 'says that' construction.[10] Whatever the puzzle comes to, it certainly runs deeper than that.

Notice that (12) does not suggest that Myles thinks that for all Professor Granger knows, she is in Prague; it expresses the thought that Myles thinks that for all he knows, that is where she is. Moreover, this is hardly a case where Granger's utterance is of doubtful appropriateness: (12) is one of the ways canonically available for Granger to express that thought. But if we assume that what is reported in a belief report of this kind is belief in the proposition the reporter expresses by *I might be in Prague*, and we assume a broad-reaching speaker-inclusion constraint, we must concede that the proposition Granger expresses by uttering (12) is that Myles believes that for all *Professor Granger* knows, Professor Granger is in Prague.

If the speaker-inclusion constraint holds universally, then anyone making such a report is wrong. There are two ways for this to happen—either they know what the sentences they're using to make the attributions mean, and they have radically false views about what other people believe, or they have non-crazy views about what people believe, but they're wrong about the meanings of the sentences they're using. The first option is incredibly implausible. So our first contextualist theory needs to postulate a widespread semantic blindness; in general speakers making reports are mistaken about the semantics of their own language. In particular, it requires that such speakers are often blind to semantic differences between sentence tokens involving epistemic modals. It is possible that some theories that require semantic blindness are true, but other things being equal we would prefer theories that do not assume this.[11] In general the burden of proof is on those who think that the folk don't know the meaning of their own words. More carefully: the burden of proof is on those who think that the folk are severely handicapped in their ability to discriminate semantic sameness and difference in their home language.

[10] For what its worth, we also note that 'S claimed that P' has less laxity (of the sort being discussed) than 'S said that P'.

[11] Note that the negation of semantic blindness concerning some fragment of the language is *not* the theory that speakers know all the semantic equivalences that hold between terms in that fragment. All we mean by the denial of semantic blindness is that speakers not have false beliefs about the semantics of their terms.

So the plausibility of (1) counts as evidence against the first contextualist theory, and provides a suggestion for our second contextualist theory. The cases that provide the best intuitive support for the speaker-inclusion constraint and the case we used above, involved unembedded epistemic modals. Perhaps this constraint is true for epistemic modals in simple sentences, but not for epistemic modals in 'that' clauses. Perhaps, that is, when S sincerely asserts *X Vs that a might be F*, she believes that X Vs that for all X (and her community) knows, *a* is *F*. (This is not meant as an account of the logical form of *X Vs that a might be F*, just an account of its truth conditions. We defer consideration of what hypothesis, if any, about the underlying syntax could generate those truth conditions.) To motivate this hypothesis, note how we introduced poor Jack, above. We said that he thinks he might have eight fingers. We certainly didn't mean by that that Jack thinks something about *our* epistemic state.

The other problem with the speaker-inclusion constraint is that it does not seem to hold when epistemic modals are bound by temporal modifiers, as in the following example. A military instructor is telling his troops about how to prepare for jungle warfare. He says, "Before you walk into an area where there are lots of high trees, if there might be snipers hiding in the branches, clear away the foliage with flamethrowers." Whatever the military and environmental merits of this tactic, the suggestion is clear. The military instructor is giving generic conditional advice: in any situation of type *S*, if *C* then do *A*. The situation *S* is easy to understand, it is when the troops are advancing into areas where there are high trees. And A, too, is clear: blaze 'em. But what about *C*? What does it mean to say that there might be snipers in the high branches? Surely not that it's compatible with the military instructor's knowledge that there are snipers in the high branches – he's sitting happily in West Point, watching boats sail lazily along the Hudson. What *he* thinks about where the snipers are is neither here nor there. Intuitively, what he meant was that the troops should use flamethrowers if *they* don't know whether there are snipers in the high branches. (Or if they know that there *are*.) So as well as leading to implausible claims about speech reports, the speaker-inclusion constraint seems clearly false when we consider temporal modifiers.

Here is a way to deal with both problems at once. There are constraints on the application of the speaker-inclusion constraint. It does not apply when the epistemic modal is in the scope of a temporal modifier (as the flamethrower example shows) and it does not apply when the epistemic modal is in a 'that' clause.[12] Our second contextualist theory then accepts the speaker-inclusion constraint, but puts constraints on its appli-

[12] This theory looks like one in which propositional attitude operators become monsters, since the content of *Jack thinks that Jill might be happy* is naturally generated by applying the operator *Jack thinks* to the proposition that *that Jill might be happy* denotes when it is expressed in Jack's context. But this is not the easiest, or obviously the best, way to look at the theory. For one thing, that way of looking at things threatens to assign the wrong content to *Jack thinks that Jill might have stolen my car*. The content of *Jill might have stolen my car* in Jack's context is that for all Jack knows, Jill stole Jack's car, which is not what is intended. That is to say, thinking of propositional attitude operators as monsters here ignores the special status of epistemic modals in the semantics. It is better, we think, to hold that on this theory epistemic modals are impure indexicals whose value is fixed, inter alia, by their location in the sentence as well as their location in the world. But even if this theory does not *officially* have monsters, the similarity to monstrous theories is worth bearing in mind as one considers the pros and cons of the theory.

Thanks to Ernest Lepore for helpful discussions here.

cation.

This kind of theory, with a speaker-inclusion constraint only applying to relatively simple epistemic modals, allows us to accept (1). The problematic claim on this theory turns out to be (4):

(4) If Myles speaks truly when he says that I might be in Prague, then I might be in Prague.

When Myles said that Professor Granger might be in Prague, he was speaking truly. That utterance expressed a true proposition. So the antecedent of (4) is true. But the consequent is false: the "might" that appears there is not in a that-clause or in the scope of a temporal modifier; so the speaker-inclusion constraint requires that Professor Granger be included in the relevant community; and since she knows that she is not in Prague, it's not true that she might be. We would similarly have to reject:

(4') If Myles has a true belief that I might be in Prague, then I might be in Prague.

But there are reasons to be worried about this version of contextualism, beyond the uneasiness that attaches to denying (4), and, worse still, (4'). For one, this particular version of the speaker-inclusion constraint seems a bit *ad hoc*: why should there be just *these* restrictions on the relevant community? More importantly, the theory indicts certain inferential patterns that are intuitively valid. Suppose a bystander in our original example reasoned[13]:

(13) Myles believes that it might be that Professor Granger is in Prague.
(14) Myles's belief is true; therefore,
(15) It might be that Professor Granger is in Prague.

But this version of contextualism tells us that while (13) and (14) are true, (15) is false. In general, there are going to be counter-intuitive results whenever we reason from cases where the speaker-inclusion constraint does not apply to cases where it does.

Finally, the theory is unable to deal with certain sorts of puzzle cases. The first kind of case directly challenges the speaker-inclusion constraint for simple sentences, although we are a little sceptical about how much such a case shows.[14] Tom is stuck in a maze. Sally knows the way out, and knows she knows this, but doesn't want to tell Tom. Tom asks whether the exit is to the left. Sally says, "It might be. It might not be." Sally might be being unhelpful here, but it isn't clear that she is *lying*. Yet if the speaker-inclusion constraint applies to unembedded epistemic modals, then Sally is clearly saying something that she knows to be false, for she knows that she knows which way is out.

This case is not altogether convincing, for there is something slightly awkward about Sally's speech here. For example, if Sally knows the exit is not to the left, then even if she is prepared to utter, "It might be [to the left]," she will not normally self-ascribe knowledge that it might be to the left. And normally speakers don't sincerely assert

[13] What follows is a belief theoretic version of Charles' reasoning.
[14] A similar case to the following appears in (Hawthorne 2004, 27).

things they don't take themselves to know. So it is natural to suppose that a kind of pretense or projection is going on in Sally's speech that may well place it beyond the purview of the core semantic theory.

The following case makes more trouble for our second contextualist theory, though it too has complications. Ann is planning a surprise party for Bill. Unfortunately, Chris has discovered the surprise and told Bill all about it. Now Bill and Chris are having fun watching Ann try to set up the party without being discovered. Currently Ann is walking past Chris's apartment carrying a large supply of party hats. She sees a bus on which Bill frequently rides home, so she jumps into some nearby bushes to avoid being spotted. Bill, watching from Chris's window, is quite amused, but Chris is puzzled and asks Bill why Ann is hiding in the bushes. Bill says

(16) I might be on that bus.

It seems Bill has, somehow, conveyed the *correct* explanation for Ann's dive—he's said something that's both true and explanatory. But in his mouth, according to either contextualist theory we have considered, it is not true (and so it can't be explanatory) that he might have been on the bus. He knows that he is in Chris's apartment, which is not inside the bus.

