Honors Logic, Lecture 18 - Nature of Epistemic Logic

Brian Weatherson

11/3/22

Models

So far we've mostly talked about **models**, which have three parts.

1. A set of worlds, which are really just featureless points.
2. An r relation on those worlds.
3. An assignment v of atomic sentences to worlds, intuitively meaning that the atomic is true at those points.

Frames

A **frame** is just a model without the last part. That is, it just has

1. A set of worlds.
2. An r relation on those worlds.

Truth on a Frame

Say that a sentence A is true on a frame $\langle W, R \rangle$ iff **any** way of adding a V and turning the frame into a model will make A true at every point in the model.

Example

- Consider the frame where the points are the real numbers, and xry holds iff $x \geq y$.
- And consider the sentence $\Box p \supset p$.
- Is that true on the frame?

Example

Yes!

- It doesn't matter what set you assign to p, at any point, either p is in the set or it isn't.
- If it is, then $\Box p \supset p$ is true at the point, since p is true.
- If it is not, then $\Box p$ is false at the point, since one point (namely that very point) that the point can see makes p false.
- And that implies $\Box p \supset p$ is true at the point.

Another Example

Same frame, but consider now the sentence $\Box p \supset \Box\Box p$.

- For any point, we can consider the question, is p true at that point and all earlier points?
- If so, $\Box\Box p$ will be true, since anywhere you can get in two steps from the point will make p true.
- If not, $\Box p$ will be false, since somewhere you can get makes p false.
- Either way, $\Box p \supset \Box\Box p$ will be true.

Characteristic Sentences

Consider not just an individual frame, but a whole set of frames. A sentence A **characterises** that set if the following two things hold.

- At **any** model built from any frame in the set, A is true.
- For any frame not in the set, there is **some** point in **some** model built on the frame where A is false.
- So there is this claim about A that is true at **all and only** frames in the set.

Charactistic Sentences

I'm not going to go over the proofs of all these I think (unless we have more time than I expect), but it turns out that there are quite a lot of sets of frames that have characteristic sentences in this (rather demanding) sense.

Charactistic Sentences

- The class of **reflexive** frames is characterised by $\Box p \supset p$.
- The class of **symmetric** frames is characterised by $p \supset \Box\Diamond p$.
- The class of **transitive** frames is characterised by $\Box p \supset \Box\Box p$.
- These can be mix and matched: The class of **reflexive, transitive** frames is characterised by $(\Box p \supset p) \wedge (\Box p \supset \Box\Box p)$.

Philosophical Significance

These results mean we can go back and forth easily between thinking about what sentences are logical truths (for a particular interpretation of Box), and what properties the relevant possibility relation has.

- And this is useful, because sometimes it is much more intuitive to think about one than the other.

Thinking about Possibility

When we think about what these models might really mean, the r relation is a kind of possibility.

- wrx means that if w is how things actually are, then x is a way things might be, in the sense of might that we care about.

Thinking about Possibility

When we think about what these models might really mean, the r relation is a kind of possibility.

- wrx means that if w is how things actually are, then x is a way things might be, in the sense of might that we care about. The point of the results in chapter 4 is that we now have two ways to think about these possibility claims.

1. We can look at them directly, and this may tell us something about certain modal sentences.
2. Or we can look at the sentences, and that can tell us something about the possibility claims.

Epistemic Modality

I want to walk through a bit how this plays out in contemporary debates in epistemology (the study of knowledge).

- We will treat $\Box A$ as meaning *Hero knows that A*.
- So $\Diamond A$ means *For all hero knows, A*, or (roughly) *A might be true (from Hero's perspective)*.
- And wrx means *If w is actual, then x is possible for Hero*. That is, if w is actual, then for all Hero knows, they are actually in x.

Reflexivity

Remember there is a tight connection between these two claims.

1. r is reflexive, i.e., wrw for all w.
2. It is always true that $\Box A \supset A$.

Reflexivity and Epistemic Modals

Both of these look like they should hold.

1. If Hero is in world w, then for all she knows, she could be in world w

Reflexivity and Epistemic Modals

Both of these look like they should hold.

