Recent Work in Epistemic Logic

Logic Group Colloquium

Wes Holliday
Philosophy
UC Berkeley

March 1, 2013

- Preliminaries
- Basic Epistemic Logic
  - Modeling Knowledge
  - Common Knowledge
  - Information Update
- Mathematical Themes
  - Preservation Theorems
  - Uniform Substitution
  - Relations to FOL
- Conclusion

What is Epistemic Logic?

Epistemic logic provides a formal framework for modeling the knowledge of agents. Used by philosophers, theoretical computer scientists, AI researchers, game theorists, and others, epistemic logic has become one of the main application areas for modal logic.

A Very Brief History

12-15th c. Informal EL investigated by logicians of the Middle Ages.

A Very Brief History

12-15th c. Informal EL investigated by logicians of the Middle Ages.

1960s EL takes off in philosophy after Jaakko Hintikka's *Knowledge and Belief* (1962) (cf. von Wright's *An Essay on Modal Logic*, 1951).

A Very Brief History

12-15th c. Informal EL investigated by logicians of the Middle Ages.

1960s EL takes off in philosophy after Jaakko Hintikka's *Knowledge and Belief* (1962) (cf. von Wright's *An Essay on Modal Logic*, 1951).

1970s Extensive philosophical discussion sparked by Hintikka is reviewed in Wolfgang Lenzen's *Recent Work in Epistemic Logic* (1978).

A Very Brief History

12-15th c. Informal EL investigated by logicians of the Middle Ages.

1960s EL takes off in philosophy after Jaakko Hintikka's *Knowledge and Belief* (1962) (cf. von Wright's *An Essay on Modal Logic*, 1951).

1970s Extensive philosophical discussion sparked by Hintikka is reviewed in Wolfgang Lenzen's *Recent Work in Epistemic Logic* (1978).

1980-90s EL is "rediscovered" in theoretical computer science, leading to the classic textbook *Reasoning about Knowledge* (1995) by Ronald Fagin, Joseph Halpern, Yoram Moses, and Moshe Vardi.

A Very Brief History

12-15th c. Informal EL investigated by logicians of the Middle Ages.

1960s EL takes off in philosophy after Jaakko Hintikka's *Knowledge and Belief* (1962) (cf. von Wright's *An Essay on Modal Logic*, 1951).

1970s Extensive philosophical discussion sparked by Hintikka is reviewed in Wolfgang Lenzen's *Recent Work in Epistemic Logic* (1978).

1980-90s EL is "rediscovered" in theoretical computer science, leading to the classic textbook *Reasoning about Knowledge* (1995) by Ronald Fagin, Joseph Halpern, Yoram Moses, and Moshe Vardi.

1990s EL is discovered again in game theory by Robert Aumann (recipient of the 2005 Nobel Memorial Prize in Economics) leading to his "Interactive Epistemology I: Knowledge" (1999).

A Very Brief History

- **12-15th c.** Informal EL investigated by logicians of the Middle Ages.
- **1960s** EL takes off in philosophy after Jaakko Hintikka's *Knowledge and Belief* (1962) (cf. von Wright's *An Essay on Modal Logic*, 1951).
- **1970s** Extensive philosophical discussion sparked by Hintikka is reviewed in Wolfgang Lenzen's *Recent Work in Epistemic Logic* (1978).
- **1980-90s** EL is "rediscovered" in theoretical computer science, leading to the classic textbook *Reasoning about Knowledge* (1995) by Ronald Fagin, Joseph Halpern, Yoram Moses, and Moshe Vardi.
- **1990s** EL is discovered again in game theory by Robert Aumann (recipient of the 2005 Nobel Memorial Prize in Economics) leading to his "Interactive Epistemology I: Knowledge" (1999).
- **90s-now** *Dynamic* epistemic logic is developed in CS (Jan Plaza) and logic, especially by the "Amsterdam School" around Johan van Benthem.

Now I am interested in bringing epistemic logic back into contact with philosophy, using it as a tool to formalize existing theories of knowledge and develop new and improved ones. See, for example:

Wesley Holliday. Forthcoming.
"Epistemic Logic and Epistemology." *Handbook of Formal Philosophy*.

Wesley Holliday. Forthcoming.
"Epistemic Closure and Epistemic Logic I." *Journal of Philosophical Logic*.