Chris's question, like most questions asking for an explanation of an action, was ambiguous. Chris might have been asking what *motivated* Ann to hide in the bushes, or he might have been asking what *justified* her hiding in the bushes. This ambiguity is often harmless, because the same answer can be given for each. This looks to be just such a case. Bill seems to provide both a motivation and a justification for Ann's leap by uttering (16). That point somewhat undercuts a natural explanation of what's going on in (16). One might think that what he said was elliptical for *She believed that I might be on the bus*. And on our second contextualist theory, that will be *true*. If Bill took himself to be answering a question about motivation, that might be a natural analysis. (Though there's the underlying problem that Ann presumably wasn't thinking about her mental states when she made the leap. She was thinking about the bus, and whether Bill would be on it.) But that analysis is less natural if we think that Bill was providing a justification of Ann's actions.[15] And it seems plausible that he could utter (16) in the course of providing such a justification. This suggests that (16) simply means that for all *Ann* knew, Bill was on that bus. Alternatively, we could say that (16) is elliptical for *Because I might be on that bus*, and that the speaker-inclusion constraint does not apply to an epistemic modal connected to another sentence by 'because'. This may be right, but by this stage we imagine some will be thinking that the project of trying to find all

[15] Though the theory will allow for the truth of, "I might *have been* on that bus" (since the epistemic modal clause doesn't occur on its own, but in the scope of a temporal operator). So if we think that (i) that's enough to do the justificatory and explanatory work, and (b) Bill's utterance of "I might be on that bus" is best understood as a clumsy stab at "I might have been on that bus", then perhaps we can account for this kind of case using our second contextualist theory. Two worries: First, it is a cost of the theory that we have to reinterpret Bill's utterance in this way, as a clumsy attempt to say something that the theory can accommodate. Second, there might be cases where the interpretation is less plausible: As a response to, "Why is Ann getting ready to jump over the hedge?", "I might have been on that bus" sounds worse to us than "I might be on that bus".

the restrictions on the speaker-inclusion constraint is a degenerating research program, and a paradigm shift may be in order.

So our final contextualist theory is that DeRose's original semantic theory, before the addition of any sort of speaker-inclusion constraint, was correct and complete. So 'might' behaves like 'local' and 'nearby'. If Susie says "There are snipers nearby," the truth condition for that might be that there are snipers near Susie, or that there are snipers near us, or that there are snipers near some other contextually salient individual or group. Similarly, if she utters "Professor Granger might be in Prague" the truth condition for that might be that for all she knows Professor Granger is in Prague, or that for all we know Professor Granger is in Prague, or that for all some other community knows, Professor Granger is in Prague. There are no universal rules requiring or preventing the speaker from being included in the class of salient epistemic agents.

According to the third version of contextualism, if Professor Granger does not equivocate when working through her paradox, then the problem lies with (6):

(6) It's not the case that I can know I'm not in Prague if I might be in Prague.

At the start of her reasoning process, Professor Granger's use of 'might' means (roughly) 'is compatible with what Myles and his friends know'. And if it keeps that meaning to the end, then the antecedent of (6) is true, because Professor Granger might (in that sense) be in Prague, even though she knows she is not. Any attempt to show that (1) through (6) form an inconsistent set will commit a fallacy of equivocation.[16]

But (6) as uttered by Professor Granger sounds extremely plausible. And there are other, more general problems as well. It is difficult on such a theory to explain why it is so hard to get the relevant community to exclude the speaker in present tense cases: Why, for instance, can't Jill's statement about Jack, "He might have eight fingers," be a statement about Jack's epistemic state rather than her own? The third theory offers us no guidance.[17]

[16] The same kind of equivocation can be seen in other arguments involving contextually variable terms. Assume that Nomar lives in Boston, Derek lives in New York, and Nomar, while talking about Fenway Park in Boston says, "I live nearby." Derek, at home in New York, hears this on television and runs through the following argument.

1. In saying "I live nearby" Nomar says that he lives nearby. (Plausible disquotational premise about 'nearby')
2. Nomar speaks truly when he says "I live nearby" (Follows from the setup)
3. If Nomar speaks truly when he says "I live nearby" and in saying "I live nearby" he says that he lives nearby, then he lives nearby. (I.e. if he speaks truly then what he says is true.)
4. If Nomar lives nearby, then he lives in New York (Since everywhere that's nearby to Derek's home is in New York.); therefore
5. Nomar lives in New York.

The right thing to say about this argument is that it equivocates. Every premise has a true reading. Perhaps every premise is true on its most natural reading, but the denotation of 'nearby' has to change throughout the argument for every premise to be true. The current view is that 'might' behaves like 'nearby', and that Professor Granger's argument equivocates, like Derek's.

[17] There also seems to be a past/future asymmetry about epistemic modals which the third contextualist theory will have trouble explaining. Consider this case involving past tense epistemic modals. Romeo sees Juliet carrying an umbrella home on a sunny afternoon. When he asks her why she is carrying an umbrella, she replies "It might have rained today." There's a scope ambiguity in Juliet's utterance. If the epistemic

We'll close this section with a discussion of the interaction between syntax and semantics in these contextualist theories. As is well known, in the last decade many different contextualist theories have been proposed for various philosophically interesting terms. Jason Stanley (2000) has argued that the following two constraints should put limits on when we posit contextualist semantic theories.

Variable Any contextual effect on truth-conditions that is not traceable to an indexical, pronoun, or demonstrative in the narrow sense must be traceable to a structural position occupied by a variable. (Stanley 2000, 401)[18]

Syntactic Evidence The only good evidence for the existence of a variable in the semantic structure corresponding to a linguistic string is that the string, or another that we have reason to believe is syntactically like it, has interpretations that could only be accounted for by the presence of such a variable.

---

modal takes wide scope with respect to the tense operator, Juliet would be claiming that she doesn't know whether it has rained today (implicating, oddly, that this is why she now has an umbrella.) Or, as Juliet presumably intends, the temporal operator could take wide scope with respect to the epistemic modal. In that case Juliet says that it was the case at some earlier time (presumably when she left for work this morning) that it was compatible with her knowledge that it would rain today. And that seems both true and a good explanation of her umbrella-carrying.

It is much harder, if it is even possible, to find cases involving future tense operators where the temporal operator takes wide scope with respect to the pistemic modal. If S says, "It might rain tomorrow", that seems to unambiguously mean that it's compatible with S's *current* knowledge (and her community's) that it rains tomorrow. For a more dramatic case, consider a case where two people, Othello and Desdemona, have discovered that a giant earthquake next week will destroy humanity. No one else knows this yet, but there's nothing that can be done about it. This rather depresses them, so they decide to take memory-wiping drugs so that when they wake up tomorrow, they won't know about the earthquake. Othello can't say, "Tomorrow, humanity might survive," even though it is true that tomorrow, for all anyone will know, humanity will survive. If the temporal modifier could take wide scope with respect to the epistemic modal, Othello's utterance could have a true reading. But it does not. It's possible at this point that our policy, announced in footnote 2, of ignoring issues relating to DeRose's second clause will come back to haunt us. One possibility here is that tomorrow it will still be false that humanity might survive because it's not compatible with what people tomorrow know and knew that humanity survives. We don't think that's what is going on, but it's possible. Here's two quick reasons to think that the problem is not so simple. First, if Othello and Desdemona commit suicide rather than take the memory-wiping drugs, it will be compatible tomorrow with all anyone ever knew that humanity survives. But still Othello's speech seems false. Second, it's not obviously right that what people ever knew matters for what is epistemically possible now. Presumably at one stage Bill Clinton knew what he had for lunch on April 20, 1973. (For example, when he was eating lunch on April 20, 1973.) But unless he keeps meticulous gastronomical records, this bit of knowledge is lost to humanity forever. So there will be true sentences of the form *Bill Clinton might have eaten x for lunch on April 20, 1973* even though someone once knew he did not. Now change the earthquake case so that it will happen in thirty years not a week, and no one will then know about it (because Othello and Desdemona took the memory-wiping drugs and destroyed the machines that could detect it). Still it won't be true if Othello says, "In thirty years, humanity might survive." This suggests to us that some kind of constraints on epistemic modals will be required. The existence of these constraints seems to refute the 'no constraints' version of contextualism. It also undermines the argument that the second version of contextualism is too ad hoc. Once some constraints are in place, others may be appropriate.

[18] We assume here, following Stanley, a 'traditional syntax involving variables' (Stanley 2000 fn. 13). At least one of us would prefer a variable-free semantics along the lines of Jacobson (1999) Adopting such a semantics would involve, as Stanley says, major revisions to the presentation of this argument, but would not *clearly* involve serious changes to the argument. Most contextualists happily accept the existence of variables so we do not beg any questions against them, but see Pagin (2005) for an important exception.

If any contextualist theory of epistemic modals is to be justifiably believed, then **Variable** and **Syntactic Evidence** together entail the existence of sentences where the 'relevant community' is bound by some higher operator. So ideally we would have sentences like (17) with interpretations like (18).

(17) Everyone might be at the party tonight.
(18) For all x, it is consistent with all x knows that x will be at the party tonight.

Now (17) cannot have this interpretation, which might look like bad news for the contextualist theory. It's natural to think that if 'might' includes a variable whose value is the relevant community, that variable could be bound by a quantifier ranging over it. But if such a binding were possible, it's natural to think that it would be manifested in (17). So **Variable** and **Syntactic Evidence** together entail that we ought not to endorse contextualism about epistemic modals.

This argument against contextualism fails in an interesting way, one that bears on the general question of what should count as evidence for or against a contextualist theory. The reason that any variable associated with 'might' in (17) cannot be bound by 'everyone' is that 'might' takes wider scope than 'everyone'. Note that (17) does not mean (19), but rather means (20).

(19) For all x, it is consistent with what we know that x will be at the party tonight.
(20) It is consistent with what we know that for all x, x will be at the party tonight.

As Kai von Fintel and Sabine Iatridou (2003) have shown, in any sentence of the form *Every F might be G*, the epistemic modal takes wide scope. For instance, (21) has no true reading if there is at most one winner of the election, even if there is no candidate that we know is going to lose.

(21) Every candidate might win.

More generally, epistemic modals take wide scope with respect to a wide class of quantifiers.[19] This fact is called the Epistemic Containment Principle by von Fintel and Iatridou. Even if there is a variable position for the relevant community in the lexical entry for 'might', this might be unbindable because the epistemic modal always scopes over a quantifier that could bind it. If that's true then the requirement imposed by **Syntactic Evidence** is too strong. If the evidence from binding is genuinely neutral between the hypothesis that this variable place exists and the hypothesis that it does not, because there are no instances of epistemic modals that take narrow scope with respect to quantifiers, it seems reasonable to conclude that there are these variable places on the basis of other evidence.

