

305 Lecture 11.7 - Frames and Philosophy

Brian Weatherson

Plan

- To go over why we should care about frames in philosophy classes, not just in logic classes.

Associated Reading

- Boxes and Diamonds, section 4.3 to 4.5 (though a lot of what I say here isn't in this book.)

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.

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 $\Box A \rightarrow A$.

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 $A \rightarrow \Box \Diamond A$

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 $A \rightarrow \Box \Diamond A$.

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.

- 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

- $p \rightarrow \Box \Diamond p$

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

- $p \rightarrow \Box \Diamond p$

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 $\Box A \rightarrow \Box \Box A$

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 $\Box A \rightarrow \Box \Box A$.

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 $\Box A \rightarrow \Box \Box A$

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 $\Box A \rightarrow \Box \Box A$. 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.

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 is 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 $w_{220}Rw_{230}$?

What's At Stake

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

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

What's At Stake

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

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

But if we don't have $w_{220}Rw_{230}$, 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.

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.

For Next Time

Next week, we'll look at how we prove things in modal logic.