# Modal Logic: Logics of Knowledge and Belief

Eric Pacuit, University of Maryland

November 15, 2023

# The "Problem" of Logical Omniscience

The rule

$$\mathrm{RK}_i \ \frac{(\varphi_1 \wedge \cdots \wedge \varphi_m) \to \psi}{(K_i\varphi_1 \wedge \cdots \wedge K_i\varphi_m) \to K_i\psi}$$

reflects so-called (*synchronic*) *logical omniscience*: the agent knows (at time $t$) all the consequences of what she knows (at $t$).

# The "Problem" of Logical Omniscience

The rule

$$\mathsf{RK}_i \ \frac{(\varphi_1 \wedge \cdots \wedge \varphi_m) \to \psi}{(K_i\varphi_1 \wedge \cdots \wedge K_i\varphi_m) \to K_i\psi}$$

reflects so-called (*synchronic*) *logical omniscience*: the agent knows (at time $t$) all the consequences of what she knows (at $t$).

Given this, there are two ways to view $K_i$: as representing either the idealized (implicit, "virtual") knowledge of ordinary agents, or the ordinary knowledge of idealized agents. For discussion, see

R. Stalnaker (1991). *The Problem of Logical Omniscience, I*. Synthese.

R. Stalnaker (2006). *On Logics of Knowledge and Belief*. Philosophical Studies.

# The "Problem" of Logical Omniscience

The rule

$$\mathsf{RK}_i \ \frac{(\varphi_1 \wedge \cdots \wedge \varphi_m) \rightarrow \psi}{(K_i\varphi_1 \wedge \cdots \wedge K_i\varphi_m) \rightarrow K_i\psi}$$

reflects so-called (*synchronic*) *logical omniscience*: the agent knows (at time $t$) all the consequences of what she knows (at $t$).

There is now a large literature on alternative frameworks for representing the knowledge of agents with bounded rationality, who do not always "put two and two together" and therefore lack the logical omniscience reflected by $\mathsf{RK}_i$. See, for example:

J. Y. Halpern and R. Pucella. 2011. *Dealing with Logical Omniscience: Expressiveness and Pragmatics*. Artificial Intelligence.

# Logical Omniscience

- From $\varphi \leftrightarrow \psi$ infer $K_i\varphi \leftrightarrow K_i\psi$

# Logical Omniscience

- From $\varphi \leftrightarrow \psi$ infer $K_i\varphi \leftrightarrow K_i\psi$
- From $\varphi \rightarrow \psi$ infer $K_i\varphi \rightarrow K_i\psi$

# Logical Omniscience

- From $\varphi \leftrightarrow \psi$ infer $K_i\varphi \leftrightarrow K_i\psi$
- From $\varphi \rightarrow \psi$ infer $K_i\varphi \rightarrow K_i\psi$
- $(K_i(\varphi \rightarrow \psi) \wedge K_i\varphi) \rightarrow K_i\psi$

# Logical Omniscience

- From $\varphi \leftrightarrow \psi$ infer $K_i\varphi \leftrightarrow K_i\psi$
- From $\varphi \rightarrow \psi$ infer $K_i\varphi \rightarrow K_i\psi$
- $(K_i(\varphi \rightarrow \psi) \wedge K_i\varphi) \rightarrow K_i\psi$
- From $\varphi$ infer $K_i\varphi$

# Logical Omniscience

- From $\varphi \leftrightarrow \psi$ infer $K_i\varphi \leftrightarrow K_i\psi$
- From $\varphi \rightarrow \psi$ infer $K_i\varphi \rightarrow K_i\psi$
- $(K_i(\varphi \rightarrow \psi) \wedge K_i\varphi) \rightarrow K_i\psi$
- From $\varphi$ infer $K_i\varphi$
- $K_i\top$