Having said all that, there still may be direct evidence for the existence of a variable position for relevant communities. Consider again our example of the military instructor, reprinted here as (22).

[19] It is not entirely clear what the relevant class of quantifiers is, although von Fintel and Iatridou have some intriguing suggestions about what it might be.

(22) Before you walk into an area where there are lots of high trees, if there might be snipers hiding in the branches use your flamethrowers to clear away the foliage.

As von Fintel and Iatridou note, it is possible for epistemic modals to take narrow scope with respect to generic quantifiers. That's exactly what happens in (22). And it seems that the best interpretation of (22) requires a variable attached to 'might'. Intuitively, (22) means something like (23).

(23) Generally in situations where you are walking into an area where there are lots of high trees, if it's consistent with *your party*'s knowledge that there are snipers hiding in the branches use your flamethrowers to clear away the foliage.

The italicised *your party* seems to be the semantic contribution of the unenunciated variable. We are *not* saying that the existence of sentences like (23) shows that there are such variables in the logical form of sentences involving epistemic modals.[20] We just want to make two points here. First, if you are a partisan of **Syntactic Evidence**, then (22) should convince you not to object to semantic accounts of epistemic modals that appeal to variables, as our contextualist theories do. Second, we note a general concern that principles like **Syntactic Evidence** presupposes that a certain kind of construction, where the contextually variable term is bound at a level like LF, is always possible. Since there are rinciples like the Epistemic Containment Principle, we note a mild concern that this presupposition will not always be satisfied.

3 Invariantist Solutions

The most plausible form of invariantism about epistemic modals is that DeRose's semantics is broadly correct, but the relevant community is not set by context - it is invariably the world. We will call this position *universalism*. Of course when we say *a might be F* we don't normally communicate the proposition that no one in the world knows whether *a* is *F*. The analogy here is to pragmatic theories of quantifier domain restriction, according to which when we say *Everyone is F*, we don't communicate the proposition that everyone in the world is *F*, even though that is the truth condition for our utterance.

The universalist position denies (2) in Professor Granger's argument. Myles did not speak truly when he said "Professor Granger might be in Prague" because someone, namely Professor Granger, knew she was not in Prague. Although (2) is fairly plausible, it probably has weaker intuitive support than the other claims, so this is a virtue of the universalist theory.

The big advantage (besides its simplicity) of the universalist theory is that it explains some puzzle cases involving eavesdropping. Consider the following kind of case. Holmes and Watson are using a primitive bug to listen in on Moriarty's discussions

[20] As previously noted, we are not all convinced that semantics *ever* needs to appeal to such variables, let alone that it does to account for the behaviour of epistemic modals.

with his underlings as he struggles to avoid Holmes's plan to trap him. Moriarty says to his assistant,

(24) Holmes might have gone to Paris to search for me.

Holmes and Watson are sitting in Baker Street listening to this. Watson, rather inexplicably, says "That's right" on hearing Moriarty uttering (24). Holmes is quite perplexed. Surely Watson knows that he is sitting right here, in Baker Street, which is definitely not in Paris. But Watson's ignorance is semantic, not geographic. He was reasoning as follows. For all Moriarty (and his friends) know, Holmes is in Paris searching for him. If some kind of contextualism is true, then it seems that (24) is true in Moriarty's mouth. And, thought Watson, if someone says something true, it's OK to say "That's right."

Watson's conclusion is clearly wrong. It's not OK for him to say "That's right," in response to Moriarty saying (24). So his reasoning must fail somewhere. The universalist says that where the reasoning fails is in saying the relevant community only contains Moriarty's gang members. If we include Holmes and Watson, as the universalist requires, then Moriarty speaks falsely when he says (24).

There are a number of serious (and fairly obvious) problems with the universalist account. According to universalism, the following three claims are inconsistent.

(25) x might be F.
(26) x might not be F.
(27) Someone knows whether x is F.

Since these don't *look* inconsistent, universalism looks to be false.

The universalist's move here has to be to appeal to the pragmatics. If (27) is true then one of (25) and (26) is false, although both might be appropriate to express in some contexts. But if we can appropriately utter sentences expressing false propositions in some contexts, then presumably we can inappropriately utter true sentences in other contexts. (Indeed, the latter possibility seems much more common.) So one could respond to the universalist's main argument, their analysis of eavesdropping cases like Watson's, by accepting that Watson can't *appropriately* say "That's right" but he can *truly* say this. The universalist will have a hard time explaining why such a theory cannot work, assuming, of course, that she can explain how her own pragmatic theory can explain all the data.

The major problem here is one common to all appeals to radical pragmatics in order to defend semantic theories. If universalism is true then speakers regularly, and properly, express propositions they know to be false.[21] (We assume here that radical scepticism is *not* true, so sometimes people know some things.) Myles knows full well than *someone* knows whether Professor Granger is in Prague, namely Professor Granger. But if he's a normal English speaker, this will *not* seem like a reason for him to not say, "Professor Granger might be in Prague." Some might not think this is a deep problem for

[21] By "express" we will always mean "semantically express". We're not concerned with, and hope not to commit ourselves to any views about, for example, what's conveyed via various pragmatic processes.

the universalist theory, for speakers can be mistaken in their semantic views in ever so many ways. But many ill regard it as a serious cost of the universalist claim.

This problem becomes more pressing when we look at what universalism says about beliefs involving epistemic modals. Myles does not just say that Professor Granger might be in Prague, he believes it. And he believes Professor Granger might not be in Prague. If he also believes that Professor Granger knows where she is, these beliefs are inconsistent given universalism. Perhaps the universalist can once again invoke pragmatics. It is not literally true in the story that Myles believes that Granger might be in Prague. But in escribing the situation we use "Myles believes that Granger might be in Prague," to pragmatically communicate truths by a literal falsehood. This appeal to a pragmatic escape route seems even more strained than the previous universalist claims.

In general, the universalism under discussion here seems to run up against a constraint on semantic theorising imposed by Kripke's Weak Disquotation Principle. The principle says that if a speaker sincerely accepts a sentence, then she believes its semantic value.[22] If we have some independent information about what a speaker believes, then we can draw certain conclusions about the content of the sentences she accepts, in particular that she only accepts sentences whose content she believes. The universalist now has two options.[23] First, she can say that Myles here does accept inconsistent propositions. Second, she can deny the Weak Disquotation Principle, and say that although Myles sincerely asserts, and accepts, "Professor Granger might be in Prague" he doesn't really believe that Professor Granger might be in Prague. Generally, it's good to have options. But it's bad to have options as unappealing as these.[24]

4 Reporting Epistemic Modals

Our third class of solutions will be relatively radical, so it's worth pausing to look at the evidence for it. Consider again the dialogue between Moriarty, Holmes and Watson. Moriarty, recall, utters (24)

[22] Note that something like this had better be true if what it is to believe p is to have a sentence that means p in one's 'belief box'.

[23] We assume that it is not a serious option to deny that we ever accept unnegated epistemic modal sentences.

[24] There also a technical problem with universalism that mirrors one of the problems Stanley and Szabó (2000) raise for pragmatic theories of quantifier domain restriction. Normally (28) would be used to express a proposition like (29).

(28) Every professor enjoys every class.
(29) Every salient professor enjoys every class that s/he teaches.

Intuitively, by uttering (28) we express a proposition that contains two restricted quantifiers. Let's accept, for the sake of the argument, that a pragmatic theory of quantifier domain restriction can sometimes explain why the quantifiers in the propositions we express are more restricted than the quantifiers in the truth conditions for the sentences we use. Stanley and Szabó argue that such an explanation will not generalise to cover embedded quantifiers where the quantifier domain in the proposition expressed is bound to the outer quantifier. One such quantifier is the quantifier over classes in (28). We will not repeat their arguments here, but simply note that if they are correct, the universalist faces a problem in explaining how we use sentences with embedded epistemic modals that are (intuitively) defined with respect to a community that is bound by a higher level quantifier. As we saw, (22) provides an example of this kind of epistemic modal.

(24) Holmes might have gone to Paris to search for me.

Watson knows that Holmes is in Baker Street, as of course does Holmes. In the above case we imagined that both Watson and Holmes heard Moriarty say this. Change the story a little so Holmes does not hear Moriarty speak, instead when he comes back into the room he asks Watson what Moriarty thinks. Watson, quite properly, replies with (30).

(30) He thinks that you might have gone to Paris to search for him.

This is clearly not direct quotation because Watson changes the pronouns in Moriarty's statement. It is not as if Watson said "He sincerely said, 'Holmes might have gone to Paris to search for me.' " This might have been appropriate if Holmes suspected Moriarty was speaking in code so the proposition he expressed was very sensitive to the words he used.

Nor was Watson's quote a 'mixed' quote, in the sense of what happens in (31).[25] The background is that Arnold always uses the phrase 'my little friend' to denote his Hummer H2, despite that vehicle being neither little nor friendly. No one else, however, approves of this terminology.[26]

(31) Arnold: My little friend could drive up Mt Everest.
Chaz: Arnold believes his little friend could drive up Mt Everest.

