Iffy Knowledge

May 21, 2022

8th CSLI Workshop on Logic, Rationality, and Intelligent Interaction
Stanford University

Seth Yalcin
yalcin@berkeley.edu

1. *Puzzles of iffy knowledge*

1.1 *Cafe*

My two favorite cafes 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.

Now 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..."

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

But this is puzzling:

i. **Failure of entailment?** Apparently, (3) $\not\models$ (2). We can take (3) for granted without being under any rational pressure to accept (2).

   ✗ **Iffy entailment.** $K(\phi \to \psi) \models \phi \to \psi$

ii. **Failure of factivity?** (3) doesn't seem to presuppose the conditional it embeds. In general, it's hard to understand what it would be to presuppose this conditional.

Philosophers often use *factive* to talk about a category of attitudes. Linguists usually use it to pick out presuppositional property of certain verbs and related constructions. I'm using it in the latter way.

1.2 *Marbles*

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

Thanks here to discussion with Kai von Fintel.

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

But we wouldn't be disposed to say this is true:

(5) If it's not under A, it's under B.

—though, since it's under A, we can agree:

(6) It's under A or B.

Pressure for the material conditional analysis, which says (5) and (6) are equivalent.

1.3 *Epistemic Sly Pete*

Turn this now into something like a Sly Pete-style case.[1] Let Carl play the game too, in isolation from Jane. The marble is still under A. We make Carl guess. He picks B. "Nope, it's not under B. It's either under A or C. Guess again!"

[1] Gibbard [1981]. See also Stalnaker [2014], Perl [forthcoming].

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

Surprisingly, speakers judge the following to be an acceptable description of the situation:

(8) 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.

Given **Iffy entailment**, this sentence would imply a counterexample to **Conditional noncontradiction**. Further reason to reject **Iffy entailment**.

Conditional noncontradiction. $(\phi \to \psi) \wedge (\psi \to \neg\psi) \vDash \bot$

1.4 *Second-order multi-agent iffy knowledge*

If we are in position to say things like (4) and (7), we presumably can also be in position to *say that we know* those things:

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

(10) We know that Carl knows that if it's not under A, it's under C.

But this is less comfortable:

(11) ??We know that if it's not under A, it's under B; and we know that if it's not under A, it's under C.

Certainly we're not in position to assert the conditional contradiction 'If it's not under A, it's under B; and if it's not under A, it's under C'. It would seem odd if (11) were somehow more assertable.

This all looks like a counterexample to:

✗ **Multiagent transparency.** $K_A K_B \phi \vDash K_A \phi$

in the special case where ϕ is an indicative conditional.

This seems to draw out a kind of relativity in iffy knowledge.

2. *Dynamicness?*

Very hard to handle any of this with conditional propositions. The facts so far seem to hint at a basically dynamic, and Ramseyan, perspective on things. Take:

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

Represent Carl's epistemic state as a set of worlds. Crude procedure for interpreting (7):

> Take Carl's epistemic state, and temporarily add to it the information that it's not under A. Ask: Does the resulting state of information entail that it's under C? If so, then (7) is true.

Ramsey [1931]. I'm thinking of the essentially dynamic or local-context-based semantic implementation of the idea as in, e.g., Gillies [2004], Yalcin [2007].

If this is the procedure, it is easy to see how (7) squares with

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

So far so good. But we run into some further challenges when we ask about the unembedded case.

(5) If it's not under A, it's under B.

Suppose we are presupposing that is under A. Following the Ramseyan/dynamic line, the question whether we accept this is turns on what happens to our context set when we add the information that it's not under A. But on the face of it, it seems what happens is that we get the empty information state, where everything is trivially true. But we want to predict that (5) isn't accepted by us, not that it trivially is accepted.

Needed to explain the data: a dose of dynamicness, but also a richer conception of information states, and of knowledge in particular.