# Logical Omniscience

- From $\varphi \leftrightarrow \psi$ infer $K_i\varphi \leftrightarrow K_i\psi$
- From $\varphi \rightarrow \psi$ infer $K_i\varphi \rightarrow K_i\psi$
- $(K_i(\varphi \rightarrow \psi) \wedge K_i\varphi) \rightarrow K_i\psi$
- From $\varphi$ infer $K_i\varphi$
- $K_i\top$
- $(K_i\varphi \wedge K_i\psi) \rightarrow K_i(\varphi \wedge \psi)$

# Dealing with Logical Omniscience

- *Syntactic approaches*: an agent's knowledge is represented by a set of formulas (intuitively, the set of formulas she knows);

# Dealing with Logical Omniscience

- *Syntactic approaches*: an agent's knowledge is represented by a set of formulas (intuitively, the set of formulas she knows);
- *Awareness*: an agent knows $\varphi$ if she is aware of $\varphi$ and $\varphi$ is true in all the worlds she considers possible;

# Dealing with Logical Omniscience

- *Syntactic approaches*: an agent's knowledge is represented by a set of formulas (intuitively, the set of formulas she knows);
- *Awareness*: an agent knows $\varphi$ if she is aware of $\varphi$ and $\varphi$ is true in all the worlds she considers possible;
- *Algorithmic knowledge*: an agent knows $\varphi$ if her knowledge algorithm returns "Yes" on a query of $\varphi$; and

# Dealing with Logical Omniscience

- *Syntactic approaches*: an agent's knowledge is represented by a set of formulas (intuitively, the set of formulas she knows);
- *Awareness*: an agent knows $\varphi$ if she is aware of $\varphi$ and $\varphi$ is true in all the worlds she considers possible;
- *Algorithmic knowledge*: an agent knows $\varphi$ if her knowledge algorithm returns "Yes" on a query of $\varphi$; and
- *Impossible worlds*: an agent may consider possible worlds that are logically inconsistent (for example, where $p$ and $\neg p$ may both be true).

# Dealing with Logical Omniscience

- *Syntactic approaches*: an agent's knowledge is represented by a set of formulas (intuitively, the set of formulas she knows);
- *Awareness*: an agent knows $\varphi$ if she is aware of $\varphi$ and $\varphi$ is true in all the worlds she considers possible;
- *Algorithmic knowledge*: an agent knows $\varphi$ if her knowledge algorithm returns "Yes" on a query of $\varphi$; and
- *Impossible worlds*: an agent may consider possible worlds that are logically inconsistent (for example, where $p$ and $\neg p$ may both be true).

*Non-Normal Modal Logics*

# Dealing with Logical Omniscience

- *Syntactic approaches*: $\mathcal{M}, w \models K_i\varphi$ iff $\varphi \in \mathcal{C}_i(w)$

# Dealing with Logical Omniscience

- *Syntactic approaches*: $\mathcal{M}, w \models K_i\varphi$ iff $\varphi \in \mathcal{C}_i(w)$
- *Awareness structures*: $\mathcal{M}, w \models K_i\varphi$ iff for all $v \in W$, if $wR_iv$ then $\mathcal{M}, v \models \varphi$ and $\varphi \in \mathcal{A}_i(w)$

# Dealing with Logical Omniscience

- *Syntactic approaches*: $\mathcal{M}, w \models K_i\varphi$ iff $\varphi \in \mathcal{C}_i(w)$
- *Awareness structures*: $\mathcal{M}, w \models K_i\varphi$ iff for all $v \in W$, if $wR_iv$ then $\mathcal{M}, v \models \varphi$ and $\varphi \in \mathcal{A}_i(w)$
- *Algorithmic knowledge*: $\mathcal{M}, w \models K_i\varphi$ iff $\mathsf{A}_i(w, \varphi) = \text{Yes}$

# Dealing with Logical Omniscience

- *Syntactic approaches*: $\mathcal{M}, w \models K_i\varphi$ iff $\varphi \in \mathcal{C}_i(w)$
- *Awareness structures*: $\mathcal{M}, w \models K_i\varphi$ iff for all $v \in W$, if $wR_iv$ then $\mathcal{M}, v \models \varphi$ and $\varphi \in \mathcal{A}_i(w)$
- *Algorithmic knowledge*: $\mathcal{M}, w \models K_i\varphi$ iff $\mathsf{A}_i(w, \varphi) = \mathrm{Yes}$
- *Impossible worlds*: $\mathcal{M}, w \models K_i\varphi$ iff if $w \in N$, then for all $v \in W$, if $wR_iv$ and $v \in N$ then $\mathcal{M}, v \models \varphi$; and if $w \notin N$, then $\varphi \in \mathcal{C}_i(w)$