Wesley Holliday. 2012.
"Epistemic Logic, Relevant Alternatives, and the Dynamics of Context."
New Directions in Logic, Language, and Computation, LNCS.

Wesley Holliday and John Perry. Forthcoming.
"Roles, Rigidity, and Quantificaiton in Epistemic Logic." *Trends in Logic*.

Logic and Philosophy

The project of proving results about formalized philosophical theories has led to the development of new logical techniques.

Logic and Philosophy

The project of proving results about formalized philosophical theories has led to the development of new logical techniques.

Wesley Holliday. Forthcoming.

"Epistemic Closure and Epistemic Logic I." *Journal of Philosophical Logic*.

This paper demonstrates an alternative method of proving completeness theorems in modal logic, without the standard Henkin-style canonical model construction, that yields results on satisfiability in finite models and complexity almost for free.

Logic and Philosophy

The project of proving results about formalized philosophical theories has led to the development of new logical techniques.

Wesley Holliday. Forthcoming.

"Epistemic Closure and Epistemic Logic I." *Journal of Philosophical Logic*.

This paper demonstrates an alternative method of proving completeness theorems in modal logic, without the standard Henkin-style canonical model construction, that yields results on satisfiability in finite models and complexity almost for free.

I am working to show how this method gives simple completeness proofs for standard systems of modal logic—good for students.

Today's Topic

But my talk today won't presuppose familiarity with any philosophical theories or standard techniques in modal logic.

Today's Topic

But my talk today won't presuppose familiarity with any philosophical theories or standard techniques in modal logic.

Today I'll deal with propositional epistemic logic, but first-order epistemic logic is also philosophically and technically rich.

Definition (Propositional Epistemic Language)

Given sets $\mathrm{At} = \{p, q, r, \dots\}$ and $\mathrm{Agt} = \{a, b, c, \dots\}$ with $|\mathrm{Agt}| = \kappa$, the epistemic language $\mathcal{L}_{\mathsf{EL}}^{\kappa}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

Definition (Propositional Epistemic Language)

Given sets $\mathrm{At} = \{p, q, r, \dots\}$ and $\mathrm{Agt} = \{a, b, c, \dots\}$ with $|\mathrm{Agt}| = \kappa$, the epistemic language $\mathcal{L}^{\kappa}_{\mathrm{EL}}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

We read $K_a\varphi$ as "**agent** a **knows that** φ."

Definition (Propositional Epistemic Language)

Given sets $\mathsf{At} = \{p, q, r, \dots\}$ and $\mathsf{Agt} = \{a, b, c, \dots\}$ with $|\mathsf{Agt}| = \kappa$, the epistemic language $\mathcal{L}_{\mathsf{EL}}^{\kappa}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

We read $K_a\varphi$ as "**agent** a **knows that** φ."

Let $\hat{K}_a\varphi := \neg K_a\neg\varphi$ for "φ **is consistent with** a**'s knowledge**."

Definition (Propositional Epistemic Language)

Given sets $\mathsf{At} = \{p, q, r, \dots\}$ and $\mathsf{Agt} = \{a, b, c, \dots\}$ with $|\mathsf{Agt}| = \kappa$, the epistemic language $\mathcal{L}_{\mathsf{EL}}^{\kappa}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

We read $K_a\varphi$ as "**agent** a **knows that** φ."

Let $\hat{K}_a\varphi := \neg K_a\neg\varphi$ for "φ **is consistent with** a**'s knowledge**."

Given a finite set $A \subseteq \mathsf{Agt}$, let

$$E_A\varphi := \bigwedge_{a\in A} K_a\varphi,$$

for "**everyone in** A **knows that** φ."

Definition (Propositional Epistemic Language)

Given sets $\mathrm{At} = \{p, q, r, \dots\}$ and $\mathrm{Agt} = \{a, b, c, \dots\}$ with $|\mathrm{Agt}| = \kappa$, the epistemic language $\mathcal{L}^{\kappa}_{\mathsf{EL}}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

We read $K_a\varphi$ as "**agent** a **knows that** φ."