1. If Hero is in world w, then for all she knows, she could be in world w.
2. If Hero knows that A, then it is the case that A.

Symmetry

Remember there is a tight connection between these two claims.

1. r is symmetric, i.e., if wrx then xrw.
2. It is always true that $A \supset \Box\Diamond A$

Symmetry

Remember there is a tight connection between these two claims.

1. r is symmetric, i.e., if wrx then xrw.
2. It is always true that $A \supset \Box\Diamond A$.

Question:

- Should this hold for the epistemic interpretation of $\Box$?

Symmetry and Epistemic Modals

This does not look particularly plausible. Let p be something that's actually true, but which Hero believes is false. (Hero could be anyone, so they could have false beliefs.) We'll start thinking directly about the models.

Symmetry and Epistemic Modals

- Let w_a be the actual world, and w_b the world in which everything is exactly as Hero thinks it is (in all respects).
- Then $w_a r w_b$, since w_b is surely possible for Hero from the actual perspective.
- But if she were in w_b, then w_a would not be possible, because she would know that p is false, and in w_a it is true.
- So $w_a r w_b$ but not $w_b r w_a$, showing symmetry fails.

Symmetry and Epistemic Modals (cont)

Now think about the axiom

- $p \supset \Box\Diamond p$

Remember Hero thinks p is false, but it's actually true. Is it the case that $\Box\Diamond p$

Symmetry and Epistemic Modals (cont)

Now think about the axiom

- $p \supset \Box\Diamond p$

Remember Hero thinks p is false, but it's actually true. Is it the case that $\Box\Diamond p$. Presumably not. This means that Hero knows that p might be true, but she thinks p is false. So false beliefs are still a problem.

Transitivity

Remember there is a tight connection between these two claims.

1. r is transitive, i.e., if wrx and xry then wry.
2. It is always true that $\Box A \supset \Box\Box A$

Transitivity

Remember there is a tight connection between these two claims.

1. r is transitive, i.e., if wrx and xry then wry.
2. It is always true that $\Box A \supset \Box\Box A$.

Question:

- Should this hold for the epistemic interpretation of $\Box$?

Current Debate

- This is a much disputed question in the contemporary literature.
- Here's one way to think about it.
- Is there some state of affairs that (a) Hero knows does not obtain, but (b) would be possible if the world was some other way, and (c) for all Hero knows the world is that way?

Current Debate

- This is a much disputed question in the contemporary literature.
- Here's one way to think about it.
- Is there some state of affairs that (a) Hero knows does not obtain, but (b) would be possible if the world was some other way, and (c) for all Hero knows the world is that way?
- For a long time, philosophers (and computer scientists, economists etc) thought this was impossible, so they thought we should have $\Box A \supset \Box\Box A$

Current Debate

- This is a much disputed question in the contemporary literature.
- Here's one way to think about it.
- Is there some state of affairs that (a) Hero knows does not obtain, but (b) would be possible if the world was some other way, and (c) for all Hero knows the world is that way?
- For a long time, philosophers (and computer scientists, economists etc) thought this was impossible, so they thought we should have $\Box A \supset \Box\Box A$. But recently a number of philosophers have started thinking this isn't true.

Outside Philosophy

This is a philosophy course, so I'm going to focus on what happens in philosophy.

- But the issue has broader ramifications.
- Lots of disciplines use models that include a representation of what various agents (living or artificial) know at different points.
- And the standard way this is done doesn't even allow for the representation of transitivity failures.
- Indeed, it doesn't even allow for symmetry failures, which is a bigger problem.
- But the issue here is one that lots of theorists should worry about.

Margins of Error

Think about the following situation.

- Hero is in a large lecture - maybe intro philosophy.
- As often happens, they are a bit bored, and start estimating how many people are in the lecture theatre.
- They are good at this kind of estimation, but not perfect.
- But they are almost always within 10% of the correct count. That's their **margin of error**.
- Today they guess there are 200 students in the lecture, and (a little surprisingly) there are indeed exactly 200 students in the lecture.

Margins and Knowledge

What does Hero know?

- Presumably they don't know that there are precisely 200 students in the lecture.
- After all, they aren't usually this accurate.
- But they do know that the number of students in somewhere between 180 and 220, since they are accurate to within 10%.

