Notes on iffy knowledge*

Seth Yalcin
yalcin@berkeley.edu

April 12, 2018

E. B. White once said that it is best to have a strong curiosity, weak affiliations. I am pretty sure he was talking about the compositional semantic interaction of attitude verbs with conditionals. Anyway, it's good advice for thinking about that topic. In this exploratory note I frame some problems and questions about how 'knows' and if-clauses interact, not really going far into detailed solutions. I hope we can discuss some of these issues in the seminar.

1

I have two favorite cafes. They are across the street from one another. I want to take you to one of them but don't care which, so I pick one at random and off we go. Sitting in the cafe, I can say (1) but not (2):

(1) If we weren't at this cafe, we'd be at the one across the street.

(2) ?? If we're not at this cafe, we're at the one across the street.

There's a pattern of explanation why, familiar from Stalnaker [1975]. Among the possibilities compatible with what is presupposed in our conversation—among the worlds in the context set—none are possibilities where we aren't at this cafe. But unembedded indicatives—unlike counterfactuals—presuppose that there be antecedent worlds in the context set of the conversation. So the presupposition of (2) is not satisfied in our context; so it is marked.

Now let me rewrite the story a little bit. As before, I take you to one of my favorite two cafes, picking at random. We're in the cafe waiting for Bill to join us. Bill glances into our cafe but fails to spot us—the cafe is big and busy. He then goes looking for us in the cafe across the street. I might explain as follows: "I told him we'd be in this cafe or that one, so..."

(3) Bill thinks that if we're not at this cafe, we're at the one across the street.

*For the NYU Mind and Language seminar. References very incomplete. See especially relevant discussion in Moss [2018].

The attitude ascription looks like it embeds the indicative conditional (2). The ascription is fine, so if it does embed an indicative conditional with the presupposition recently described, the presupposition is not inherited by the whole ascription. This is not very surprising, since classically 'believes' is maybe the most famous example of a presupposition 'plug' (Karttunen [1973]).

The plug imagery is misleading, though, if it suggests that (3) has no presuppositions at all. An extremely tempting analysis of this situation would describe the presupposition of the embedded indicative as 'locally satisfied' by the attitude verb context: the presupposition carried by the conditional is now understood as placing a demand, not on the context set, but on Bill's state of belief. The presupposition of (3) is that Bill's doxastic state leaves open possibilities where we're not at this cafe. Such an analysis would predict that if none of Bill's doxastic alternatives make the antecedent true, the ascription should be marked. This idea seems to be correct. Consider for instance:

(4) Ted is absolutely certain that nobody but Oswald killed Kennedy. ??But he thinks that if Oswald didn't do it, someone else did.

Once we make it clear that Ted lacks doxastic alternatives where Oswald didn't do it, the ascription with the corresponding embedded indicative seems marked.[1] One could implement this kind of picture of presupposition projection concretely in many ways. For my purposes it is sufficient to gesture at this general direction of analysis.

Going back to the cafe story, I want to consider some things I could have said instead of (3). Again, the situation is that Bill glances into our cafe but fails to spot us. He then goes looking for us in the cafe across the street. I might explain as follows: "I told him we'd be in this cafe or that one, so..."

(6) Bill knows that we're either at this cafe or at the one across the street.

According to most everyone I've asked, I also could have said:

(7) Bill knows that if we're not at this cafe, we're at the one across the street.

(7) differs from (3) just in the replacement of 'believes' with 'knows'.

[1] Note that this discourse is fine:

(5) Ted thinks Oswald killed Kennedy. But he thinks that if Oswald didn't do it, someone else did.

In the context of the assumptions made so far, this is evidence that one can count as thinking or believing that p without being in a state that entirely eliminates $\neg p$ worlds (Hawthorne, Rothschild, and Spectre [2016], Beddor and Goldstein [2017], Yalcin [2018]). This is why I resorted to 'absolutely certain' in the above discourse.

