

PIKSI-Logic Summer 2022

Tamar Lando

July 27, 2022

Tamar Lando

PIKSI-Logic Summer 2022

The Plan

- Propositional Modal Logic
 - Early Kripke semantics
 - Frames and Models
 - Semantic Consequence Relation(s)
 - Defining Classes of Frames
 - P-morphisms of Frames
 - Limits on Expressive Power of Basic Propositional Modal Language
- Epistemic Logic
 - Language and Semantics
 - Interpretation of Accessibility Relation
 - Epistemic Principles
- Wise Words from Kripke

Tamar Lando

PIKSI-Logic Summer 2022

Paradoxes of Material Implication

In a 1912 article, C.I. Lewis raises concerns about the paradoxes of material implication:

- $\neg p \rightarrow (p \rightarrow q)$
A false proposition 'implies' any other proposition.
- $p \rightarrow (q \rightarrow p)$
A true proposition is 'implied' by any proposition.

Lewis thought these paradoxes show that material implication does not capture anything like the ordinary meaning of 'implies'.

His idea was to formalize this notion with what is now called 'strict implication' $p \leftrightarrow q$, expressing that 'it is **not possible that** p is true and q is false.' He developed (with Langford) several axiomatic systems for strict implication in the early 1930's.

Tamar Lando

PIKSI-Logic Summer 2022

Modal Operators

Nowadays we typically take as primitives the modal operator $\Box$ or $\Diamond$ (or both).

On their alethic interpretations, these formalize the English-language modalities 'necessarily' and 'possibly.'

But people also use modal logic to formalize other operators:

- 1 Epistemic Logics: 'S knows that' or 'S believes that'
- 2 Temporal Logics: 'Henceforth' or 'At all points in the future'
- 3 Deontic Logics: 'It is obligatory that' or 'It ought to be the case that'

Tamar Lando

PIKSI-Logic Summer 2022

Non-Truth Functional

Arguably, the hallmark of such operators is that they are not truth-functional.

- The fact that p is true does not settle whether 'Necessarily p ' is true.
- The fact that p is true does not settle whether 'Trump believes that p ' is true.
- The fact that p is true does not settle whether 'Henceforth p ' is true.

Relational Structures

Some view modal logics as nice tools for formalizing our reasoning about necessity, knowledge, belief, time, and other notions.

According to a quite different view of the subject, modal logics are good tools for reasoning about relational structures.

That's because the standard semantics for modal languages involves 'Kripke frames,' which are built up of nodes and a binary relation(s) on nodes (think: graphs).

Such relational structures are ubiquitous, and it's of value to have a nice language for reasoning about them.

We can use a first-order language, but using a propositional modal language brings with it significant advantages.

The Propositional Modal Language

The basic propositional modal language $\mathcal{L}$ consists of a countable set PV of propositional variables, the classical connectives $\neg$, $\vee$, and the unary operator, $\Box$.

Definition (Formulas in $\mathcal{L}$)

- 1 Every propositional variable is a formula;
- 2 If φ and ψ are formulas, then $\neg\varphi$ and $\varphi \vee \psi$ are formulas;
- 3 If φ is a formula, then $\Box\varphi$ is a formula.

To this we can add familiar shorthand:

$$\varphi \wedge \psi := \neg(\neg\varphi \vee \neg\psi)$$

$$\varphi \rightarrow \psi := \neg\varphi \vee \psi$$

$$\varphi \leftrightarrow \psi := \neg(\neg\varphi \vee \neg\psi) \vee \neg(\varphi \vee \psi)$$

$$\Diamond\varphi := \neg\Box\neg\varphi$$

Truth Tables with Deleted Rows

How can we define a semantics (way of assigning truth values to formulas) for this modal language?

A simple, early semantics was given by Saul Kripke.