# Justification Logic (1)

$t : \varphi$: "$t$ is a *justification*/*proof* for $\varphi$"

S. Artemov and M. Fitting (2019). *Justification Logic: Reasoning with Reasons*. Cambridge University Press.

S. Artemov and M. Fitting (2020). *Justification logic*. The Stanford Encyclopedia of Philosophy.

S. Artemov. *Explicit provability and constructive semantics (2001)*. The Bulletin of Symbolic Logic 7, pp. 1 - 36.

M. Fitting (2005). *The logic of proofs, semantically*. Annals of Pure and Applied Logic 132, pp. 1 - 25.

# Justification Logic (2)

$$t \;:= c \mid x \mid t + s \mid\; !t \mid t \cdot s$$

$$\varphi := p \mid \varphi \wedge \psi \mid \neg\varphi \mid t : \varphi$$

# Justification Logic (2)

$$t \ := c \mid x \mid t + s \mid !t \mid t \cdot s$$

$$\varphi := p \mid \varphi \wedge \psi \mid \neg \varphi \mid t : \varphi$$

Justification Logic:

- $t : \varphi \rightarrow \varphi$
- $t : (\varphi \rightarrow \psi) \rightarrow (s : \varphi \rightarrow t \cdot s : \psi)$
- $t : \varphi \rightarrow (t + s) : \varphi$
- $t : \varphi \rightarrow (s + t) : \varphi$
- $t : \varphi \rightarrow !t : t : \varphi$

# Justification Logic (2)

$$t \ := c \mid x \mid t + s \mid !t \mid t \cdot s$$

$$\varphi := p \mid \varphi \wedge \psi \mid \neg\varphi \mid t : \varphi$$

Justification Logic:

- $t : \varphi \to \varphi$
- $t : (\varphi \to \psi) \to (s : \varphi \to t \cdot s : \psi)$
- $t : \varphi \to (t + s) : \varphi$
- $t : \varphi \to (s + t) : \varphi$
- $t : \varphi \to !t : t : \varphi$

**Internalization**: if $\vdash_{JL} \varphi$ then there is a proof polynomial $t$ such that $\vdash_{JL} t : \varphi$
**Realization Theorem**: if $\vdash_{\mathbf{S4}} \varphi$ then there is a proof polynomial $t$ such that $\vdash_{JL} t : \varphi$

# Justification Logic (3)

**Fitting Semantics**: $\mathcal{M} = \langle W, R, \mathcal{E}, V \rangle$

- $W \neq \emptyset$
- $R \subseteq W \times W$
- $\mathcal{E} : W \times \text{ProofTerms} \rightarrow \wp(\mathcal{L}_{JL})$
- $V : \text{At} \rightarrow \wp(W)$

$\mathcal{M}, w \models t : \varphi$ iff for all $v$, if $wRv$ then $\mathcal{M}, v \models \varphi$ and $\varphi \in \mathcal{E}(w, t)$

# Justification Logic (3)

**Monotonicity** For all $w, v \in W$, if $wRv$ then for all proof polynomials $t$, $\mathcal{E}(w, t) \subseteq \mathcal{E}(v, t)$.

**Application** For all proof polynomials $s$, $t$ and for each $w \in W$, if $\varphi \rightarrow \psi \in \mathcal{E}(w, t)$ and $\varphi \in \mathcal{E}(w, s)$, then $\psi \in \mathcal{E}(w, t \cdot s)$

**Proof Checker** For all proof polynomials $t$ and for each $w \in W$, if $\varphi \in \mathcal{E}(w, t)$, then $t : \varphi \in \mathcal{E}(w, !t)$.