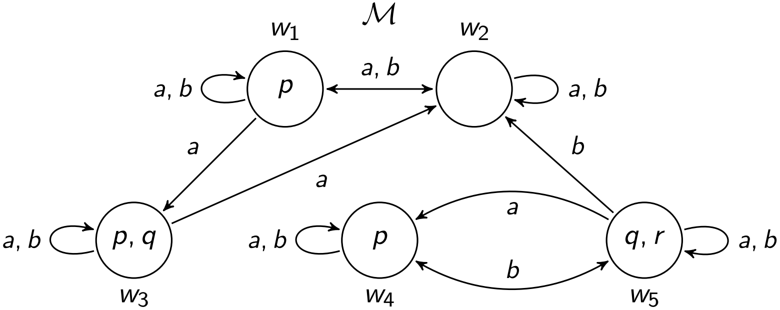
Let $\hat{K}_a\varphi := \neg K_a\neg\varphi$ for "φ **is consistent with** a**'s knowledge**."

Given a finite set $A \subseteq \mathrm{Agt}$, let

$$E_A\varphi := \bigwedge_{a\in A} K_a\varphi,$$

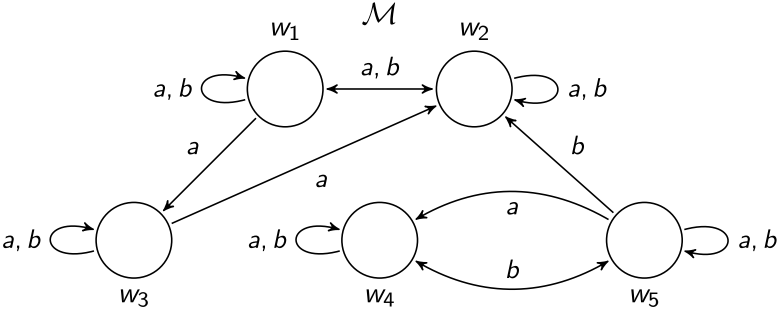
for "**everyone in** A **knows that** φ."

Where $\Box$ is K_a or E_A, $\Box^1\varphi := \Box\varphi$ and $\Box^{n+1}\varphi := \Box\Box^n\varphi$ $(n \geq 1)$.



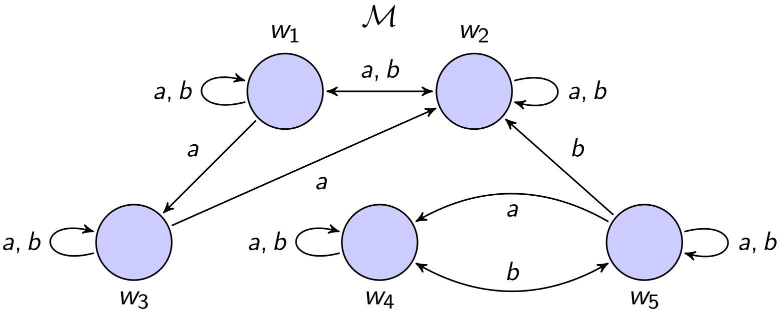
Definition (Epistemic Frames and Models)

A κ-agent epistemic **model** is a triple $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$:



Definition (Epistemic Frames and Models)

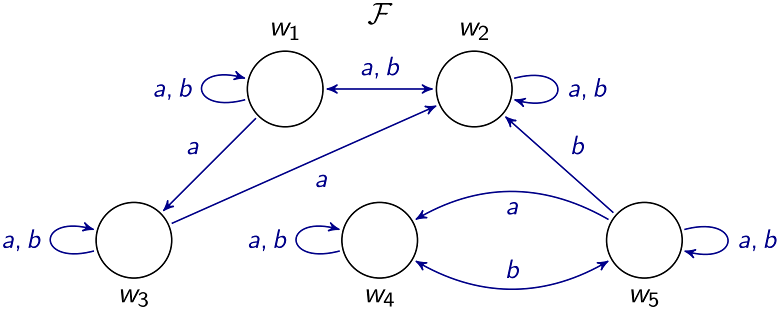
A κ-agent epistemic **frame** is a pair $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$ where:



Definition (Epistemic Frames and Models)

A κ-agent epistemic **frame** is a pair $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$ where:

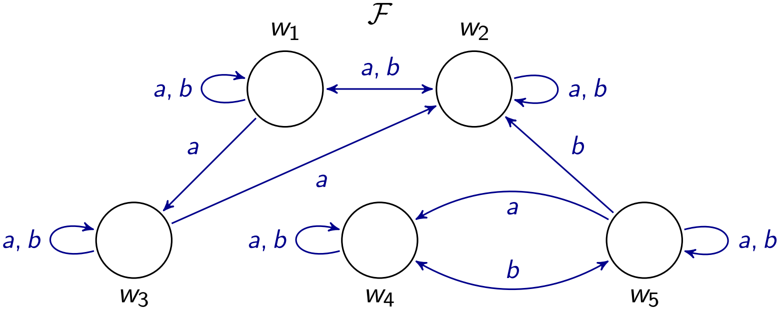
- W is a non-empty set (of "possibilities," "scenarios," "states," or "worlds");