3. *Nontriviality of counterepistemic iffy knowledge*

We are brushing up against the question what to say about the truth of indicative conditionals where the antecedent is impossible, or—what can seem close enough to that—false relative to the information known in the context (epistemically impossible). One familar idea is that in that case, any conditional built with the antecedent is true:

> ✗ **Materiality 1.** $\neg\phi \vDash \phi \to \psi$

The $\supset$ analysis of $\to$ validates this. This feature gets called "paradoxical" because of the way it lets us trivially derive what look like very nontrivial conditionals. Here I want to focus on this point that it would get the wrong result under knowledge operators. These can be true:

(12) Ivano knows that Shakespeake wrote *Hamlet*.

Cf. Ciardelli [2020].

(13) Ivano knows that if Shakespeake didn't *Hamlet*, someone else did.

While this is just false:

(14) Ivano knows that if Shakespeake didn't *Hamlet*, Betty White did.

Epistemic claim: that (12) is true doesn't imply that Ivano is in position to know just anything of the form *if Shakespeake didn't write Hamlet...* (13) reports a further substantive achievement of Ivano's knowledge state, going beyond anything the truth of (12) would put Ivano in position to know.

Discussing empty names and nonexistence, Yablo [2020] makes the same point.[2]

Let's say that in these cases, Ivano and Steve have **counterepistemic iffy knowledge**: they have iffy knowledge, plus they know the if-part is false.

[2] He observes that (15) and (16) are true, but (17) isn't:

(15) Steve knows Holmes doesn't exist.

(16) Steve knows that if Holmes does exist, he's not in this room.

(17) Steve knows that if Holmes exists, there are planets closer to the Sun than Mercury.

4. *Indicatives are not mighty*

Counterepistemic iffy knowledge can seem surprising, partly because it is common to think:

✗ **Mighty indicatives.** $\phi \rightarrow \psi \vDash \Diamond\phi$

Mighty indicatives for instance would explain:

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

But as Ciardelli [2020] has argued, **Mighty indicatives** is not correct.

(18) A: Shakespeare wrote *Hamlet*.
B: Agreed. But if he didn't, did someone else?
A: Well, obviously. But I'm sure it was Shakespeare.

An analogous sort of example is discussed in Stalnaker [2005].

Another example from Ciardelli: "If we win this match, we might win the World Cup. But that's not gonna happen."

Arguably it's common ground here that Shakespeare wrote *Hamlet*, and that if he didn't, someone else did. But it's not common ground that it might be someone else who wrote it.

The failure of **Mighty indicatives** is if anything *more* obvious in epistemic contexts. Again, these are both true:

(19) Steve knows that Holmes doesn't exist.

(20) Steve knows that if Holmes exists, he's not one of us.

But we don't we want the latter to make for an easy step to:

(21) Steve knows that Holmes might exist.

If **Mighty indicatives** is indeed wrong, we need another account of what is wrong with things like (2).

5. *Graded knowledge states*

To account for the substantivity of counterepistemic iffy knowledge, let's say that a state of knowledge fixes, not just a set of worlds, but something more structured—a set of worlds plus a total preordering over the set, reflecting something like plausibility.

The picture is: there's the elite "best" epistemic alternatives, which settle your outright factual knowledge, and which settle what you know *might* be the case. But the alternatives to this elite class are not all on epistemic par. Some are more plausible than others. And counterepistemic iffy knowledge ascriptions are ways of expressing things about how things are with an agent's less-than-elite epistemic alternatives.

Ciardelli [2020], drawing on Grove [1988], Lewis [1973], has proposed modeling states of conversation as graded information states. I'm saying: apply this idea in the semantics for 'knows'. So here's an adjusted picture, illustrating with:

(22) Steve knows that if Shakespeare didn't write *Hamlet*, someone else did.

Graded Ramseyan truth-condition for iffy knowledge.

(22) is true just in case when we update Steve's knowledge state with the information that Shakespeare didn't write *Hamlet*—when we restrict attention to the (as it happens, non-best) epistemic alternatives where that is so—the best *of these* are worlds where someone else did.

That delivers a substantive picture of counterepistemic iffy knowledge. This is all rather in the spirit of Moss [2018].

Much epistemology left to do.

6. *Loose end: Factivity*

None of the above really helps with factivity challenge. There seems to be a big problem here for any standard semantic account of the source of the presupposition of factives—any theory that tries to write into the semantic value of 'knows' that it presuppose its complement.