His idea was to take as models truth tables in which some rows are deleted, and one row is 'designated.'

	p	q	r
1	T	T	T
2	T	T	F
3	T	F	T
4	T	F	F
5	F	T	T
6	F	T	F
7	F	F	T
8	F	F	F

Rough Idea

The intuitive idea of the semantics is the Leibnizian thought that necessity is 'truth in all possible worlds.'

Think of each non-deleted row of the truth table as a crude representation of a possible world, and the designated row as crude representation of the actual world.

The formula $\Box\varphi$ is true if φ is true in all non-deleted rows.

Therefore $\Diamond\varphi$ is true if φ is true in some non-deleted row.

(Recall that $\Diamond\varphi := \neg\Box\neg\varphi$.)

Navigation icons

Satisfaction for Early Kripke Semantics

More formally, a model for a formula α is a pair $M = \langle G, K \rangle$, where K is a set of truth assignments to all propositional variables in α , and $G \in K$.

Satisfaction is a relation between a model, truth assignment, and formula.

Let $H \in K$. Then for all subformulas of α , satisfaction is determined as follows:

- $M, H \models p$ iff $H(p) = T$ where p is a propositional variable
- $M, H \models \neg\varphi$ iff $M, H \not\models \varphi$
- $M, H \models \varphi \vee \psi$ iff $M, H \models \varphi$ or $M, H \models \psi$
- $M, H \models \Box\varphi$ iff $M, J \models \varphi$ for all $J \in K$.

Navigation icons

Example Revisited

	p	q	r
1	T	T	T
2	T	T	F
3	T	F	T
4	T	F	F
5	F	T	T
6	F	T	F
7	F	F	T
8	F	F	F

Which of the following are true in the model pictured on the left, in the designated (pink) row?

- $\neg p \wedge q \wedge r$
- $\Box(q \vee r)$
- $\neg\Box q$
- $p \vee \Box q$
- $\Diamond(\neg p \wedge \neg q)$
- $\Diamond r \rightarrow \Box p$

Navigation icons

Some Validities

Question. Can you think of any formulas that come out true in every such model?

Examples:

$$\mathbf{K}: \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$$

$$\mathbf{D}: \Box p \rightarrow \Diamond p$$

$$\mathbf{T}: \Box p \rightarrow p$$

$$\mathbf{B}: p \rightarrow \Box\Diamond p$$

$$\mathbf{4}: \Box p \rightarrow \Box\Box p$$

$$\mathbf{5^*}: \neg\Box p \rightarrow \Box\neg\Box p$$

Do we *want* all of these to come out true in every model?

Navigation icons

Alethic and Epistemic Interpretations of T, 5

Let's focus on two of these and consider both alethic and epistemic interpretations.

T: $\Box\varphi \rightarrow \varphi$

- If necessarily φ , then φ .
- If Anna knows that φ , then φ .
- If Anna believes that φ , then φ .

5*: $\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$

- If φ is not necessary, then necessarily φ is not necessary.
- If Anna doesn't know that φ , then Anna knows that she doesn't know that φ .
- If Anna doesn't believe that φ , then Anna believes that she doesn't believe that φ .

Navigation icons

Moving to a more flexible semantics

Given the variety of interpretations of modal languages, we want a more flexible semantics – one that (at least in its most general implementation) doesn't validate all of the named formulas above.

In order to refute these formulas, we need to allow in 'more' models.

We'll consider now the 'mature' version of Kripke semantics.

Navigation icons

Kripke Models

Definition (Kripke Frame)

A *Kripke frame* is a pair

$$F = \langle W, R \rangle$$

where W is a non-empty set, and R is a binary relation on W .

Definition (Kripke Model)

A *Kripke model* is a triple

$$M = \langle W, R, V \rangle$$

where $\langle W, R \rangle$ is a frame, and $V : PV \rightarrow \mathcal{P}(W)$ is a function assigning to each propositional variable some subset of W . We say that the model M is built over the frame $F = \langle W, R \rangle$.