The felicity of (7) is surprising.[2] 'Knows' is factive—it is thought to presuppose its complement. It is also often thought to be a presupposition hole, allowing the presuppositions of what it embeds to pass through. It is also often thought that *A knows that* ϕ entails ϕ. All three of these ideas seem to be in a bad way here. First, (7) doesn't seem to presuppose its complement—as noted at the outset, we are in a context where it would be infelicitous to assert 'If we're not at this cafe, we're at the one across the street'. This conditional is not presupposed by us, and neither are we apt to accommodate it under the impact of (7). Second, if (2) carries the sort of presupposition we mentioned earlier, (7) evidently does not inherit this presupposition—contra the idea that presuppositionally speaking 'knows' is holey. Third, the entailment from (7) to (2) seems to fail in our context. Indeed, it seems we can rationally believe (7) while not believing (2).

At a first glance, this example seems like good news for certain dynamic or information-sensitive analyses of the interaction between the attitude verb and the conditional (e.g., Heim [1983, 1992], Gillies [2004], Yalcin [2007]). These approaches would say:

- (7) is true just in case, when you restrict attention to the subset of Bill's epistemic alternatives where we are not at this cafe, you find that all these worlds are worlds where we are across the street.
- (2) is accepted (incorporated, supported) by a state of information (such as a context set) just in case the state includes a world where the antecedent is true, and every antecedent world in the state is one where the consequent is true.

On an approach in this vein, one can see that adding the truth-conditional content expressed by (7) to a context set c needn't result in (2)'s being accepted at c. If entailment is modeled as informational consequence (in the terminology of Yalcin [2007]; see also Stalnaker [1975], Veltman [1996], Bledin [2014]), then (7) doesn't entail (2). In (7) what's happening is that we are starting with Bill's epistemic state K_b, considering the result of updating it with the information that we are not in this cafe ($K_b[\neg C]$)—in the process locally satisfying the presupposition of the indicative—and then asking whether we're at the other cafe throughout this updated state. The thing to note is that $K_b[\neg C]$ is *not* an epistemic state. Certainly it needn't include the actual world. So facts about what is true throughout this state don't imply anything about actuality. This seems to explain the absence of the presupposition, and the failure of entailment.

[2] If you actually think this sentence is infelicitous, don't worry—section 4 below is just for you.

A residual question for this view is the status of the presupposition affiliated with 'knows'. We might have written the lexical entry for 'knows' so that *A knows that* ϕ determines an update on information states that is defined only if the input state incorporates $[\phi]$ (cf. Heim [1983, 1992], Beaver [2001]). Or we might have written it in such a way as to require that 'knows' requires combination with something called a 'fact'. Clearly, both of these approaches would be nonstarters here, so we must find something different to say about what it means for 'knows' to be presuppositional.[3]

Incidentally, MacFarlane (p.c.) asks whether the same puzzle arises for bare epistemic modals. He offers this dialogue:

> A: Why is he looking for you across the street?
>
> B: Because he knows I might be there.

The phenomenon seems parallel.

2

It is useful to have a second case. Here's one due to Kai von Fintel, in reaction to an earlier version of these notes.

We randomly put a ball under one of three cups—A, B, and C. It ends up under cup A. We make Jane guess. She picks C. We say "Nope, it's not under C. It's either under A or B. Guess again!" We can say:

(8) Jane knows it's under either A or B.

But what about:

(9) Jane knows that if it's not under A, it's under B.

Is there any asymmetry in acceptability between these? Pretty much everyone who isn't a semanticist tells me they are both true (but von Fintel reports he feels a contrast).

A key difference between this case and the cafe case is that while the counterfactual (1) is true, the corresponding counterfactual 'If the ball weren't under A, it would be under B' is not true. Thus if one had hoped to blame the likeability of (7) somehow on the truth of the counterfactual corresponding to the embedded indicative, we seem to have an obstacle to that idea here.

[3] If we wanted to keep the idea that 'knows' requires its complement to be presupposed, one far-out solution might be to say that in (7), 'knows' does not, despite appearances, have an indicative as its complement after all. Rather, the if-clause composes with 'knows'—roughly, we have a complex predicate of the form 'knows that if ϕ'— and in the composition of this complex predicate, the if-clause manages to deactivate the presupposition of 'knows'.