**Sum** For all proof polynomials $s, t$ and for each $w \in W$, $\mathcal{E}(w, s) \cup \mathcal{E}(w, t) \subseteq \mathcal{E}(w, s + t)$.

# Doxastic Logic: Models

Model: $\langle W, R, V \rangle$

States/possible worlds: $W \neq \emptyset$

Quasi-partitions: $R \subseteq W \times W$ is serial, transitive and Euclidean

# Doxastic Logic: Models

Model: $\langle W, R, V \rangle$

States/possible worlds: $W \neq \emptyset$

Quasi-partitions: $R \subseteq W \times W$ is serial, transitive and Euclidean

- *serial*: for all $w \in W$, there is a $v \in W$ such that $w\,R\,v$
- *transitive*: for all $w, v, u \in W$, if $w\,R\,v$ and $v\,R\,u$, then $w\,R\,u$
- *Euclidean*: for all $w, v, u \in W$, if $w\,R\,v$ and $w\,R\,u$, then $v\,R\,u$

# Doxastic Logic: Models

Model: $\langle W, R, V \rangle$

States/possible worlds: $W \neq \emptyset$

Quasi-partitions: $R \subseteq W \times W$ is serial, transitive and Euclidean

- *serial*: for all $w \in W$, there is a $v \in W$ such that $w \, R \, v$
- *transitive*: for all $w, v, u \in W$, if $w \, R \, v$ and $v \, R \, u$, then $w \, R \, u$
- *Euclidean*: for all $w, v, u \in W$, if $w \, R \, v$ and $w \, R \, u$, then $v \, R \, u$

Valuation function: $V : \mathrm{At} \to \wp(W)$, where At is a set of atomic propositions.

# Doxastic Logic: Language and Semantics

$$p \mid \varphi \wedge \varphi \mid \neg\varphi \mid B\varphi$$

# Doxastic Logic: Language and Semantics

$$p \mid \varphi \wedge \varphi \mid \neg\varphi \mid B\varphi$$

Boolean connectives:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$
- $\mathcal{M}, w \models \neg\varphi$ iff it is not the case that $\mathcal{M}, w \models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$

# Doxastic Logic: Language and Semantics

$$p \mid \varphi \wedge \varphi \mid \neg\varphi \mid B\varphi$$

Boolean connectives:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$
- $\mathcal{M}, w \models \neg\varphi$ iff it is not the case that $\mathcal{M}, w \models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$

Belief operators: $\mathcal{M}, w \models B\varphi$ iff for all $v$, if $w\, R\, v$, then $\mathcal{M}, v \models \varphi$.

# Doxastic Logic: Language and Semantics

$$p \mid \varphi \wedge \varphi \mid \neg\varphi \mid B\varphi$$

Boolean connectives:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$
- $\mathcal{M}, w \models \neg\varphi$ iff it is not the case that $\mathcal{M}, w \models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$

Belief operators: $\mathcal{M}, w \models B\varphi$ iff for all $v$, if $w\, R\, v$, then $\mathcal{M}, v \models \varphi$.

$$\mathcal{M}, w \models B\varphi \text{ iff } R(w) \subseteq [\![\varphi]\!]_{\mathcal{M}}$$

# Doxastic Logic: Language and Semantics

$$p \mid \varphi \wedge \varphi \mid \neg\varphi \mid B\varphi$$

Boolean connectives:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$
- $\mathcal{M}, w \models \neg\varphi$ iff it is not the case that $\mathcal{M}, w \models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ iff $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$

Belief operators: $\mathcal{M}, w \models B\varphi$ iff for all $v$, if $w\,R\,v$, then $\mathcal{M}, v \models \varphi$.

$$\mathcal{M}, w \models B\varphi \text{ iff } \underbrace{R(w)}_{\{v \mid w\,R\,v\}} \subseteq \underbrace{[\![\varphi]\!]_{\mathcal{M}}}_{\{v \mid \mathcal{M}, w \models \varphi\}}$$