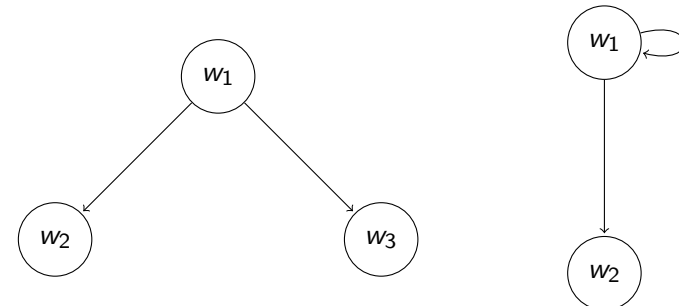
Navigation icons

Kripke Frames Pictorially

We sometimes think/speak of:

- W as a set of possible worlds
- R an 'accessibility relation'
- V a 'valuation'

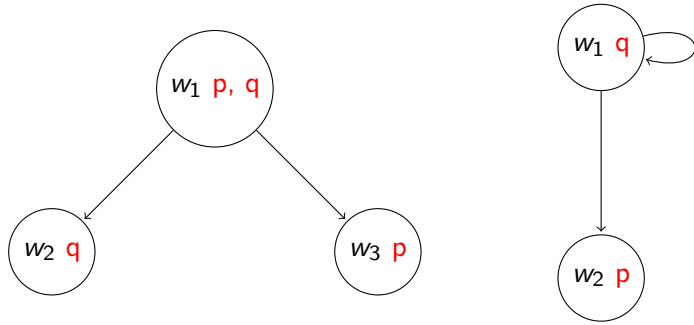
Pictorially, Kripke frames look like graphs:



Navigation icons

Kripke Models Pictorially

And Kripke models look like graphs with propositional information added to nodes:



Satisfaction

In Kripke semantics, the basic satisfaction relation holds between a model $M = \langle W, R, V \rangle$, a world $w \in W$, and a formula φ . It is defined recursively as follows.

- $M, w \models p$ iff $w \in V(p)$
- $M, w \models \neg\varphi$ iff $M, w \not\models \varphi$
- $M, w \models \varphi \vee \psi$ iff $M, w \models \varphi$ or $M, w \models \psi$
- $M, w \models \Box\varphi$ iff $M, v \models \varphi$ for **all** v such that wRv

As a consequence, we have the following clause for the $\Diamond$ -modality:

- $M, w \models \Diamond\varphi$ iff $M, v \models \varphi$ for **some** v such that wRv

Truth and Validity

Let $M = \langle W, R, V \rangle$ be a model, and $w \in W$. Here are some important truth notions in modal logic:

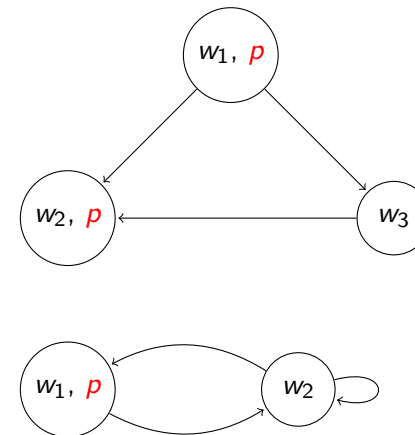
We say φ is true at world w if $M, w \models \varphi$.

We say that φ is true in M if φ is true at every world in M .

We say that φ is valid in the frame $F = \langle W, R \rangle$ if φ is true in every model built over F .

Finally, we say φ is valid in a class of frames if φ is valid in every frame in the class.

Examples



- $M, w_1 \models p \wedge \Diamond p$
- $M, w_1 \models \Diamond \neg p$
- $M, w_3 \models \neg p \wedge \Diamond p$
- $M, w_2 \models \Box p$
- $M, w_2 \models \Box \neg p$
- $M, w_1 \models p \wedge \Box \neg p$
- $M, w_2 \models \Diamond \neg p$
- $M, w_2 \models \Diamond \Diamond \neg p$
- $M, w_2 \models \neg \Box p$
- $M, w_2 \models \neg \Box \neg p$

Refuting $\mathbf{T}$ and $\mathbf{5}^*$

Let's see how the formulas $\mathbf{T}$ and $\mathbf{5}^*$ fare on this new semantics.
Can we refute them?