We've left off the punctuation here so as to not beg any questions, but there is a way this could be an acceptable report if the fourth and fifth word, and those two words only, are part of a quotation. This is clearly not ordinary direct quotation, for Arnold did not think, in English or Mentalese, "His little friend could drive up Mt Everest." Nevertheless, this is not ordinary indirect quotation. In ordinary spoken English Chaz's report will be unacceptable unless 'little friend' is stressed. The stress here seems to be just the same stress as is used in metalinguistic negation, as described in Horn (1989). Note the length of the pause between 'his' and 'little'. With an ordinary pause it sounds as if Chaz is using, not mentioning, 'little friend'. So it is possible *in principle* to have belief reports, like this one, that are neither strictly direct nor strictly indirect.[27] Nevertheless, it does not seem like (30) need such a case. In particular, there need be no distinctive metalinguistic stress on 'might' in Watson's utterance of (30), and such stress seems to be mandatory for this mixed report.

[25] Earlier we used *speech* reports to illustrate the oddities of epistemic modals inside propositional attitude ascriptions. There are well-known difficulties with connecting appropriate speech reports to the semantic content of what is said, as opposed to merely communicated. (For some discussion of these, see Soames (2002) and Capellen and Lepore (1997).) We don't think those difficulties affect the above arguments, where the evidence is fairly clear, and fairly overwhelming. But matters get a little more delicate in what follows, so we move to belief reports because they are more closely tied to the content of what is believed.

[26] In this case, as with all the belief reports discussed below, the only evidence the reporter has for the report is given by the speech immediately preceding it. We assume there is good reason from the context to assume that the speakers are sincere.

[27] There are somewhat delicate questions about what a direct belief report *means*, but we assume the notion is well enough understood, even if we could not formally explicate what is going on in all such reports.

Assuming Moriarty was speaking ordinary English, Watson's report seems perfectly accurate. This is despite the fact that the relevant community one would naturally associate with Watson's use of 'might' is quite different to the community we would associate with Moriarty's use. When reporting speeches involving epistemic modals – and the beliefs express by sincere instances of such speeches, speakers can simply *disquote* the modal terms.

As is reasonably well known, there are many terms for which this kind of disquoting report is impermissible. In every case, Guildenstern's report of Ophelia's utterance is inappropriate.

(32) Ophelia: I love Hamlet.
...
Guildenstern: *Ophelia thinks that I love Hamlet.

(33) Guildenstern: What think you of Lord Hamlet?
Ophelia: He is a jerk.
...
Rosencrantz: What does Ophelia think of the King?
Guildenstern: *She thinks that he is a jerk.

(34) Guildenstern: Are you ready to teach the class on contextualism?
Ophelia: I'm ready.
...
Rosencrantz: Does Ophelia think she is ready to defend her dissertation?
Guildenstern: *She thinks she is ready.

(35) (Guildenstern and Ophelia are on the telephone, Guildenstern is in Miami, and Ophelia is in San Francisco)
Guildenstern: What do you like best about San Francisco?
Ophelia:There are lots of wineries nearby.
...
Rosencrantz: Is it possible to grow wine in south Florida?
Guildenstern: *Ophelia thinks that there are lots of wineries nearby.[^We do not say that 'nearby' in a speech report could *never* refer to the area near the location of the original speaker. Had Rosencrantz asked a question about San Francisco, and Guildenstern given the same response, that is presumably what it would have done. We just say that it does not *automatically* refer back to that area, and in some cases, like (35), it can refer to a quite different area. 'Nearby' behaves quite differently in this respect to 'near here', which always refers to the area near the reporter.]

Even when the contextualist claim is not *obviously* true, as with 'local' and 'enemy', disquotational reports are unacceptable after context shifts.

(36) (Brian is calling from Providence, Hud and Andy are in Bellingham)
Brian: When I get all this work done, I'll head off to a local bar for some drinks.

Andy: How much work is there?
Brian: Not much. I should get to the bar in a couple of hours.
Hud: Hey, is Brian in town? Where's he going tonight?
Andy: *He thinks he'll be at a local bar in a couple of hours.

(37) The Enemy, speaking of us: The enemy have the advantage.
One of us: How are we doing?
Another of us: *Someone just informed me that the enemy have the advantage.

(38) (Terrell is an NFL player, and Dennis is his coach.)
Terrell: Why are you cutting me coach?
Dennis: Because you are old and slow.
(After this Terrell returns to academia. Kate and Leopold are students in his department.)
Kate: Do you think Terrell would do well on our department ultimate frisbee team?
Leopold: ??I'm not sure. Someone thinks he's old and slow.

This data provides us with the penultimate argument against the contextualist theory of epistemic modals. We have already seen several such arguments.

First, as seen through the difficulties with each of the options discussed in section 2, *any* version of contextualism faces serious problems, though by altering the version of contextualism we are using, we can alter what problems we have to face.

Second, there is nothing like the speaker-inclusion constraint for terms like 'local' and 'enemy' for which contextualism is quite plausible. This disanalogy tells against the contextualist theory of 'might'. With the right stage setting (and it doesn't usually take very much), we can get 'local' and 'enemy' to mean *local to x* and *enemy of x* for pretty much any x we happen to be interested in talking about. At least for 'bare' (unembedded) epistemic modals, the situation is markedly different. We can't, just by making Jack salient, make our own knowledge irrelevant to the truth of our utterance of, for example, "Jack might have eight fingers." The only way we can make our knowledge irrelevant is if we are using this sentence in an explanation or justification of Jack's actions.[28]

Third, there is a difference in behaviour between embedded and unembedded occurrences of epistemic modals. When epistemic modals are embedded in belief contexts, conditionals, etc., they behave differently—the speaker inclusion constraint seems to be lifted. (Think about belief reports and that military instructor case.) 'Local' and 'enemy' don't seem to show any analogous difference in their behaviour between their bare and embedded occurrences.

Fourth, 'local' and 'enemy' don't generate any of the peculiar phenomena about willingness to agree. If Myles (still in Cleveland), says

(39) Many local bars are full of Browns fans.

[28] And then it would probably be more natural to say "He might have eight fingers," but that's possibly for unrelated reasons.

Professor Granger (still in the South Pacific), will not hesitate to say "that's right" (as long as she knows that many bars in Cleveland really are, as usual, full of Browns fans). The fact that the relevant bars aren't local to *her* doesn't interfere with her willingness to agree with (39) in the way that the fact that *she* knew that she wasn't in Prague interfered with her willingness to agree with Myles' claim that she might be in Prague, or in the way that Watson's knowledge that Holmes was in London (should have) interfered with his willingness to assent to Moriarty's claim that Holmes might be in Paris.

Fifth, when there is a context shift, we are generally hesitant to produce belief reports by disquoting sincerely asserted sentences involving contextually variable terms. This is what the examples (32) through (36) show. For a wide range of contextually variable terms, speakers will quite naturally hesitate to make disquotational reports unless they are in the same context as the original speaker. Such hesitation is not shown by speakers reporting epistemic modals.

The sixth argument, that there is an alternative theory that does not have these flaws, will have to wait until the next section. For now, let's note that there are other words that seem at first to be contextually variable, but for which disquotational reports seem acceptable.

(40) Vinny the Vulture: Rotting flesh tastes great.
John: Vinny thinks that rotting flesh tastes great.

(41) Ant Z: He's huge (said of 5 foot 3 141 lb NBA player Muggsy Bogues)
Andy: Ant Z thinks that Muggsy's huge.

(42) Marvin the Martian: These are the same colour (said of two colour swatches that look alike to Martians but not to humans.)
Brian: Marvin thinks that these are the same colour.

In all three cases the report is accurate, or at least extremely natural. And in all three cases it would have been inappropriate for the reporter to continue "and he's right". But crucially, in none of the three cases is it *clear* that the original speaker made a mistake. In his context, it seems Vinny utters a truth by uttering, "Rotting flesh tastes great", for rotting flesh does taste great to vultures. From Ant Z's perspective, Muggsy Bogues is huge. We assume here, a little controversially, that there is *a* use of comparative adjectives that is not relativised to a comparison class, but rather to a perspective. Ant Z does not say that Muggsy is huge for a human, or for an NBA player, but just relative to him. And he's right. Even Muggsy is huge relative to an ant. Note the contrast with (36) here. There's something quite odd about Leopold's statement, which intuitively means that someone said Terrell is old and slow for a graduate student, when all that was said was that he is old and slow for an NFL player.[29] And, relative to the Martian's classification of objects into colours, the two swatches are the same colour. So there's something very odd going on here.

The following *very* plausible principle looks like it is being violated.

[29] Or perhaps something more specific than that, such as that he is old and slow for a player at his position.

Truth in Reporting If X has a true belief, then Y's report *X believes that S* accurately reports that belief only if in the context Y is in, S expresses a true proposition.[30]

Not only do our three reports here seem to constitute counterexamples to **Truth in Reporting**, Watson's report in (30) is also such a counterexample, if Moriarty speaks truly (and sincerely). One response here would be to give up **Truth in Reporting**, but that seems like a desperate measure. And we would still have the puzzle of why we can't say "and he's right" at the end of an accurate report.

Another response to these peculiar phenomena would be to follow the universalist and conclude that Moriarty, Vinny, Ant Z and Marvin all believe something false. It should be clear how to formulate this kind of position: something tastes great iff every creature thinks it tastes great; something is huge iff it is huge relative to all observers; and two things are the same colour iff they look alike (in a colour kind of way) to every observer (in conditions that are normal for them). As we saw, there are problems for the universalist move for epistemic modals. And the attractiveness of the other universal seems to dissipate when we consider the cases from a different perspective.

(43) Brian: Cognac tastes great.
Vinny: Brian believes that cognac tastes great.

(44) Andy:He's huge (said of Buggsy Mogues, the shortest ever player in the Dinosaur Basketball Association).
Tyrone the T-Rex: Andy believes that Buggsy's huge.