We can easily turn this into something like a Sly Pete-style case (Gibbard [1981]; see also [Stalnaker, 2014, 162-3]). Let Carl play the game too, in isolation from Jane. We make Carl guess. He picks B. "Nope, it's not under B. It's either under A or C. Guess again!" Again, people apparently find these fine:

(10) Carl knows it's under either A or C.

(11) Carl knows that if it's not under A, it's under C.

Now given all that background, can we also say this?

(12) Jane knows that if it's not under A, it's under B, and Carl knows that if it's not under A, it's under C.

When I ask ordinary speakers not under the influence of training in linguistics or epistemology, they tell me that are fine with it—to my surprise, and to the disgust of many semanticists I know.

Incidentally, in an effort to provoke at least some eyebrow-raising from my non-linguist consultants, I have tried things like these:

(13) We're not in the cafe across the street. But Bill knows that if we're not in this cafe, we're in the cafe across the street.

(14) The ball is under cup A. But Jane knows that if it is not under cup A, it's under cup B.

These occasionally do trigger objection, but I don't get the level of outrage one might have expected.

Stalnaker [2014] understandably describes sentences like (12) as "paradoxical" (205), and indeed some theories of conditionals predict that this sentence says that one agent knows a proposition that is incompatible with a proposition a second agent knows. But unless ordinary speaker judgments are being systematically thrown off from tracking this fact, it appears those theories are missing something. Meanwhile (12) is not a contradiction on a dynamic or information-sensitive account—apparently not an unwelcome result.

(Another kind of view that could handle these data is one which basically takes indicative conditionals to be attitude reports. Some would like to say that indicative conditionals often express propositions about what some agent or agents know. One version of this view says that in (12), Jane is knowledge-related to a proposition about her epistemic state and Carl is knowledge-related to a different proposition about his epistemic state. So they are not said by (12) to know incompatible propositions, and there is no tension. I think such theories don't interact beautifully with epistemic modals as compared to rivals, but that's a long story.)

3

The dynamic or information-sensitive account I've been imagining does have a problem, though, which is that it makes iffy knowledge too easy. Consider:

(15) Ted is absolutely certain that nobody but Oswald killed Kennedy. So he doesn't think that if Oswald didn't do it, someone else did. ??But he knows that if Oswald didn't do it, someone else did.

Suppose that at the actual world, somebody other than Oswald killed Kennedy. Then since Ted's epistemic alternatives must include actuality, he has epistemic alternatives where somebody other than Oswald killed Kennedy. But then it's rather unclear why something of the form (15) can't be true. It seems perfectly possible to posit an epistemic state for Ted which is such that, when you update it with 'Oswald didn't do it', the resulting state accepts 'Someone else did'.

This might recommend upgrading the information-sensitive account of conditionals in Yalcin [2007] with selection functions. See for instance Santorio [2018].

4

I would expect some would want to uphold the view that *pace* the folk I have consulted, things like (7), (9), (11) and (12) ought really to get hashes in the contexts I set out for them, and really we should deliver theories that predict markedness. A less aggressive view would be that these sentences admit of two readings, one of which is marked and one of which isn't, and for reasons as yet unclear, speakers divide in respect of what reading they get. One version of this second view would predict a judgment survey on these questions to yield a bimodal distribution.

I think it is worthwhile to consider what the theoretical situation would look like if our aim was actually to predict markedness for (7), (9) and so on. We would still face some real problems; they are just different ones. (These sentences raise problems if you think they are fine, but they also raise problems if you think they are marked.)

If we wanted to say the target sentences are marked, a natural impulse would be to blame it on some kind of presupposition failure. One then needs to explain what the presuppositions are and why they are problematic. This is not as straightforward as it might look.