Consider the model, M , below:



Note that:

$$M, w_1 \models \Box p$$

$$M, w_1 \not\models p.$$

$$\text{So } M, w_1 \not\models \Box p \rightarrow p.$$

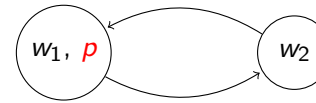
Question: can you refute $\mathbf{T}$ with just one world? Sure!



(Cont.)

What about $\mathbf{5}^*$?

Consider the model, N , below:



Note that:

$$N, w_1 \models \neg \Box p$$

$$N, w_2 \not\models \neg \Box p$$

$$\text{So } N, w_1 \not\models \Box \neg \Box p$$

$$\text{So } N, w_1 \not\models \neg \Box p \rightarrow \Box \neg \Box p.$$

Validities

Do any formulas come out valid in the mature Kripke semantics?

Examples:

- All tautologies. E.g.

- $(p \wedge q) \rightarrow p$

- $(\Diamond p \wedge \Diamond q) \rightarrow \Diamond p$

- **K**: $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$

- **M**: $\Box(p \wedge q) \rightarrow (\Box p \wedge \Box q)$

- **C**: $(\Box p \wedge \Box q) \rightarrow \Box(p \wedge q)$

Note: None of D , T , B , 4 , or 5 come out valid.

Early vs. Mature Kripke semantics

Aside: What is the relationship between the 'early' Kripke semantics (w/deleted rows of truth tables) and the 'mature' Kripke semantics?

Defining classes of frames

Now that we've developed a very flexible semantics for the basic propositional modal language, we may want to carve out restrictions on that semantics for particular applications.

For example, suppose we want to validate the formula:

$$\mathbf{T}: \Box\varphi \rightarrow \varphi$$

Is there a restricted class of frames in which this formula *is* valid?

Definition

A formula φ **defines** a class of frames $\mathcal{C}$ if:
 φ is valid in F if and only if $F \in \mathcal{C}$

(cont.)

What class of frames do the following formulas define:

$$\mathbf{D}: \Box p \rightarrow \Diamond p$$

$$\mathbf{T}: \Box p \rightarrow p$$

$$\mathbf{B}: p \rightarrow \Box\Diamond p$$

$$\mathbf{4}: \Box p \rightarrow \Box\Box p$$

$$\mathbf{5^*}: \neg\Box p \rightarrow \Box\neg\Box p$$

Answers:

- The class of serial frames.
- The class of reflexive frames.
- The class of symmetric frames.
- The class of transitive frames.
- The class of Euclidean frames.



Semantic Consequence*

There are multiple semantic consequence relations in modal logic, but the basic one is the following 'local' relation.

Definition (Local Consequence)

A set of formulas Γ implies a formula φ if for any model $M = \langle W, R, V \rangle$ and any $w \in W$, if $M, w \models \gamma$ for all $\gamma \in \Gamma$, then $M, w \models \varphi$.

Examples:

- $\Box p \wedge \Box q \models \Box(p \wedge q)$
- $\Diamond p, \Box(p \rightarrow q) \models \Diamond q$
- $\Diamond p, \Box p \rightarrow \Box q \not\models \Diamond q$
- $(\Box p \vee \Box q) \models \Box(p \vee q)$
- $\Box(p \vee q) \not\models \Box p \vee \Box q$



Consequence Relations for Different Classes of Frames*

Definition ($\models_{\mathcal{C}}$)

Let $\mathcal{C}$ be a class of frames. Then $\Gamma \models_{\mathcal{C}} \varphi$ if for every model $M = \langle W, R, V \rangle$ over a frame in $\mathcal{C}$, and every $w \in W$, if $M, w \models \gamma$ for all $\gamma \in \Gamma$, then $M, w \models \varphi$.

