Logic 2: Modal Logic

Lecture 9

Wolfgang Schwarz

February 16, 2022

University of Edinburgh

Kripke models for knowledge

Kripke models for knowledge

In epistemic logic, the box represents knowledge.

Possible-worlds analysis of knowledge

S knows that P iff P is true at all worlds compatible with S's knowledge.

In epistemic Kripke models, wRv means v is compatible with the agent's knowledge at w.

Kripke models for knowledge

More knowledge = fewer open possibilities

The duke has been murdered. There are four suspects: the gardener, the butler, the cook, and the maid.

Kripke models for knowledge

More knowledge = fewer open possibilities

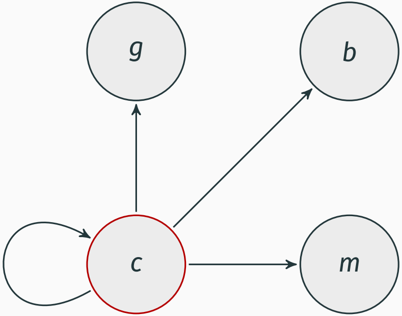
The duke has been murdered. There are four suspects: the gardener, the butler, the cook, and the maid.

The gardener has an alibi.

Kripke models for knowledge

More knowledge = fewer accessible worlds

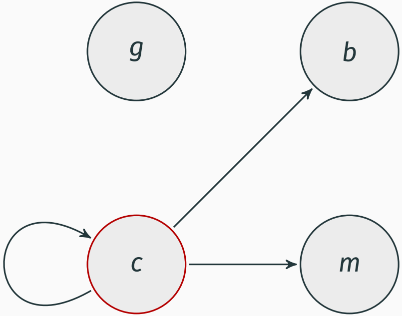
The cook has murdered the duke. The detective investigates.



Kripke models for knowledge

The cook has murdered the duke. The detective investigates.

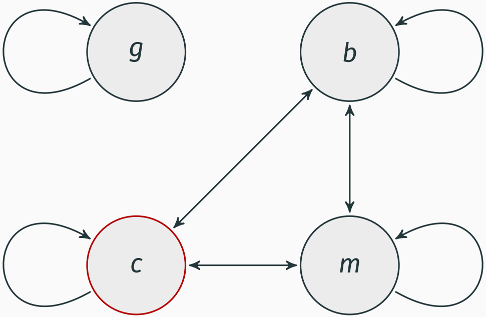
The gardener has an alibi.



Kripke models for knowledge

The cook has murdered the duke. The detective investigates.

The gardener has an alibi.



Kripke models for knowledge

Kripke semantics

$M, w \models \Box A$ iff $M, v \models A$ for all v such that wRv.
$M, w \models \Diamond A$ iff $M, v \models A$ for some v such that wRv.

In epistemic logic, we usually write the box as 'K'.

I write the diamond as 'M'.

Kripke semantics

$M, w \models \mathrm{K}A$ iff $M, v \models A$ for all v such that wRv.
$M, w \models \mathrm{M}A$ iff $M, v \models A$ for some v such that wRv.

Logical Omniscience

Logical Omniscience

A consequence of our semantics:

- Knowledge is **closed under known consequence**: $\mathsf{K}A, \mathsf{K}(A \to B) \models \mathsf{K}B$.
- Knowledge is **closed under logical consequence**: If $A \models B$ then $\models \mathsf{K}A \to \mathsf{K}B$.

This seems wrong.

Response 1: Our semantics is only adequate for ideal agents.

Response 2: We are modelling **implicit knowledge**.

Logical Omniscience

Our ordinary concept of knowledge is ill-behaved.

Case 1.

Pierre learns that a city called “Londres” is the capital of the UK and that it was founded by the Romans.

Later, Pierre travels to London, and believes it was founded in the middle ages.

Does Pierre know that London was founded by the Romans?

Logical Omniscience

Our ordinary concept of knowledge is ill-behaved.

Case 2.

If ‘S knows that p’ is true, and q is trivially equivalent to p, can we infer that ‘S knows that q’?

- Claire knows that her arm is broken.
- Claire knows that her arm is fractured.

Logical Omniscience

Our ordinary concept of knowledge is ill-behaved.

Case 3.

(P) *Nobody knows the sentence labelled '(P)' on slide 11.*

If (P) is false, then somebody does know (P); that's impossible.

So (P) is true. So nobody knows (P).

???

Logical Omniscience

Fred Dretske's (1970) argument against (K):

1. I know that I have hands. $\mathsf{K}p$
2. I know that if I have hands then I'm not a brain in a vat. $\mathsf{K}(p \rightarrow \neg q)$
3. I do not know that I'm not a brain in a vat. $\neg\mathsf{K}\neg q$

$\models \mathsf{K}(A \rightarrow B) \rightarrow (\mathsf{K}A \rightarrow \mathsf{K}B)$

$\mathsf{K}(A \rightarrow B), \mathsf{K}A \models \mathsf{K}B$

Epistemic Accessibility

Epistemic Accessibility

A world v is epistemically accessible for an agent at w (wRv) iff

- the agent's knowledge at w is compatible with the hypothesis that v is the actual world;
- v might be the actual world, for all the agent knows;
- whatever the agent knows at w is true at v.