For example, on one kind of Kratzer-inspired view (mentioned parenthetically at the end of section 2), the sentence (7) embeds a covert necessity modal under 'knows' which is indexed to the subject. The sentence says that Bill is

knowledge-related to a proposition that is true in w just in case, in all of the (nearest) worlds w' compatible with Bill knows in w and such that we are not in this cafe in w', we're at the cafe across the street in w'. If the knowledge ascription presupposes this proposition, that *per se* shouldn't present any special problem from the point of view of a context where it is assumed that we are in the cafe we're in. As we're sitting in the cafe discussing the situation, we'd have no problem presupposing this. A dynamic or informational account of (7), one which required that its update be undefined on contexts which fail to incorporate the update associated with the embedded conditional, would clearly do better in this regard.

But I think ultimately this too would be unsatisfactory. If the problem with the target sentences were that they are marked owing to their presuppositions, they should become acceptable if we can find a way to eliminate those presuppositions. One way to do this is with disjunctions. Let's agree that

(8) Jane knows the ball is under either A or B.

presupposes that the ball under A or B. Observe that by contrast,

(16) Either the ball evaporated into thin air, or else Jane knows it's under either A or B.

does not inherit this presupposition. Note also that it is true in the context stipulated above (where the ball is under A, and Jane is told it's under A or B). Now compare that to

(17) Either the ball evaporated into thin air, or Jane knows if it's not under A, it's under B.

This sentence plausibly does *not* presuppose that if it's not under A, it's under B. If the subjects who find (9) to be marked did so only owing to its presuppositions, then the prediction should be that they will judge (17) to be unmarked and true. But I seriously doubt that anyone who finds (9) marked hears (17) as better.

So I think that if it's right to say that (7), (9), (11) and (12) should come out as semantically/pragmatically bad, the best of explanation of this markedness will not appeal merely to facts about presupposition projection. The problem must have to do with the (non-presuppositional) content of these sentences. I suspect that speakers who find these sentences marked just doubt that the relevant subjects actually have the knowledge the sentences ascribe—it's not, or not just, that they have unwanted presuppositions. For example, I doubt such speakers would be happy to say:

(18) I believe that Jane knows if it's not under A, it's under B.

This sentence doesn't presuppose the embedded conditional or the embedded knowledge ascription, so it shouldn't involve the speaker in any crash-inducing presupposition failures. But my hunch is that if you don't like (7), (9), (11) or (12), then you're not apt to affirm this sentence, and this is simply because you don't have the belief you'd be reporting you have if you said the sentence.

Retreat to belief ascription might help to bring out some of the oddness of iffy knowledge, at least as it should strike those who want to say that the target sentences are marked. Suppose I tell the guessing game story involving Jane exactly as above, except I leave out the fact about where the ball actually is, as follows. I randomly put a ball under one of three cups—A, B, and C. I make Jane guess. She picks C. I say "Nope, it's not under C. It's either under A or B." What I say is true—it is under A or B. Jane knows I'm being honest and that I'm fully informed, etc. If this is all the information I give you about the story, you won't find any problem saying that (9) is true. In the world as you believe it to be, Jane knows that if the ball is not under A, it is under B. Furthermore, *your* epistemic and doxastic position is on par with Jane's: you too think that if the ball is not under A, it is under B.

However, suppose now I give you the information that the ball is under A, so that you know it's under A. You are no longer correctly described as thinking that if the ball is not under A, it is under B. Nothing relevant has changed about *Jane's* epistemic situation—she hasn't gained or lost evidence, whether in reality or from your point of view—but your epistemic situation has been upgraded, and this change apparently produce a change in what iffy knowledge you are willing to credit Jane with—or so it seems if you want to say (9) is marked. This line of thinking would suggest that iffy knowledge ascriptions are not informationally persistent: you might accept such an ascription in one state of information, but go on to reject that ascription under the impact of strict information growth.

5

6

I'll end with a brief comparison to embedded counterfactuals. Compare:

(7) Bill knows that if we're not at this cafe, we're at the one across the street.

(19) Bill knows that if we weren't at this cafe, we'd be at the one across the street.