Let $\mathcal{C}_1$ be the class of reflexive frames. Note that while in the general Kripke semantics, $\Box p$ does not imply p ,

$$\Box p \models_{\mathcal{C}_1} p$$

Let $\mathcal{C}_2$ be the class of transitive frames. Note that while in the general Kripke semantics, $\Box p$ does not imply $\Box\Box p$,

$$\Box p \models_{\mathcal{C}_2} \Box\Box p$$

Let $\mathcal{C}_3$ be the class of transitive and Euclidean frames. Note that while $\Box\Box p \vee \Box\neg\Box p$ is not valid generally,

$$\models_{\mathcal{C}_3} \Box\Box p \vee \Box\neg\Box p$$



The right modal consequence relation?*

So we can begin to ask, what is the right modal consequence relation for understanding, e.g., knowledge? Belief? Time? And so on.

These are of course substantive philosophical questions, about which there is a lot of disagreement.

That other perspective...

Flipping the question on its head, we can also ask for a given class of frames (or property): Is there a modal formula that defines it?

- Is there a formula that defines the class of infinite frames?
- Is there a formula that defines the class of irreflexive frames?

Such questions begin to explore the expressive power of the (basic propositional) modal language.

How could we answer one of the above questions in the negative?

P-morphism of Frames

Definition

Let $F = \langle W, R \rangle$ and $F' = \langle W', R' \rangle$ be frames. We say that a function $f : W \rightarrow W'$ is a p-morphism of frames if

- (F) If wRv then $f(w)R'f(v)$;
- (B) If $f(w)R'v'$, then there is some $v \in W$ such that wRv and $f(v) = v'$.

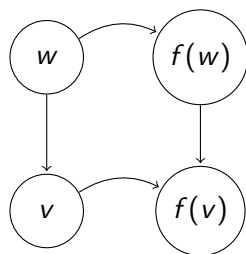


Figure: Forward

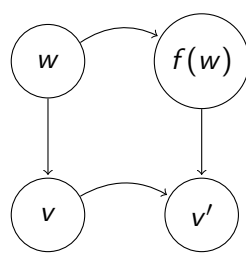
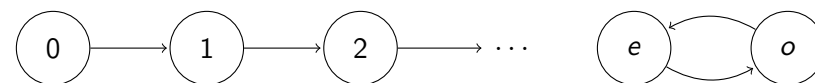


Figure: Backward

Examples of Frame P-morphisms

Example 1.



$$\text{Let } f(n) = \begin{cases} e & \text{if } n \text{ is even} \\ o & \text{if } n \text{ is odd} \end{cases}$$

Example 2.



$$\text{Let } f(w) = f(v) = u$$

Both functions defined above are p-morphisms of frames. And they are **surjective**: every world in the RHS frame is 'hit' by the function.

Proposition

If F and F' are frames, and $f : F \rightarrow F'$ is a surjective p -morphism, then any formula valid in F is valid in F' .

As a consequence of this simple theorem, we can discover some important limitations on the expressive power of the basic propositional modal language.

- No formula (in the basic propositional modal language) defines the class of infinite frames. Why?
- No formula (in the basic propositional modal language) defines the class of irreflexive frames. Why?

Epistemic Logics

Language of Epistemic Logics

The simplest language for epistemic logic is very similar to the basic propositional modal language, except here we take K instead of $\Box$ to be our modal operator.

Definition (Basic Language of Epistemic Logic)

The set of all formulas in $\mathcal{L}_K$ is defined inductively:

- 1 Every propositional variable is a formula;
- 2 If φ and ψ are formulas, then $\neg\varphi$ and $\varphi \vee \psi$ are formulas;
- 3 If φ is a formula, then $K\varphi$ is a formula.