Some worlds

Let w_n be the world where there are precisely n students in the lecture.

- So the actual world is w_{200}.
- And every world between w_{180} and w_{220} is possible.

Some worlds

Let w_n be the world where there are precisely n students in the lecture.

- So the actual world is w_{200}.
- And every world between w_{180} and w_{220} is possible.
- But the worlds outside that range are not possible - Hero knows they do not obtain.

Some worlds

Let w_n be the world where there are precisely n students in the lecture.

- So the actual world is w_{200}.
- And every world between w_{180} and w_{220} is possible.
- But the worlds outside that range are not possible - Hero knows they do not obtain. Now think about which worlds are possible from w_{220}.

Some worlds

Let w_n be the world where there are precisely n students in the lecture.

- So the actual world is w_{200}.
- And every world between w_{180} and w_{220} is possible.
- But the worlds outside that range are not possible - Hero knows they do not obtain. Now think about which worlds are possible from w_{220}.
- In particular, should we think $w_{220} r w_{230}$?

What's At Stake

If $w_{220} r w_{230}$ then transitivity fails.

- We have $w_{200} r w_{220}$.
- And we have, by hypothesis, $w_{220} r w_{230}$.
- But we don't have $w_{200} r w_{230}$, since Hero actually knows that there are less than 230 students in the lecture.

What's At Stake

If $w_{220} r w_{230}$ then transitivity fails.

- We have $w_{200} r w_{220}$.
- And we have, by hypothesis, $w_{220} r w_{230}$.
- But we don't have $w_{200} r w_{230}$, since Hero actually knows that there are less than 230 students in the lecture.

But if we don't have $w_{220} r w_{230}$, then transitivity holds here. And this kind of case might be the best case for a transitivity failure.

So What's Right?

Think about what Hero knows in the world where

1. There are actually 220 students in the lecture, and
2. She guesses there are 200 students.

So What's Right?

Think about what Hero knows in the world where

1. There are actually 220 students in the lecture, and
2. She guesses there are 200 students.

- On the one hand, her estimations are only accurate to within 10%, so it would be funny if she knew there were less than 230 students. That would mean she could rule out something that was within the margin of error of her estimation.

So What's Right?

Think about what Hero knows in the world where

1. There are actually 220 students in the lecture, and
2. She guesses there are 200 students.

So What's Right?

Think about what Hero knows in the world where

1. There are actually 220 students in the lecture, and
2. She guesses there are 200 students.

- On the other hand, she knows that she guessed 200. And she knows (we can assume) that she's almost always within 10% of the truth. While she does not know this is one of the cases where she is this accurate, it actually is such a case. So maybe she can knowingly deduce (from her guess) that the crowd size is between 180 and 220.

Other Applications

This might seem like a weird special case, but it comes up literally all the time.

- Consider any case where Hero uses a measuring device (scales to weigh something, a gauge in their car to measure speed, or anything of the sort) that has a known margin of error m.
- And consider what happens when the real value is a, and the displayed value is b, with $|a - b| < m$.

What does Hero know?

- One option: Hero knows that real value is in $[b-m, b+m]$. The machine is working, and when it does, the user knows that the real value is within the margin of error of the diplayed value.
- Another option: Hero knows that the real value is in $[b-m-e, b+m+e]$, where $e = |a-b|$ is the error of the machine on this occasion. This way, Hero can never rule out anything closer to reality than the margin of error of the machine.
- And there are other options.

Implications for Epistemic Logic

- Option 1 strongly supports r being transitive; indeed if we can (somehow) rule out false beliefs, it even suggests that r is an equivalence relation.
- Option 2 means that r will not be transitive.
- And the case is hardly an obscure one; most measuring devices have a non-zero margin of error and on an occasion are not accurate to arbitrarily many decimal places.

A Verdict?

I’m not going to adjudicate this here. Indeed it’s an open debate in the literature.

- What I do think is that thinking about these models, and in particular how they can be used to model noisy estimation of point values, is a helpful way to approach the problem.
- It’s more helpful than thinking directly about whether someone could know something without knowing that they knew it.