(45) John: These are the same colour (said of two colour swatches that look alike to humans but not to pigeons).
Pete the Pigeon: John believes that these are the same colour.

Again, every report seems acceptable, and in every case it would seem strange for the reporter to continue "and he's right." The universalist explanation in every case is that the original utterance is false. That certainly explains the data about reports, but look at the cost! All of our utterances about colours and tastes will turn out false, as will many of our utterances about sizes. It seems we have to find a way to avoid both contextualism and universalism. Our final suggestions for how to think about epistemic modals attempt to explain all this data.

5 Relativism and Centred Worlds

John MacFarlane (2003) has argued that believers in a metaphysically open future should accept that the truth of an utterance is *relative* to a context of evaluation.[31] For

[30] One might also consider a 'says that' version of **Truth in Reporting:** If X speaks true, then Y's report *X says that S* is accurate only if in the context Y is in, S expresses a true proposition. This is more questionable, since it is questionable whether 'says that' constructions must report what is semantically expressed by a speech, as opposed to what is merely communicated. See again the papers mentioned in footnote 25.

[31] We are very grateful in this section to extensive conversations with John MacFarlane. His (2003) was one of the main inspirations for the relativist theory discussed here. His (ms), which he was kind enough to show

example, if on Thursday Emily says, "There will be a sea battle tomorrow", the believer in the open future wants to say that at the time her utterance is neither determinate true nor determinately false. One quick objection to this kind of theory is that if we look back at Emily's statement while the sea battle is raging on Friday, we are inclined to say that she got it right. From Friday's perspective, it looks like what Emily said is true. The orthodox way to reconcile these intuitions is that the only sense in which Emily's statement is *indeterminate* on Thursday is an *epistemic* sense – we simply don't know whether there will be a sea battle. MacFarlane argues instead that we should simply accept the intuitions as they stand. From Friday's perspective, Emily's statement is determinately true, from Thursday's it is not. Hence the truth of statements is relative to a context of evaluation.

There is a natural extension of this theory to the cases described above. Moriarty's statement is true relative to a context C iff it is compatible with what the people in C know that Holmes is in Paris. So in the context he uttered it, the statement is *true*, because it is consistent with what everyone in his context knows that Holmes is in Paris. But in the context of Watson's report, it is false, because Watson and Holmes know that Holmes is not in Paris.

We will call any such theory of epistemic modals a *relativist* theory, because it says that the truth of an utterance containing an epistemic modal is *relative* to a context of evaluation. As we will see, relativist theories do a much better job than contextualist theories of handling the data that troubled contextualist theories. Relativist theories are also plausible for the predicates we discussed at the end of the last section: 'huge', 'color' and 'tastes'. On such a theory, any utterance that *x tastes F* is true iff *x* tastes *F to us*. Similarly, an utterance *x is huge* that doesn't have a comparison class, as in (41) or (44), is true iff *x* is huge relative *to us*. And *Those swatches are the same color* is true iff they look the same colour *to us*. The reference to us in the truth conditions of these sentences isn't because there's a special reference to us in the lexical entry for any of these worlds. Rather, the truth of any utterance involving these terms is relative to a context of evaluation, and when that is *our* context of evaluation, *we* get to determine what is true and what is false. If the sentences were being evaluated in a different context, it would be the standards of that context that mattered to their truth.

So far we have not talked about the pragmatics of epistemic modals, assuming that their assertability conditions are given by their truth conditions plus some familiar Gricean norms. But it is not obvious how to apply some of those norms if utterance truth is contextually relative, because one of the norms is that one should say only what is true.

One option is to say that utterance appropriateness is, like utterance truth, relative to a context of evaluation. This is consistent, but it does not seem to respect the data. Watson might think that Moriarty's utterance is false, at least relative to his context

---

us a copy of while we were drafting this paper, develops the argument for a relativist approach to epistemic modals in greater detail than we do here. Mark Richard also has work in progress that develops a relativist view on related matters, which he has been kind enough to show us, and which has also influenced our thinking.

of evaluation[32], but if he is aware of Moriarty's epistemic state he should think it is *appropriate*. So if something like truth is a norm of assertion, it must be truth relative to one or other context. But which one?

We could say that one should only say things that are true relative to all contexts. But that would mean John's statement about the two swatches being the same colour would be inappropriate, and that seems wrong.

We could say that one should only say things that are true relative to some contexts. But then Brian could have said, "Rotting carcases taste great" and he would have said something appropriate, because that's true when evaluated by vultures.

The correct norm is that one should only say something that's true when evaluated in the context you are in. We assume here that contexts can include more than just the speaker. If Vinny the Vulture is speaking to a group of humans he arguably cannot say *Rotting flesh tastes great*. The reason is that rotting flesh does not taste great to the group of speakers in the conversation, most of whom are humans. This norm gives us the nice result that Myles's statement is appropriate, as is Moriarty's, even though in each case their most prominent audience member knows they speak falsely.[33]

This helps explain, we think, the somewhat ambivalent attitude we have towards speakers who express epistemic modals that are false relative to our context, but true relative to their own. What the speaker said wasn't true, so we don't want to endorse what they said. Still, there's still a distinction between such a speaker and someone who says that the sky is green or that grass is blue. That speaker would violate the properly relativised version of the *only say true things* rule, and Myles and Moriarty do not violate that rule.

As MacFarlane notes, relativist theories deny **Absoluteness of Utterance Truth**, the claim that if an utterance is true relative to one context of evaluation it is true relative to all of them. It is uncontroversial of course that the truth value of an utterance *type* can be contextually variable, the interesting claim that relativists make is that the truth value of utterance tokens can also be different relative to different contexts. So they must deny one or more premises in any argument for **Absoluteness of Utterance Truth,** such as this one.

1. **Absoluteness of Propositional Content**: If an utterance expresses the proposition *p* relative to some context of evaluation, then it expresses that proposition relative to all contexts of evaluation.
2. **Absoluteness of Propositional Truth Value**: If a proposition *p* is true relative to one context in a world it is true relative to all contexts in that world; therefore,
3. **Absoluteness of Utterance Truth**

[32] We do not assume here that ordinary speakers, like Watson, explicitly make judgments about the truth of utterances relative to a context of evaluation, as such. They do make judgments about the truth of utterances, and those judgments are made in contexts, but they don't explicitly makes judgments of *truth relative to context of evaluation*. One of the nice features, however, of the relativist account is that it is possible to do an attractive rational reconstruction of most of their views in terms of contexts.

[33] Can we even say that someone speaks falsely here now that truth and falsity is always relative to a context of evaluation? It turns out we can, indeed we must, although the matter is a little delicate. We return to this point below.

This argument provides a nice way of classifying relativist theories. One relativist approach is to say that Moriarty (or anyone else who utters an epistemic modal) says something different relative to each context of evaluation. Call this approach *content relativism*. Another approach is to say that there is a single proposition that he expresses with respect to every context, but the truth value of that proposition is contextually variable. Call this approach *truth relativism*. (So that the meaning of 'proposition' is sufficiently understood here, let us stipulate that we understand propositions to be the things that are believed and asserted and thus, relatedly, the semantic values of 'that'-clauses.)

It might look like some of our behaviour is directly inconsistent with *any* sort of relativism. Consider the following dialogue.

(46) Vinny: Rotting flesh tastes great
Vinny's brother: That's true.
John: That (i.e. what Vinny's brother said) is not true.

If what Vinny's brother is saying is that Vinny's utterance *Rotting flesh tastes great* is true in his context, then John is *wrong* in saying that what Vinny's brother said isn't true. For it is true, we claim, that *Rotting flesh tastes great* is true in Vinny's context.[34] But this prediction seems unfortunate, because John's utterance seems perfectly appropriate in his context.

The solution here is to recognise a *disquotational* concept of truth, to go alongside the *binary* concept of truth that is at the heart of the relativist solution.[35] The binary concept is a relation between an tterance and a context of evaluation. Call this $true_B$. So Vinny's utterance is true_B relative to his context, and to his brother's context, and false_B relative to John's context. One crucial feature of the binary concept is that it is not a relativist concept. If it is true relative to one context that an utterance is true_B relative to context C, it is true relative to all contexts that the utterance is true_B relative to context C. The disquotational concept is unary. Call this $true_T$. As far as is permitted by the semantic paradoxes, it claims that sentences of the form *S is* $true_T$ *iff S* will be true_B relative to any context (note here the primacy of truth_B for semantic explanation) True_T is a relative concept. An utterance can be true_T relative to C and not true_T relative to C′. When an utterance is given the honorific *true* in ordinary discourse, it is the unary relative concept true_T that is being applied. That explains what is going on in (46). Vinny's brother says that Vinny's utterance is true_T. Relative to his context, that's right, since Vinny's utterance is true in his context. But relative to John's context, that's false, because an utterance is true_T relative to John's context iff it is true relative to John's context. So John spoke truly relative to his own context, so he spoke correctly. The important point is that assignments of truth_T are relative rather than contextually rigid, so they might be judged true relative to some contexts and false relative to others.

[34] We assume here the vultures are talking mainly to other vultures, and John is talking mainly to other humans.

[35] We are grateful to John Macfarlane for helpful correspondence that influenced what follows.

Although both truth relativism and content relativism can explain (46) if they help themselves to the distinction between truth_B and truth_T, there are four major problems for content relativism that seem to show it is not the correct theory.

The first problem concerns embeddings of "might" clauses in belief contexts. Suppose Watson says,

(47) Moriarty believes that Holmes might be in Paris.