Definition (Epistemic Frames and Models)

A κ-agent epistemic **frame** is a pair $\mathcal{F} = \langle W, \{R_a\}_{a \in \text{Agt}} \rangle$ where:

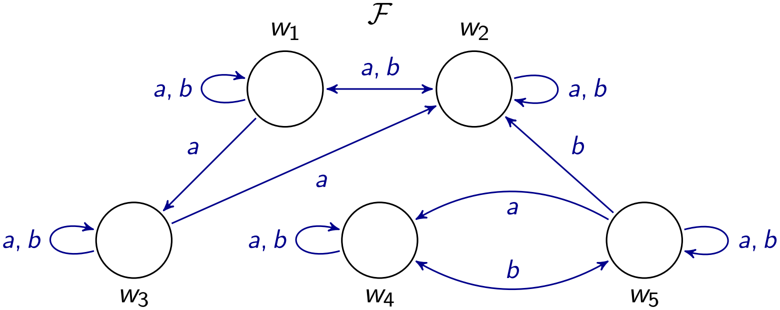
- for each $a \in \text{Agt}$ ($|\text{Agt}| = \kappa$), R_a is a binary relation on W; R_a can be thought of as representing agent a's **uncertainty**.



Definition (Epistemic Frames and Models)

A κ-agent epistemic **frame** is a pair $\mathcal{F} = \langle W, \{R_a\}_{a \in \text{Agt}} \rangle$ where:

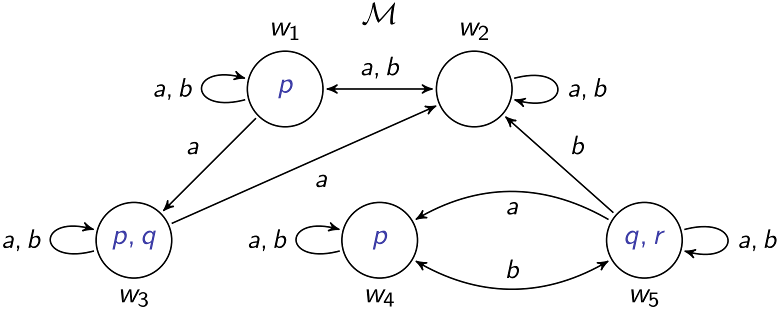
- for each $a \in \text{Agt}$ ($|\text{Agt}| = \kappa$), R_a is a binary relation on W; let's call R_a agent a's "**epistemic accessibility**" relation; wR_av means that everything a knows in w is true in v.



Definition (Epistemic Frames and Models)

A κ-agent epistemic **frame** is a pair $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$ where:

- for each $a \in \mathsf{Agt}$ ($|\mathsf{Agt}| = \kappa$), R_a is a binary relation on W; let's call R_a agent a's "**epistemic accessibility**" relation; wR_av means that v is compatible with a's knowledge in w.



Definition (Epistemic Model)

A κ-agent epistemic **model** is a triple $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$:

- $V : \mathsf{At} \to \wp(w)$ assigns to each atomic sentence the set of scenarios in which it holds. We say $\mathcal{M}$ is *based on* $\mathcal{F}$.

Three Ideas for the Semantics

The semantics will be based on three ideas from philosophy:

Three Ideas for the Semantics

The semantics will be based on three ideas from philosophy:

1. The knowledge operator K_a can be interpreted as a restricted quantifier ranging over those possibilities compatible with a's knowledge (J. Hintikka, *Knowledge and Belief*, 1962).

Three Ideas for the Semantics

The semantics will be based on three ideas from philosophy:

1. The knowledge operator K_a can be interpreted as a restricted quantifier ranging over those possibilities compatible with a's knowledge (J. Hintikka, *Knowledge and Belief*, 1962).

2. It is *common knowledge* that φ when everyone knows that φ, everyone knows that everyone knows that φ, everyone knows that everyone knows that everyone knows that φ, etc. (inspired by D. Lewis, *Convention*, 1969).