We can say this more succinctly by putting:

$$\varphi := p \mid \neg\varphi \mid \varphi \vee \psi \mid K\varphi$$

where $p \in PV$.

Multiple Agents

This language allows us to reason about what a single agent knows.

But often we want to reason about multiple agents at a time – about what each one knows, what each knows about what the other knows (and so on).

We can expand the language to allow for this.

Let AG be some finite set (of ‘agents’).

Our new language includes a modal operator for each agent:

$$\varphi := p \mid \neg\varphi \mid \varphi \vee \psi \mid K_a\varphi$$

for $p \in PV$ and $a \in AG$.

Semantics for epistemic logics

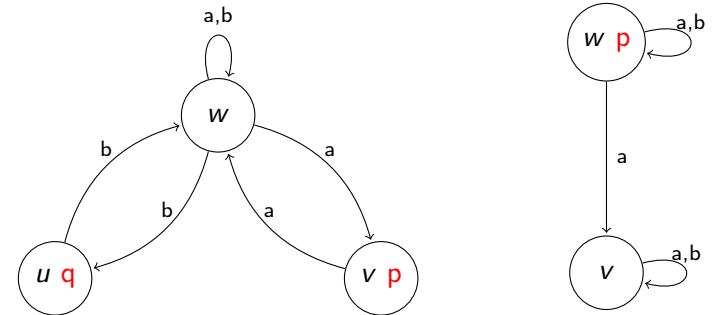
Models for this expanded language are like the Kripke models we've already studied, but here we have a separate accessibility relation for each agent/modal operator.

Suppose we are in a language that contains modal operators K_a and K_b for two agents. A model for this language consists of:

- A non-empty set W (of 'states')
- Binary relations $R_a, R_b \subseteq W \times W$
- A valuation function $V : PV \rightarrow \mathcal{P}(W)$

Again, we can picture such models as graphs (nodes and arrows), except that here arrows are labeled for different agents.

Models Pictorially



LHS: In world w , agent a sees w and v , but agent b sees w and u .

RHS: In world w , agent a sees w and v , but agent b sees only w .

Notice that we place no constraints here on the accessibility relation.

Satisfaction for Epistemic Logics

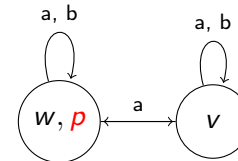
As before, the satisfaction relation holds between a model $M = \langle W, \{R_a\}_{a \in AG}, V \rangle$, a state $w \in W$, and a formula φ . It is defined recursively exactly as you would expect.

- $M, w \models p$ iff $w \in V(p)$
- $M, w \models \neg\varphi$ iff $M, w \not\models \varphi$
- $M, w \models \varphi \vee \psi$ iff $M, w \models \varphi$ or $M, w \models \psi$
- $M, w \models K_a\varphi$ iff $M, v \models \varphi$ for all v such that wR_av

Example

Example. Imagine that both Anna and Ben are looking at a box on the table. Anna is color blind and cannot tell the difference between red and green. Ben can. The box is green.

Let p be the proposition that the box is green. Then consider the model, M :

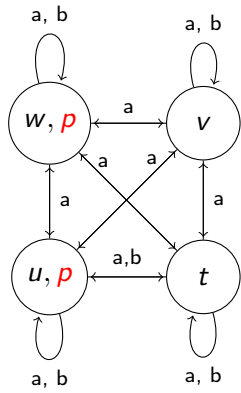


- $M, w \models K_bp$
- $M, w \not\models K_ap$
- $M, v \models K_b\neg p$
- $M, v \not\models K_a\neg p$
- $M, w \models K_a(K_bp \vee K_b\neg p)$.

According to our model, Anna knows that Ben knows whether the box is green! But what if Anna thinks that Ben might be color blind too?

(cont.)

We can fix this by adding additional worlds.