On the content relativist view, (47) will say, relative to Watson, that Moriarty believes that, as far as Watson knows, Holmes is in Paris. That would be a crazy thing for Watson to assert. Suppose Watson is talking to Holmes. Then, relative to Holmes, Watson will have claimed that Moriarty believes that, as far as Holmes knows, Holmes is in Paris. That would *also* be a crazy thing for Watson to assert. But, given what he's just overheard, it would be perfectly natural—and pretty clearly *correct*, so long as nothing funny is going on behind the scenes—for Watson to assert (47). A view that tells us that Watson's saying something crazy relative to everybody who's likely to be a member of his audience is in pretty serious conflict with our pretheoretical judgements about the case. (Enlarging the context to include both Holmes and Watson obviously doesn't help, either.)

The second problem concerns the social function of assertion. In particular, it causes difficulties for an attractive part of the Stalnakerian story about assertion, that the central role of an assertion is to add the proposition asserted to the stock of conversational presuppositions (Stalnaker 1978). On the content relativist view, it can't be that the essential effect of assertion is to add *the proposition asserted* to the stock of common presuppositions, because there's no such thing as *the* proposition asserted. There will be a different proposition asserted relative to each audience member. That's not part of an attractive theory. And it's not terribly clear what the replacement story about the essential effect of assertion—about the fundamental role of assertion in communication—is going to be. It may be that there's a story to be told about *assertability*—about when Moriarty is entitled to assert, for example, "it might be that Holmes is in Paris"—but there's no obvious story about what he's *up to* when he's making that assertion—about what the assertion is supposed to accomplish. (And if you think that appropriateness of assertion's got to be tied up with what your assertion's supposed to accomplish, then you'll be sceptical about even the first part.)

The third problem concerns epistemic modals in the scope of temporal modifiers. The content relativist has difficulties explaining what's going on with sentences like (48).

(48) The Trojans were hesitant in attacking because Achilles might have been with the Greek army.

On the content relativist view, (48) will be false relative to pretty much everybody—certainly relative to everybody alive today. It's certainly false that the Trojans were hesitant because, as far as *we* know, Achilles was with the Greek army. (Or worse, because,

as far as we knew *then*, Achilles was with the Greek army.) But, depending on how the Trojan war went, (48) could be *true* relative to everybody.[36]

Finally, content relativism has a problem with commands. Keith's Mom says:

(49) For all days *d*, you should carry an umbrella on d if and only if it might rain on *d*.

Suppose on Monday Keith checks the forecast and it says there's a 50% chance of rain. So he takes an umbrella. It doesn't rain, and on Tuesday he wonders whether what he did on Monday was what his Mom said he should. On the content relativist view, we get the following strange result: on Monday, it would have been true to say that he was doing what his Mom said he should, since at the time, the embedded clause expressed a proposition that was true relative to him. Looking back on Tuesday, though, it looks like he did what his Mom said he shouldn't, because now the embedded clause expresses a proposition that's false relative to him. But that's not right. He just plain did what his Mom told him to do.

The same thing happens with the soldiers trying to follow the imperative issued as (22). Assume one of them attempts to follow the command by burning down some trees that seem to contain snipers. Relative to the time she is doing the burning, she will be complying with the command. But later, when it turns out the trees were sniper-free, she will not have been following the command. If we assume there's an overarching command to not use flamethrowers unless explicitly instructed to do so, then it will turn out that, as of *now*, she violated her orders *then*. But that's not right. She just plain followed her orders.

There's a similar problem with the other terms about which relativism seems plausible. Consider the following commands:

(50) Don't pick fights with huge opponents.
(51) Stack all of the things that are the same color together.
(52) If it tastes lousy, spit it out.

It's possible to sensibly issue these commands, even in relevantly mixed company. And if we're going to get the right compliance conditions, we don't want *content* relativism about great-tastingness, hugeness, and same-coloredness here. When we hear a command like (52), we take (a) the same command to have been issued to everybody, and (b) everybody to be following it if we all spit out the things that taste lousy to us. On the content relativist view, we've each gotten different commands, and the philosopher who spits out the chunk of week-old antelope hasn't complied with the command that Vinny was given. This seems wrong.

So the content relativist theory has several problems. The truth relativist theory does much better. Let us begin with the familiar notion of a function from worlds to truth values. Call any such function a Modal Profile. On the standard way of looking at

[36] We don't take any stand here on just how the war went, if it happened at all. The important point is that whether (48) is true when said of a particular battle is a wide-open empirical question, not one that can be settled by appeal to the semantics of *might*. The content relativist says, falsely, that it can be thus settled.

things, propositions – the objects of belief and assertion, the semantic values of 'that'-clauses – are, or at least determine a Modal Profile. The truth relativist denies this. According to the truth-relativist, the relevant propositions are true or false not relative to *worlds*, but relative to *positions* within worlds—that is, they're true or false relative to *centered worlds*. (A centered world is a triple of a possible world, an individual, and a time.) There's a few ways to formally spell out this idea. One is to replace talk of Modal Profiles with Centring Profiles, i.e. functions from *centred* worlds to truth values. Another is to say that a centred world and proposition combine to determine a Modal Profile, so propositions determine functions from centred worlds to Modal Profiles. Each of these proposals has some costs and benefits, and we postpone discussion of their comparative virtues to an appendix. For now we are interested in the idea, common to these proposals, that propositions only determine truth values relative to something much more fine-grained than a world. (We take no stand here on whether propositions should be *identified* with either Modal Profiles or Centering Profiles or functions from Centred Worlds to Modal Profiles).

Truth relativism is not threatened by the four problems that undermine content relativism.

According to truth relativism, Watson and Moriarty express the very same proposition by the words *Holmes might be in Paris*, so it is no surprise that Watson can report Moriarty's assertive utterance by using the very same words. Similarly, it is no surprise that if Moriarty has a belief that he would express by saying *Holmes might be in Paris*, Watson can report that by (53).

(53) Moriarty believes that Holmes might be in Paris.

Above we noted that it's unlikely that Watson could use this to express the proposition that for all *Watson* knows Holmes is in Paris. We used that fact to argue that DeRose's constraint did not apply when an epistemic modal is inside a propositional attitude report. The truth relativist theory predicts not only that DeRose's constraint should not apply, but that a different constraint should apply. When one says that *a believes that b might be F*, one says that *a* believes the proposition *b might be F*. And *a* believes that proposition iff *a* believes it is consistent with what they know that *b* is *F*. And that prediction seems to be entirely correct. It is impossible for Watson to use (53) to mean that Moriarty believes that for all Holmes knows he is in Paris, or that for all Watson knows Holmes is in Paris. This seems to be an interesting generalisation, and while it falls out nicely from the truth relativist theory, it needs to be imposed as a special constraint on contextualist theories.

Since there is a proposition that is common to speakers and hearers when an epistemic modal is uttered, we can keep Stalnaker's nice idea that the role of assertion is to add propositions to the conversational context. Since propositions are no longer identified with sets of possible worlds we will have to modify other parts of Stalnaker's theory, but those parts are considerably more controversial.

The truth relativist can also explain how (48) can be true, though the explanation requires a small detour through the nature of psychological explanations involving rel-

ativist expressions go.

(48) The Trojans were hesitant in attacking because Achilles might have been with the Greek army.

All of the following could be true, and not because the things in question are rude, huge or great tasting for us.

(54) Marvin the Martian dropped his pants as the Queen passed by because it would have been rude not to.
(55) Children are scared of adults because they are huge.
(56) Vultures eat rotting flesh because it tastes great.

In general it seems that the truth of an explanatory claim of the form, X φed because p depends only on whether p is true in X's context (plus whether the truth of p in X's context bears the right relation to X's φing).. Whether or not p is true in *our* context is neither here nor there. Adults are not huge, rotting flesh does not taste great, and it is rude to drop one's pants as the Queen passes by, but (54)-(56) could still be true, and could all count as good explanations. Similarly, (48) can be true because *Achilles might have been with the Greek army* could be true relative to the Trojans.

Similarly, what it is to comply with a command is to act in a way that makes the command true in the context of action. This is not a particular feature of epistemic modals, but just a general property of how commands involving propositions with centered-worlds truth conditions behave. If Don picks a fight with Pedro after Don has shrunk so much that Pedro is now relatively huge, he violates (50), even if Pedro was not huge when the command was issued. And he still violates it from a later perspective when Pedro and Don are the same size. The general point is that whether the command is violated depends on the applicability of the salient terms from the perspective of the person to whom the command applies. Similarly, Keith does not violate his Mom's command if he takes an umbrella where *It might rain* is true in the context the action is performed. And this, of course, matches up perfectly with intuitions about the case.

It's a little tricky to say just which statement in Professor Granger's original hexalemma gets denied by the truth relativist. It all depends what we mean by *spoke truly*. If *Myles spoke truly* means that Myles said something true_{T}, then (2) is false (relative to Granger's context), for its right-hand-side is true but its left-hand-side is false. If, on the other hand, it means he said something true_{B} relative to his own context, then (4) is false, for he did speak truly_{B} relative to his context, but it's not the case that Professor Granger might be in Prague. This is awkward, but we might expect that any good solution to the paradox will be awkward.