# Doxastic Logic: **KD45**

| | |
|---|---|
| $K$ | $B(\varphi \to \psi) \to (B\varphi \to B\psi)$ |
| $D$ | $B\varphi \to \neg B \neg \varphi$ |
| 4 | $B\varphi \to BB\varphi$ |
| 5 | $\neg B\varphi \to B\neg B\varphi$ |

# Doxastic Logic: **KD45**

$$\begin{array}{ll} K & B(\varphi \to \psi) \to (B\varphi \to B\psi) \\ D & B\varphi \to \neg B \neg \varphi \\ 4 & B\varphi \to BB\varphi \\ 5 & \neg B\varphi \to B\neg B\varphi \end{array}$$

The logic **KD45** adds the above axiom schemes to an axiomatization of classical propositional logic with the rules Modus Ponens, Substitution of Equivalents, and Necessitation (from $\varphi$ infer $B\varphi$).

**KD45** is sound and strongly complete with respect to all quasi-partition frames.

**Exercise**: Show that the following axiom schemes and rules are valid on quasi-partition models and are theorems of **KD45**:

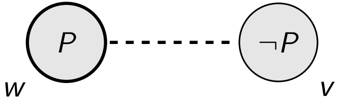
- agglomeration: $(B\varphi \wedge B\psi) \rightarrow B(\varphi \wedge \psi)$
- consistency: $\neg B\bot$
- monotonicity: From $\varphi \rightarrow \psi$ infer $B\varphi \rightarrow B\psi$

**Exercise**: Show that the following axiom schemes and rules are valid on quasi-partition models and are theorems of **KD45**:

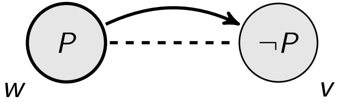
- agglomeration: $(B\varphi \wedge B\psi) \to B(\varphi \wedge \psi)$
- consistency: $\neg B\bot$
- monotonicity: From $\varphi \to \psi$ infer $B\varphi \to B\psi$
- secondary-reflexivity: for all $w, v \in W$, if $w\,R\,v$ then $v\,R\,v$
  $B(B\varphi \to \varphi)$

**Exercise**: Show that the following axiom schemes and rules are valid on quasi-partition models and are theorems of **KD45**:

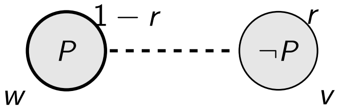
- agglomeration: $(B\varphi \wedge B\psi) \to B(\varphi \wedge \psi)$
- consistency: $\neg B\bot$
- monotonicity: From $\varphi \to \psi$ infer $B\varphi \to B\psi$
- secondary-reflexivity: for all $w, v \in W$, if $w\, R\, v$ then $v\, R\, v$
  $B(B\varphi \to \varphi)$
- correctness of own beliefs:
  $B\neg B\varphi \to \neg B\varphi$
  for all $w$, there is a $v$ such that $w\, R\, v$ and for all $z$ if $v\, R\, z$ then $w\, R\, z$
  $BB\varphi \to B\varphi$
  density: for all $w$ and $v$ if $w\, R\, v$ then there is a $z$ such that $w\, R\, z$ and $z\, R\, v$




Ann does not know that $P$




Ann does not know that $P$, but she believes that $\neg P$




Ann does not know that $P$, but she believes that $\neg P$ is true to degree $r$.

# Combining Logics of Knowledge and Belief

$\mathcal{M} = \langle W, \sim, R, V \rangle$ where

- $W \neq \emptyset$ is a set of states;
- each $\sim$ is an equivalence relation on $W$;
- each $R$ is a serial, transitive, Euclidean relation on $W$; and
- $V$ is a valuation function.

# Combining Logics of Knowledge and Belief

$\mathcal{M} = \langle W, \sim, R, V \rangle$ where

- $W \neq \emptyset$ is a set of states;
- each $\sim$ is an equivalence relation on $W$;
- each $R$ is a serial, transitive, Euclidean relation on $W$; and
- $V$ is a valuation function.

What is the relationship between knowledge ($K$) and believe ($B$)?