Can we be more informative?

Epistemic Accessibility

Proposal 0

A world v is epistemically accessible for an agent at w (wRv) iff the agent's evidence at w is compatible with v.

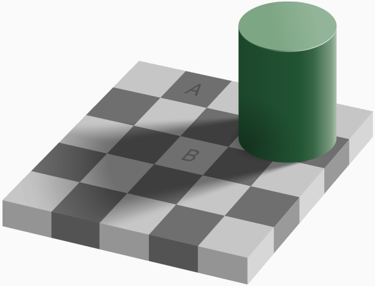
An agent's evidence is what her senses tell her.

Epistemic Accessibility

Proposal 1

A world v is epistemically accessible for an agent at w (wRv) iff the agent's evidence at w is compatible with v.

An agent's evidence is what her senses and memory tell her.



My senses tell me that square A is darker than square B.

But I don't believe that square A is darker.

$\mathrm{K}p \to p$ would become invalid.
$\mathrm{K}p \to \mathrm{B}p$ would become invalid.

Epistemic Accessibility

Proposal 2

A world v is epistemically accessible for an agent at w (wRv) iff the agent's sense experiences and memory at v are the same as at w.

- Is R reflexive? (For all w, wRw)
- Is R transitive? (If wRv and vRu then wRu)
- Is R symmetric? (If wRv then vRw)

Yes. R is an equivalence relation.

Epistemic Accessibility

Proposal 2

A world v is epistemically accessible for an agent at w (wRv) iff the agent's sense experiences and memory at v are the same as at w.

We get an S5 logic.

(**K**) $\mathrm{K}(A \rightarrow B) \rightarrow (\mathrm{K}A \rightarrow \mathrm{K}B)$

(**T**) $\mathrm{K}A \rightarrow A$

(**4**) $\mathrm{K}A \rightarrow \mathrm{K}\mathrm{K}A$

(**B**) $A \rightarrow \mathrm{K}\mathrm{M}A$

(**5**) $\mathrm{M}A \rightarrow \mathrm{K}\mathrm{M}A$

Epistemic Accessibility

Proposal 2

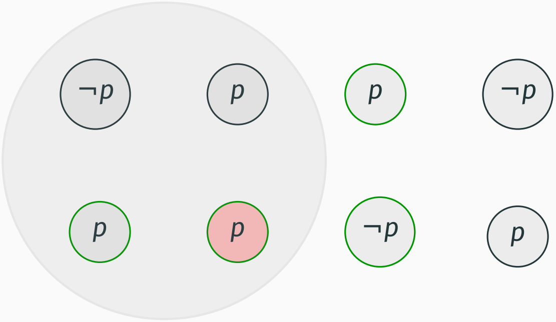
A world v is epistemically accessible for an agent at w (wRv) iff the agent's sense experiences and memory at v are the same as at w.

We also get scepticism about the external world.

Epistemic Accessibility

Proposal 3 (Lewis 1996)

A world v is epistemically accessible for an agent at w (wRv) iff the agent's sense experiences and memory at v are the same as at w and v is not properly ignored



Reflexivity, Seriality, Symmetry, Transitivity, Euclidity

Almost everyone wants the logic of knowledge to validate

$$(\mathbf{T}) \quad \mathsf{K}A \to A$$

So R should be reflexive. We then automatically get

$$(\mathbf{D}) \quad \mathsf{K}A \to \mathsf{M}A$$

Reflexivity, Seriality, Symmetry, Transitivity, Euclidity

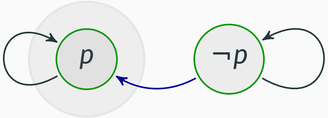
Should R be symmetric? Do we want (B) to come out valid?

$$(\mathbf{B})\quad A \rightarrow \mathsf{K}\,\mathsf{M}\,A$$

Suppose you falsely believe $\neg p$.

- p is true.
- You believe that you know $\neg p$.
- You don't believe that you don't know $\neg p$.
- You don't know that you don't know $\neg p$.
- $\Box\neg\Box\neg p$ is false.
- $\Box\Diamond p$ is false.

Also, this would lead to skepticism.



Reflexivity, Seriality, Symmetry, Transitivity, Euclidity

Positive Introspection:

$$(\mathbf{4}) \quad \mathrm{K}A \rightarrow \mathrm{K}\mathrm{K}A$$

Negative Introspection:

$$(\mathbf{5}) \quad MA \rightarrow \mathrm{K}MA$$

(5) corresponds to euclidity. Euclidity and reflexivity entail symmetry. So philosophers mostly reject (5).

(4) corresponds to transitivity. It is controversial.