6 Objections to Truth Relativism

It might be thought that the truth relativist has to deny **Truth in Reporting,** but in fact this can be retained in its entirety provided we understand it the right way. The

following situation is possible on the truth relativist theory. *X* has a belief that is true in her context, and *Y* properly reports this by saying *X believes that S*, where *S* in *Y*'s mouth expresses a proposition that is false in *Y*'s mouth in her context. But this is no violation of **Truth in Reporting**. What would be a violation is if *X*'s belief was true in *Y*'s context, and still *Y* could report it as described here. But there's no case where, intuitively, we properly report an epistemic modal but violate that constraint. And the same holds for reports of uses of *huge*, *color* or *tastes*. Even if Vinny (truly) believes that rotting flesh tastes great, and the words "Rotting flesh tastes great" in John's mouth express a false proposition, John's report, "Vinny believes that rotting flesh tastes great" would only violate **Truth in Reporting** if Vinny's belief is still true in *John's* context. And it is not.

Given that the relativist has the concept of truth_T, or as we might put it truth *simpliciter*, what should be done with it? The answer seems to be not much. We certainly shouldn't restate the norms of assertion in terms of it, because that will lead to the appropriateness of assertion being oddly relativised. Whether it was appropriate for Vinny to *say* "Rotting flesh tastes great," is independent of the context of evaluation, even if the truth of what he uttered is context-relative. (It would not at all be appropriate for him to have said "Rotting flesh tastes terrible" even though we should think he would have said something true by that remark, and something false by what he actually said.) And the same thing seems to hold for generalisations about truth as the end of belief. It is entirely appropriate for Myles to *believe* that Granger might be in Prague, because it's true_B relative to his context. Relatedly, if knowledge is tied to truth_T rather than truth_B, knowledge can't be the norm of assertion or end of belief.[37] On the other hand, using truth_T we can say that **Truth in Reporting** is true in the truth relativist theory without reinterpreting it in terms of relative truth concepts. Moreover, we can invoke truth_T to explain why we got confused when thinking about the original puzzle: It is arguable that, even if we should distinguish truth_T from truth_B in our semantic theorizing, we aren't unreflectively as clear about that distinction as we might be. No wonder then that we get a little confused as we think about the Granger case. We want to say Myles doesn't make a mistake. And we also want to say "That's wrong" speaking of the object of his assertion and belief, and what's more, when we say that, we don't seem to be making a binary claim about the relation between ourselves and what is believes. Once we clearly distinguish truth_T from truth_B things become clearly. Using the disquotational notion, we can say 'That is false_T', which is a monadic claim, and not a binary one. The binary truth_B explains why that claim is assertable (it is assertable because 'That is false_T' is truth_B at my context), but doesn't figure in the proposition believed. Meanwhile, the relevant notion of mistake – that of an agent believing a proposition that is not true_B at her context, can only be properly articulated once the distinction between the more explanatory truth_T is carefully distinguished from the (arguably) conceptually more basic

[37] Arguably, then, one will have to distinguish (and posit an ordinary conflation between) knowledge_T from knowledge_B, the latter being needed to make good on the normative importance of knowledge, the former being need to make sense of the validity of the inference from knowing that p to p. Is trouble lurking here for the truth relativist, especially given link between the truth_B of 'might' claims and facts about knowledge? We shall not pursue the matter further here.

truth$_{B.}$

One final expository point. In general, truth relativism makes for irresolvable disputes. Let us say that two conversational partners are in deadlock concerning a claim when the following situation arises: There is a pair of conversational participants, x and y, and a sentence S, under dispute, such that each express the same proposition (in the sense explained) by S but that S is true$_B$ at each of the contexts x is in during the conversation, and false$_B$ at each of the contexts y is in during the conversation. Neither speaks past one another in alternately asserting and denying the same sentence, since each expresses the same proposition by it. And each asserts what they should be asserting when each says: What I say is truth$_T$ and what the other says is false$_{T.,}$ since each makes a speech that is true$_T$ at the respective contexts. In general, truth relativism about a term will lead one to predict deadlock for certain conversations, traceable to the truth relativity of the term. But in the case of 'might', it is arguable that conversation tends to force a situation where, even if at the outset, a 'might' sentence was true relative to x and not to y (on account of the truth-relativity of the 'might' sentence), x and y will, in the course of engagement and dispute, be quickly put into a pair of contexts which do not differ with respect to truth$_B$ (unless the 'might' sentence contained other terms that themselves made for deadlock). This is not merely because the conversational participants will, through testimony, pool knowledge about the sentence embedded in the 'might' claim. It is in any case arguable that the relevant community whose body of knowledge determines whether a 'might' claim is true$_B$ at a context always includes not just that of the person at that context but also that of his conversational partners. In the special case of 'might', then, Truth Relativism may well generate far less by way of deadlock than in other cases.

There are two primary objections to the truth relativist theory: it doesn't quite handle all the cases and that it is too radical.

There are some cases that seem to tell directly against the truth relativist position. Consider the case again of Tom and Sally stuck in a maze. Sally knows the way out, but doesn't want to tell Tom. She says, inter alia, (57), and does not seem to violate any *semantic* norms in doing so, even though she knows the exit is some other way.

(57) The exit might be that way.

This seems to directly contradict the relativist claim that the norm for assertion is speaking truly in one's own context. We suspect that what's going on here is that Sally is projecting herself into Tom's context. She is, we think, merely trying to verbalise thoughts that are, or should be, going through *Tom's* head, rather than making a simple assertion. As some evidence for this, note (as was mentioned above) that it would be wrong to take (57) as evidence that Sally believes the exit might be that way, whereas when a speaker asserts that *p* that is usually strong evidence that she believes that *p*. It is unfortunate for the relativist to have to appeal to something like projection, but we think it is the simplest explanation of these cases that any theorist can provide.

The idea that utterances have their truth value absolutely is well-entrenched in contemporary semantics, so it should only be overturned with caution. And it might be

worried that once we add another degree of relativisation, it will be open to relativise in all sorts of directions. We are sensitive to these concerns, but we think the virtues of the relativist theory, and the vices of the contextualist and invariantist theories, provides a decent response to them. Invariantist theories are simply implausible, and any contextualist theory will have to include so many ad hoc conditions, conditions that seem to be natural consequences of relativism, that there are methodological considerations telling in favour of relativism. (Let us be clear: we are not recommending a general preference for relativism over contextualism in semantic theory. As we have been trying to make clear, for example, the case of 'might' is very different from, say, the case of 'ready'.) It is (as always) hard to tell which way the balance tips when all these methodological considerations are weighed together, but we think the relativist has a good case.

Appendix on Types of Content

Robert Stalnaker has long promoted the idea that the content of an assertoric utterance is a set of possible worlds, or a function from worlds to truth values. This idea has been enormously influential in formal semantics, although it has come in for detailed criticism by various philosophers. (See especially Soames (1987) and King (1994, 1995, 1998).) But even philosophers who think that there is more to content than a set of possible worlds would agree that propositions determine a function from worlds to truth values. Some would agree that such a function exhausts the 'discriminatory role' of a proposition, although this depends on the (highly contestable) assumption that the role of propositions is to discriminate amongst *metaphysical* possibilities. Still, even philosophers who disagree with what Stalnaker says about the nature of propositions could agree that if all we wanted from a proposition was to divide up some metaphysical possibilities, propositions could be functions from worlds to truth values, but they think some propositions that divide up the metaphysical possibilities the same way should be distinguished.

We don't want to take sides in that debate, because our truth relativism means we are in conflict with even the idea that a proposition determines a function from worlds to truth values. To see this, consider a sentence whose truth value is relative to a context of evaluation, such as *Vegemite tastes great*. The truth relativist says that this sentence should be evaluated as true from a context where people like the taste of Vegemite (call this the Australian context) and should be evaluated as false from a context where people dislike this taste (call that the American context) and both evaluations are correct (from their own perspective) even though the Australians and Americans agree about what the content of *Vegemite tastes great* is, and they are in the same world. So there's just no such thing as *the* truth value of *Vegemite tastes great* in the actual world, so it does not determine a function from worlds to truth values. What kind of function does it determine then?

One option, inspired by Lewis's work on *de se* belief, is to say that it determines a function from centred worlds to truth values. The idea is that we can identify a context of evaluation with a centred world, and then *Vegemite tastes great* will be true relative

to a centred world iff it is properly evaluated as true within that context. Alternatively, the content of *Vegemite tastes great* will determine a set of centred worlds, the set of contexts from which that sentence would be evaluated as true. Just as propositions were traditionally thought to determine (or be) sets of possible worlds, properties were traditionally thought to determine (or be) functions from worlds to sets of individuals.[38] Now if we identify centred worlds with ⟨individual, world⟩ pairs, a function from worlds to sets of individuals just is a set of centred worlds.[39] So the content of *Vegemite tastes great* could just be a **property**, very roughly the property of being in a context where most people are disposed to find Vegemite great-tasting.

This proposal has three nice features. First, even though the content of *Vegemite tastes great* is not, and does not even determine, a proposition as Stalnaker conceived of propositions, it does determine a property. So the proposal is not as radical as it might at first look. Second, properties are the kind of thing that divide up possibilities. The possibilities they divide are individuals, not worlds, but the basic idea that to represent is to represent yourself as being in one class of possible states rather than another is retained. The only change is that instead of representing yourself as being in one class of worlds rather than another, you represent yourself as being in one class of ⟨individual, world⟩ pairs rather than another. Third, the proposal links up nicely with David Lewis's account of *de se* belief, and offers some prospects for connecting the contents of beliefs with the contents of assertions, even when both of these contents have ceased to be propositions in Stalnaker's sense.[40]

But there's a problem for this account. Consider what we want to say about *Possibly Vegemite tastes great*, where context makes it clear that the 'possibly' is a metaphysical modal. There's a trivial problem and a potentially deep problem for this account. The trivial problem is that we know what the meaning of *possibly* is. It's a function that takes propositions as inputs and delivers as output a proposition that is true iff the input proposition is true at an accessible world. If the content of *Vegemite tastes great* is a property rather than a proposition, then we have a type-mismatch. This is a trivial problem because it's a fairly routine exercise to convert the meanings of words like *possibly* so they are the right kind of things to operate on what we now take the meaning of *Vegemite tastes great* to be.