It can feel to some like these are saying the same thing. I get the sense that in the target scenario, many would prefer (19). Maybe this is a sign that (7) does inherit the problematic presupposition of the indicative (to the effect that our not being at this cafe is an open possibility). If so, that might suggest that the preference for (19) is really about avoiding that presupposition.

One common idea is that indicative conditionals express things about knowledge states, evidence, or epistemic relations, whereas 'subjunctives' or counterfactuals express things about causal or modal dependence. But these embedded conditionals might help to bring out a different and rather more mundane function for the indicative/counterfactual contrast to be playing, one along the lines suggested by Stalnaker [1975]. Sometimes we have a need to describe knowledge states that leave open possibilities we have eliminated in our conversation, and we want to do so in a way that doesn't let those possibilities back in. In such cases, it might be that we reach for the 'subjunctive' mood—not necessarily because we're interested in ascribing to Bill knowledge of a counterfactual dependence rather than something more epistemic in character, but just because we don't want to get caught up in the presuppositions the mood of the embedded conditional would involve us in.[4]

On that note, suppose again we're in the cafe, but we see Bill checking out the cafe across the street for us *before* he comes over to the cafe we're in fact in. In that situation it's fine to say:

(21) Bill knows that if we're not at the cafe across the street, we're at this one.

The embedded antecedent is one of our presuppositions, so it's not a problem. But I take it this is a bit odd:

(22) ?Bill knows that if we weren't at the cafe across the street, we'd be at this one.

We could try to say that it is odd because it embeds a counterfactual we wouldn't assert; but then, (7) embeds an indicative we wouldn't assert, and we've been saying that doesn't cause too much trouble. With (22) the issue is perhaps that the presupposition of the counterfactual is not locally satisfied by Bill's knowledge state: (22) presumes that Bill has no epistemic alternatives where we aren't at the cafe across the street, but we aren't at that cafe, so he does.

[4] To test this, we'd need to ask how people react to things like:

(20) Jane knows that if the ball weren't under A, it would be under B.

in the epistemic situation set out for Jane as above.

References

David Beaver. *Presupposition and Assertion in Dynamic Semantics*. Studies in Logic, Language and Information. CSLI Publications, 2001.

Bob Beddor and Simon Goldstein. Believing epistemic contradictions. *Review of Symbolic Logic*, Online first:1–27, 2017.

Justin Bledin. Logic informed. *Mind*, 130(490):277–316, 2014.

Allan Gibbard. Two recent theories of conditionals. In W. Harper, R. Stalnaker, and G. Pearce, editors, *Ifs: Conditionals, belief, decision, chance, and time*, pages 211–47. D. Reidel, 1981.

Anthony Gillies. Epistemic conditionals and conditional epistemics. *Noûs*, 38(4):585–616, 2004.

John Hawthorne, Daniel Rothschild, and Levi Spectre. Belief is weak. *Philosophical Studies*, 173(5):1393–1404, 2016.

Irene Heim. On the projection problem for presuppositions. In Michael Barlow, Daniel P. Flickinger, and Michael T. Wescoat, editors, *Proceedings, Second West Coast Conference on Formal Linguistics*, pages 114–25, Stanford, California, 1983. Stanford Linguistics Association.

Irene Heim. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics*, 9(3):183–221, 1992.

Lauri Karttunen. Presuppositions of compound sentences. *Linguistic inquiry*, 4(2): 169–193, 1973.

Sarah Moss. *Probabilistic Knowledge*. Oxford University Press, 2018.

Paolo Santorio. Conditional excluded middle in path semantics. Pacific APA, 2018.

Robert Stalnaker. Indicative conditionals. *Philosophia*, 5(3):269–286, 1975.

Robert Stalnaker. *Context*. Oxford University Press, Oxford, 2014.

Frank Veltman. Defaults in update semantics. *Journal of Philosophical Logic*, 25(3): 221–61, 1996.

Seth Yalcin. Epistemic modals. *Mind*, 116(464):983–1026, 2007.

Seth Yalcin. Conditional belief and conditional assertion. 2018.