(Note: I wouldn't say counterepistemic knowledge ascriptions presuppose *nothing*. They seem to presuppose at least the corresponding material conditional.)

Appendix: One implementation

Ciardelli [2020] upgrades the system in Yalcin [2007] with graded information states. I just want to add a knowledge operator to his upgraded system.

We have a set of atomic sentences P. Where p ranges over P, L_0 is:

$$\alpha ::= p \mid \neg\alpha \mid \alpha \wedge \alpha$$

Extend this language of factual sentences to L by adding epistemic modals, indicative conditionals, and a knowledge operator:

$$\phi ::= \alpha \mid \alpha \rightarrow \phi \mid \Diamond\alpha \mid \neg\phi \mid \phi \wedge \psi \mid K\phi$$

Semantics-sketch: A **model** is $\langle W, V, K\rangle$, where: W is a (for simplicity finite) set of worlds; V maps elements of P and worlds to truth values; K is a function from worlds to **graded information states**.

A **graded information state** is a pair $s = \langle D_s, \leq_s\rangle$ where D_s is a set of worlds and $\leq_s$ is a total pre-ordering of D_s. (Rough gloss: "at least as plausible as".) $\mathsf{Best}(s)$ is the set of $\leq_s$-minimal elements in D_s.

We model both conversational states and epistemic states as graded. If s is an epistemic state for x, $\mathsf{Best}(s)$ is the set of x's epistemic alternatives—what might be the case according to x. K is a mapping from worlds to epistemic states, so we want to require that $w \in \mathsf{Best}(K(w))$—and probably more, depending on your preferred modal logic of knowledge.

Let's explain **updating a graded state** (with factual information), a notion that will matter for interpreting antecedents. If $s = \langle D, \leq\rangle$ is a graded state and $\alpha \in L_0$, the **update** of s with α, written $s[\alpha]$, is just the state $\langle D_\alpha, \leq_\alpha\rangle$ where D_α is the set of α-worlds in D and $\leq_\alpha$ is the restriction of $\leq$ to D_α.

We also need the idea of a sentence's being **accepted** (**supported, incorporated**) by a graded information state. Crucially, this is a generally matter of what's the case at the *best* worlds in s:

Acceptance (def). $s \vDash \phi$ iff $\forall w \in \mathsf{Best}(s) : w, s \vDash \phi$

Then here's a semantics:

$w, s \vDash p$ iff $w(p) = 1$

$w, s \vDash \neg\phi$ iff $w, s \nvDash \phi$

$w, s \vDash \phi \wedge \psi$ iff $w, s \vDash \phi$ and $w, s \vDash \psi$

$w, s \vDash \Box\psi$ iff $s \vDash \alpha$

$w, s \vDash \alpha \rightarrow \phi$ iff $s[\alpha] \vDash \phi$

$w, s \vDash K\phi$ iff $K(w) \vDash \phi$ (i.e., iff $\forall w' \in \mathsf{Best}(K(w)) : w', K(w) \vDash \phi$)[3]

[3] Further issues arise when a belief operator is added.

References

Ivano Ciardelli. Indicative conditionals and graded information. *Journal of Philosophical Logic*, 49(3):509–549, 2020.

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.

Adam Grove. Two modellings for theory change. *Journal of Philosophical Logic*, pages 157–170, 1988.

David K. Lewis. *Counterfactuals*. Blackwell, Malden, 1973.

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

Caleb Perl. How to outfox Sly Pete. In William Dunaway and David Plunkett, editors, *Meaning, Decision, and Norms: Themes from the Work of Allan Gibbard*. Maize Books, University of Michigan, forthcoming.

Frank Ramsey. General propositions and causality. In *The Foundations of Mathematics and Other Logical Essays*. Routledge and Kegan Paul, London, 1931.

Robert Stalnaker. Conditional propositions and conditional assertions. In J. Gajewski, V. Hacquard, B. Nickel, and S. Yalcin, editors, *New Work on Modality*, pages 207–31. MIT Working Papers in Linguistics, v. 51, 2005.

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

Stephen Yablo. Nonexistence and aboutness: The bandersnatches of Dubuque. *Crítica. Revista Hispanoamericana de Filosofía*, 52(154):77–100, 2020.

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