The deep problem is that when we go through that routine exercise, we get the wrong results. We don't want *Possibly Vegemite tastes great* to be true in virtue of there being an accessible world where the people *there* like the taste of Vegemite. We want it to be true

[38] Lewis preferred the theory on which properties were sets of individuals, potentially from different worlds. This theory has difficulties accounting for individuals that exist in more than one world. And since properties exist in more than one world, and properties have to be treated as individuals in some contexts (e.g. when they are the subjects of predication) this is a serious problem. Treating properties as functions from worlds to sets of individuals removes this problem without introducing any other costs. (See Egan (2004) for more details.)

[39] Matters are a little more complicated when we introduce times into the story. For purposes of this appendix we ignore all matters to do with tense. As you'll see, the story is complicated enough as it is, and this omission doesn't seriously affect the dialectic to follow.

[40] It might be that propositions just are whatever things are the contents of assertions and beliefs, so we shouldn't say that the contents of sentences like *Vegemite tastes great* are not propositions. But they will be very different kinds of propositions to what we are used to. Thanks here to John MacFarlane.

in virtue of there being a world where Vegemite's taste is a taste that in this context we'd properly describe as great. And it's not clear how to get that on the current story. To see how big a problem this is, consider (58), where the modal is meant to be metaphysical and have wide scope.

(58) Possibly everyone hates Vegemite but it tastes great.

That's true, on its most natural reading. But the content of *Everyone hates Vegemite but it tastes great* will be the empty set of centred worlds, for there is no centred world on which this is true. Now it's not clear just what the meaning of *possibly* could be that delivers the correct result that (58) is true.

So we are tempted to consider an alternative proposal. Start with a very natural way of thinking about why the relativist has to modify the Stalnakerian story about content. The problem is that (even given a context of utterance) *tastes great* does not determine a property. Rather, relative to any context of evaluation, i.e. centred world, it determines a property. That is, its content is (or at least determines) a function from centred worlds to properties. So given our actual context, it determines the property of having a taste that people around here think is great. Now properties combine with individuals to form Stalnakerian propositions. So *tastes great* is a function from centred worlds to functions from individuals to sets of worlds. Hence *Vegemite tastes great* is a function from centred worlds to sets of worlds, the previous function with the value for the 'individual' being fixed as Vegemite.

Our second option then is that in general that sentences containing 'relative' terms like 'tastes' or 'huge' or 'might' determines a function from centred worlds to sets of worlds. This makes it quite easy to understand how (58) could work. *Possibly* type-shifts so that it is now a function from functions from centred worlds to sets of worlds to functions from centred worlds to sets of worlds. It's fairly easy to say what this function is. If the content of p is (or determines) f, a function from centred worlds to sets of worlds, then the content of $\Diamond p$ is (or determines) g, the function such that for any centred world c, $w \in g(c)$ iff for some w' accessible from w, $w' \in f(c)$. The core idea is just that we ignore the role of the centred worlds until the end of our semantic evaluation, and otherwise just treat $\Diamond$ as we'd treated it in traditional semantics. This is a rather nice position in many ways, but there are two issues to be addressed.

First, it is not clear that functions from centred worlds to sets of worlds are really kinds of content. They are not things that divide up intuitive possibilities, in the way that sets of individuals, and sets of ⟨individual, world⟩ pairs do. It's no good to say that relative to a centred world a content is determined. That would be fine if we were content relativists, and we said the content was meant to be determined relative to a centred world. But as argued in the text the content of *Vegemite tastes great* should be the same across various contexts of evaluation. A better response is to say functions from centred worlds to sets of worlds do determine a kind of content. For any such function f, we can determine the set of centred worlds $\langle i, w\rangle$ such that $w \in f(\langle i, w\rangle)$. These will be the centred worlds that the proposition is true at. It's not necessarily a problem that the proposition does *more* than determine this set. (It's not an objection to King's ac-

count of propositions that on his theory propositions do more than determine a set of possibilities.)

Second, it isn't exactly clear how to fill out these functions when we get back to our core case: epistemic modals. It's easy to say what it is for *Vegemite tastes great* to be true in a world relative to our context of evaluation; indeed we did so above. It's a lot harder to say what it is for *Granger might be in Prague* to be true in an arbitrary world w relative to an arbitrary context of evaluation c. As a first pass, we might say this is true in w iff for all the people in c know, it is true in w that Granger is in Prague. But the problem is that whenever c is not a centre in w, it's very hard to say just what the people in c know about w. Under different descriptions of w they will know different things about it. If w is described as a nearby world in which Granger is in Cleveland, they will know Granger is not in Prague in w. If it is described as a nearby world in which Myles knows where Granger is they may not know anything about whether Granger is in Prague is in w, even if those descriptions pick out the same worlds. Ideally we would cut through this by talking about their *de re* knowledge about w, but most folks have very little *de re* knowledge about other possible worlds. It's not clear this is a huge problem though. Remember that a sentence containing an epistemic modal is meant to determine a function from centred worlds to functions from worlds to truth values. Provided we have a semantics that allows for semantic indeterminacy, we can just say that the functions from worlds to truth values are partial functions, and they simply aren't determined when it's unclear what the people in c know about w. Or we can say there's a default semantic rule such that w is not in $f(c)$ (where f is the function determined by the sentence) whenever this is unclear. Since the sentences whose meanings are determined by these values of the function, like *Possibly Granger might be in Prague* are similarly vague, it is no harm if the function is a little vague.

So we have two options on the table for what kind of functions sentences might determine if they don't determine functions from world to truth values. One option is that they determine functions from centred worlds to truth values, another that they determine functions from centred worlds to functions from worlds to truth values. Neither is free from criticism, and the authors aren't in agreement about which is the best approach, so it isn't entirely clear what the best way to formally implement truth relativism is. But it does not look like there are no possible moves here. Moving to truth relativism does not mean that we will have to totally abandon the fruitful approaches to formal semantics that are built on ideas like Stalnaker's, although it does mean that those semantic theories will need to be modified in places.

References

Capellen, Herman, and Ernest Lepore. 1997. "On an Alleged Connection Between Indirect Quotation and Semantic Theory." *Mind and Language* 12 (3-4): 278–96. doi: 10.1111/j.1468-0017.1997.tb00075.x.

DeRose, Keith. 1991. "Epistemic Possibilities." *Philosophical Review* 100 (4): 581–605. doi: 10.2307/2185175.

———. 1998. "Simple Might's, Indicative Possibilities, and the Open Future." *The Philosophical Quarterly* 48 (190): 67–82. doi: 10.1111/1467-9213.00082.

Egan, Andy. 2004. "Second-Order Predication and the Metaphysics of Properties." *Australasian Journal of Philosophy* 82 (1): 48–66. doi: 10.1080/713659803.

Fintel, Kai von, and Sabine Iatridou. 2003. "Epistemic Containment." *Linguistic Inquiry* 34 (2): 173–98. doi: 10.1162/002438903321663370.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Horn, Laurence. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.

Jacobson, Pauline. 1999. "Towards a Variable Free Semantics." *Linguistics and Philosophy* 22 (2): 117–84. doi: 10.1023/A:1005464228727.

King, Jeffrey. 1994. "Can Propositions Be Naturalistically Acceptable?" *Midwest Studies in Philosophy* 19 (1): 53–75. doi: 10.1111/j.1475-4975.1994.tb00279.x.

———. 1995. "Structured Propositions and Complex Predicates." *Noûs* 29 (4): 516–35. doi: 10.2307/2216285.

———. 1998. "What Is a Philosophical Analysis?" *Philosophical Studies* 90 (2): 155–79. doi: 10.1023/A:1004254128428.

Lewis, David. 1976. "The Paradoxes of Time Travel." *American Philosophical Quarterly* 13 (2): 145–52. Reprinted in his *Philosophical Papers*, Volume 2, Oxford: Oxford University Press, 1986, 67-80. References to reprint.

———. 1979. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8 (1): 339–59. doi: 10.1007/bf00258436. Reprinted in his *Philosophical Papers*, Volume 1, Oxford: Oxford University Press, 1983, 233-249. References to reprint.

MacFarlane, John. 2003. "Future Contingents and Relative Truth." *The Philosophical Quarterly* 53 (212): 321–36. doi: 10.1111/1467-9213.00315.

Pagin, Peter. 2005. "Compositionality and Context." In *Contextualism in Philosophy: Knowledge, Meaning, and Truth*, edited by Gerhard Preyer and Georg Peter, 303–48. Oxford: Oxford University Press.

Soames, Scott. 1987. "Direct Reference, Propositional Attitudes and Semantic Content." *Philosophial Topics* 15 (1): 47–87. doi: 10.5840/philtopics198715112.

———. 2002. *Beyond Rigidity*. Oxford: Oxford University Press.

Stalnaker, Robert. 1978. "Assertion." *Syntax and Semantics* 9: 315–32.

Stanley, Jason. 2000. "Context and Logical Form." *Linguistics and Philosophy* 23 (4): 391–434. doi: 10.1023/A:1005599312747.

Stanley, Jason, and Zoltán Gendler Szabó. 2000. "On Quantifier Domain Restriction." *Mind and Language* 15 (2&3): 219–61. doi: 10.1111/1468-0017.00130.

Published in *Contextualism in Philosophy*, 2005, pp. 131-169.