Note that:

$M, w \models K_b p \vee K_b \neg p$

$M, v \models K_b p \vee K_b \neg p$

$M, u \not\models K_b p \vee K_b \neg p$

$M, t \not\models K_b p \vee K_b \neg p$

Ben knows whether p in w, v .

Ben doesn't know whether p in u, t .

So at all worlds, Anna doesn't know if Ben knows whether p .

So Anna knows that she doesn't know if Ben knows whether p .

Navigation icons

How to Understand the Accessibility Relation

One way of understanding the accessibility relation in epistemic logics is as an **indistinguishability** relation.

An arrow from world w to world v indicates that the agent's information does not distinguish between w and v .

In the above example, because Anna is color blind, her information does not distinguish between a world in which the box is green and a world in which the box is, e.g., red.

An agent is represented as knowing a proposition φ in world w if φ is true throughout the worlds which she cannot distinguish from w .

So, e.g., since Anna cannot distinguish the world in which the box is green from the one in which it is red, she does not know that it is green.

Navigation icons

Properties of Indistinguishability Relations

This approach is not philosophically innocent.

Taking the accessibility relation to be a relation of indistinguishability ensures that it has certain properties:

- Every world is indistinguishable from itself, so the relation is reflexive.
- If w is indistinguishable from v , then v is indistinguishable from w . So the relation is symmetric.
- If w is indistinguishable from v , and v is indistinguishable from u , is w indistinguishable from u ? If so, the relation is transitive.

Thus, our models would validate the schemas corresponding to these classes of frames.

Navigation icons

Another Approach

A different approach to epistemic logics is to start by considering what general principles of knowledge one is committed to, and imposing constraints on the accessibility relation(s) accordingly.

Let's examine some of the schemas we saw above:

- **D:** $K_a \varphi \rightarrow \neg K_a \neg \varphi$
- **T:** $K_a \varphi \rightarrow \varphi$
- **4:** $K_a \varphi \rightarrow K_a K_a \varphi$
- **5*:** $\neg K_a \varphi \rightarrow K_a \neg K_a \varphi$
- **B:** $\varphi \rightarrow K_a \neg K_a \neg \varphi$

Navigation icons

Veridicality and Consistency

Let's take these in turn.

- **T:** $K_a\varphi \rightarrow \varphi$
If Anna knows that φ , then φ is true.
This axiom is uncontroversial.
We sometimes refer to this as veridicality or factivity of knowledge.
- **D:** $K_a\varphi \rightarrow \neg K_a\neg\varphi$
If Anna knows that φ , then she doesn't know $\neg\varphi$.
In other words, Anna does not know both φ and its negation.

Note that the second principle follows classically from the first. If Anna knew both φ and $\neg\varphi$, by factivity both φ and $\neg\varphi$ would be true, but that is not (classically) possible!



Positive and Negative Introspection

- **4:** $K_a\varphi \rightarrow K_aK_a\varphi$
If Anna knows φ , then she knows that she knows φ .
This is sometimes called the principle of positive introspection.
- **5*:** $\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$
If Anna doesn't know φ , then she knows that she doesn't know φ .
This is sometimes called the principle of negative introspection.

Both of these principles are controversial, and have been rejected by some epistemologists.



The Problem of Logical Omniscience

Whatever principles we adopt, if we model knowledge in Kripke models, we will at least end up validating all tautologies.

Moreover, if a formula φ is valid in a class of frames, then so is $\Box\varphi$ – or, if we're in the language of epistemic logics, $K_a\varphi$.

Taken together, this means that agents are modeled as knowing all tautologies – and indeed all formulas valid in the class of all Kripke frames.

It also means that agents are modeled as knowing all of the (classical) consequences of what they know.

These commitments don't seem to reflect the cognitive state of actual human reasoners.



Certainly the philosopher of 'possible worlds' must take care that his technical apparatus not push him to ask questions whose meaningfulness is not supported by our original intuitions of possibility that gave the apparatus its point.

- Saul Kripke



I wish I could have skipped college.

- Saul Kripke