- $K$ is **S5**

# Combining Logics of Knowledge and Belief

$\mathcal{M} = \langle W, \sim, R, V \rangle$ where

- $W \neq \emptyset$ is a set of states;
- each $\sim$ is an equivalence relation on $W$;
- each $R$ is a serial, transitive, Euclidean relation on $W$; and
- $V$ is a valuation function.

What is the relationship between knowledge $(K)$ and believe $(B)$?

- $K$ is **S5**
- $B$ is **KD45**

# Combining Logics of Knowledge and Belief

$\mathcal{M} = \langle W, \sim, R, V \rangle$ where

- $W \neq \emptyset$ is a set of states;
- each $\sim$ is an equivalence relation on $W$;
- each $R$ is a serial, transitive, Euclidean relation on $W$; and
- $V$ is a valuation function.

What is the relationship between knowledge $(K)$ and believe $(B)$?

- $K$ is **S5**
- $B$ is **KD45**
- $K\varphi \to B\varphi$? "knowledge implies belief"

# Combining Logics of Knowledge and Belief

$\mathcal{M} = \langle W, \sim, R, V \rangle$ where

- $W \neq \emptyset$ is a set of states;
- each $\sim$ is an equivalence relation on $W$;
- each $R$ is a serial, transitive, Euclidean relation on $W$; and
- $V$ is a valuation function.

What is the relationship between knowledge $(K)$ and believe $(B)$?

- $K$ is **S5**
- $B$ is **KD45**
- $K\varphi \to B\varphi$? "knowledge implies belief"
- $B\varphi \to BK\varphi$? "positive certainty"

# Combining Logics of Knowledge and Belief

$\mathcal{M} = \langle W, \sim, R, V \rangle$ where

- $W \neq \emptyset$ is a set of states;
- each $\sim$ is an equivalence relation on $W$;
- each $R$ is a serial, transitive, Euclidean relation on $W$; and
- $V$ is a valuation function.

What is the relationship between knowledge ($K$) and believe ($B$)?

- $K$ is **S5**
- $B$ is **KD45**
- $K\varphi \rightarrow B\varphi$? "knowledge implies belief"
- $B\varphi \rightarrow BK\varphi$? "positive certainty"
- $B\varphi \rightarrow KB\varphi$? "strong introspection"

# An Issue - Negative Introspection and Positive Certainty

- Suppose that $p$ is something you are certain of (you *believe* it with probability one), but is false: $\neg p \land Bp$

# An Issue - Negative Introspection and Positive Certainty

- Suppose that $p$ is something you are certain of (you *believe* it with probability one), but is false: $\neg p \wedge Bp$
- $Bp \rightarrow BKp$

# An Issue - Negative Introspection and Positive Certainty

- Suppose that $p$ is something you are certain of (you *believe* it with probability one), but is false: $\neg p \wedge Bp$
- $Bp \rightarrow BKp$
- $\neg p \rightarrow \neg Kp$

# An Issue - Negative Introspection and Positive Certainty

- Suppose that $p$ is something you are certain of (you *believe* it with probability one), but is false: $\neg p \wedge Bp$
- $Bp \rightarrow BKp$
- $\neg p \rightarrow \neg Kp \rightarrow K\neg Kp$

# An Issue - Negative Introspection and Positive Certainty

- Suppose that $p$ is something you are certain of (you *believe* it with probability one), but is false: $\neg p \wedge Bp$

- $Bp \rightarrow BKp$

- $\neg p \rightarrow \neg Kp \rightarrow K\neg Kp \rightarrow B\neg Kp$

# An Issue - Negative Introspection and Positive Certainty

- Suppose that $p$ is something you are certain of (you *believe* it with probability one), but is false: $\neg p \wedge Bp$
- $Bp \rightarrow BKp$
- $\neg p \rightarrow \neg Kp \rightarrow K\neg Kp \rightarrow B\neg Kp$
- So, $BKp \wedge B\neg Kp$ also holds, but this contradictions $B\varphi \rightarrow \neg B\neg\varphi$.