Comments on Moss, 'Epistemology Formalized'

Formal Epistemology Workshop, USC, 5/19/2011

Seth Yalcin
yalcin@berkeley.edu

Moss's idea:

Properties of one's credence distribution can constitute knowledge.

1 Factivity

A key linguistic property of *knows*:

ENTAILING. $K_x\phi$ **entails** ϕ.

It's a wide-open, high-level empirical question of natural language semantics how best to formalize **entails**. Some approaches:

Tarskian: Sentences semantically determine functions from some class of points of evaluation (within some class of models) to truth values. Entailment is a matter of truth preservation.

Dynamic: Sentences semantically determine operations on information states. Entailment preserves informational update potential.

Informational: Sentences semantically determine constraints on states of information. Entailment is a matter of the constraints' expressed by the premises subsuming the constraint expressed by the conclusion.

As Moss notes, by "factive", philosophers often have in mind (i) ENTAILING + (ii) a specific (Tarskian) hypothesis about how entailment is to be defined, roughly:

FACTIVE$_1$. for all w: $[\![K_x\phi]\!]^w = 1 \ \supset \ [\![\phi]\!]^w = 1$.

Three points:

- It's no objection *per se* if a theory fails to predict that *knows* is FACTIVE$_1$. What theory has to predict is ENTAILING. (Here I just echo Moss.)
- A similar response can be made to the objection that knowledge must be a relation to *truths*. Here's a datum in the vicinity:

  ENTAILS TRUE. $K_x\phi$ **entails** `It's true that` ϕ.

  One doesn't need a Tarskian notion of entailment to cover this data point.
- Moss's positive story works with something like informational entailment. That is a natural choice for her project. But it's worth noting that any of the above formal models of entailment, and probably others, can in principle be made compatible with Moss's idea.

2 Nonfactualism about knowledge

Moss holds that the contents of some knowledge states—especially, those expressed with probability operators and indicative conditionals—can be represented as conditions on probability spaces.[1] This entails a view we could call:

content probabilism: Some contents are representable as (nontrivial) constraints on probability spaces.

Content probabilism about a domain of discourse calls for a sort of nonfactualism about that domain of discourse, insofar as it requires thinking of the sentences of the discourse as not characterizing a state of the world, or a state of a position in the world (i.e., as having *truth-conditions*, in one usual sense).

This brings us to a simple point: If there is anything nonfactual/expressivist about the complement of *knows*, then—thanks to the entailingness of *knows*—this nonfactual/expressivist character will percolate up to the whole ascription. So we should read Moss as defending a kind of nonfactualism about knowledge states/expressivism about knowledge ascriptions.

This may shed light on some initially puzzling features of the account. A case:

Suppose a pair of fair dice are rolled. At t, let it be the case that A and B both know that the dice probably did not come up snake eyes. Let them know it in the same way, and let them each know that the other knows it, too. Thus at t, we would say:

(1) A knows that B knows that the dice probably did not come up snake eyes.

Now suppose at $t+1$, A learns (knows) that the dice in fact came up snake eyes.

In this kind of case, various surprises ensue on Moss's picture.

- It ceases to be the case that B knows that the dice probably did not come up snake eyes, since this is incompatible with A's knowing that the dice came up snake eyes.
- Consequently, A ceases to possess the knowledge ascribed to her in (1)—despite the fact that A seems not to have lost any information about how B is epistemically situated in the world (and B's epistemic situation in the world has not changed).
- We are unable to ascribe the knowledge A once had by placing a past tense on the attitude verb:

[1] Not really true; she takes them to be "constraints on credences" but doesn't take a stand on how to model credences. But I'll work with the probability space representation anyway, to fix ideas.

(2) A once knew that the dice probably did not come up snake eyes.

For (2) entails 'The dice probably did not come up snake eyes', which is incompatible with our stipulation that A learned the dice came up snake eyes.

3 Tensions with classical logic

Another way into things. This pattern is classically valid:

Remixed MT

$\neg\phi \rightarrow \neg\psi$

ψ

$\therefore \phi$

On Moss's view, Remixed MT appears subject to failure in the following kind of case:

$\neg\phi \rightarrow (\neg K_x\text{PROBABLY } \phi)$

$K_x\text{PROBABLY } \phi$

$\therefore \phi$

(Instances of the major premise here are virtual tautologies for Moss, thanks to ENTAILING + the semantics of the indicative.) Example:

(P1) If it is not raining, then it not the case that John knows it is probably raining.

(P2) John knows that it's probably raining.

(C) It is raining.

Can you be in a nontrivial state of information that accepts the premises, while (rationally and after full reflection) failing to accept the conclusion? Classical logic: no. Moss: yes. Who's right? I can see intuitions pulling in both directions.

—Or does Moss not say yes? It does seem like she should. Again, she thinks (P1) is essentially trivial. If only a triviality plus (P2) constitute one's information, how could one rationally conclude (C), something well beyond (P2)?

On the other hand, the conjunction of (P1) and (P2) leads to a puzzle. (P1) very plausibly entails/presupposes:

(3) It might not the case that John knows it is probably raining.

and this is uncomfortable to conjoin with (P2):

(4) ? John knows that it's probably raining, and that might not be true.

So, two related questions for Moss:

- Is the inference from (P1), (P2) to (C) valid or invalid?
- Are the premises jointly knowable?

  If no: why not?

  If yes: how should the defect in (4) be explained?

4 Characterizing safety

Moss says we can replace (Safe$_1$) with (Safe$_2$):

(Safe$_1$) For all cases α and β, if β is close to α and in α one knows that C obtains, then in β one does not falsely believe that C obtains.

(Safe$_2$) For all cases α and β, if β is close to α and in α one knows that C, then the following is not the case in β: that one believes that C and it is not the case that C.

I think to do this, she must pun on "cases".

Cases in the sense of (Safe$_1$) are centered worlds. In (Safe$_2$), however, they need to be the sorts of things that sentences which express constraints on probability space can be true or false (or hold or not hold, etc.) with respect to. That, presumably, is a probability space—or anyway, some kind of state of information. But if cases are states of information, then the import of (Safe$_2$) is quite different than that of (Safe$_1$), and they each involve different concepts of closeness. (Safe$_1$) cashes safety out in terms of what's going on at metaphysically nearby possibilities. (Safe$_2$) cashes safety out in terms of what's would be accepted on some kind of minimal revision to a given initial state of information. Is that what we really have in mind by safety?