Three Ideas for the Semantics

The semantics will be based on three ideas from philosophy:

1. The knowledge operator K_a can be interpreted as a restricted quantifier ranging over those possibilities compatible with a's knowledge (J. Hintikka, *Knowledge and Belief*, 1962).

2. It is *common knowledge* that φ when everyone knows that φ, everyone knows that everyone knows that φ, everyone knows that everyone knows that everyone knows that φ, etc. (inspired by D. Lewis, *Convention*, 1969).

3. *Knowledge acquisition* (information update) can be understood in terms of the elimination of possibilities from an agent's initial epistemic state (R. Stalnaker, *Inquiry*, 1987).

Idea 1: Knowledge Operators as Restricted Quantifiers

Example (Berkeley and Copenhagen)

Let K_b stand for **agent** b **knows that** and K_c stand for **agent** c **knows that**. Suppose agent b, who lives in Berkeley, knows that agent c lives in Copenhagen. Let r stand for 'it's raining in Copenhagen'. Although b doesn't know whether it's raining in Copenhagen, b knows that c knows whether it's raining there:

Idea 1: Knowledge Operators as Restricted Quantifiers

Example (Berkeley and Copenhagen)

Let K_b stand for **agent** b **knows that** and K_c stand for **agent** c **knows that**. Suppose agent b, who lives in Berkeley, knows that agent c lives in Copenhagen. Let r stand for 'it's raining in Copenhagen'. Although b doesn't know whether it's raining in Copenhagen, b knows that c knows whether it's raining there:

$$\neg(K_b r \vee K_b \neg r) \wedge K_b(K_c r \vee K_c \neg r).$$

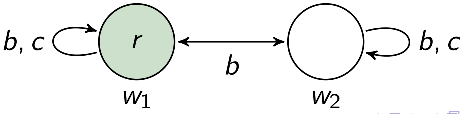
Idea 1: Knowledge Operators as Restricted Quantifiers

Example (Berkeley and Copenhagen)

Let K_b stand for **agent** b **knows that** and K_c stand for **agent** c **knows that**. Suppose agent b, who lives in Berkeley, knows that agent c lives in Copenhagen. Let r stand for 'it's raining in Copenhagen'. Although b doesn't know whether it's raining in Copenhagen, b knows that c knows whether it's raining there:

$$\neg(K_b r \vee K_b \neg r) \wedge K_b(K_c r \vee K_c \neg r).$$

The following picture depicts a situation in which this is true, where an arrow represents *compatibility with one's knowledge*:

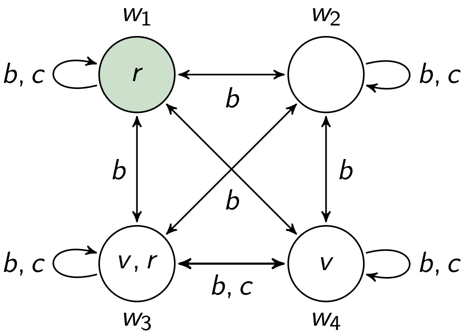


Now suppose that agent b doesn't know whether agent c has left Copenhagen for a vacation. (Let v stand for 'c has left Copenhagen on vacation'.) Agent b knows that if c is not on vacation, then c knows whether it's raining in Copenhagen; but if c is on vacation, then c won't bother to follow the weather.

$$K_b(\neg v \to (K_c r \vee K_c \neg r)) \wedge K_b(v \to \neg(K_c r \vee K_c \neg r)).$$

Now suppose that agent b doesn't know whether agent c has left Copenhagen for a vacation. (Let v stand for 'c has left Copenhagen on vacation'.) Agent b knows that if c is not on vacation, then c knows whether it's raining in Copenhagen; but if c is on vacation, then c won't bother to follow the weather.

$$K_b(\neg v \to (K_c r \vee K_c \neg r)) \wedge K_b(v \to \neg(K_c r \vee K_c \neg r)).$$



Idea 1: Knowledge Operators as Restricted Quantifiers

Definition (Truth)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}_{\mathsf{EL}}^{\kappa}$, we define $\mathcal{M}, w \vDash \varphi$ ("φ is true in $\mathcal{M}$ at w") as follows:

$$
\begin{array}{lll}
\mathcal{M}, w \vDash p & \text{iff} & w \in V(p); \\
\mathcal{M}, w \vDash \neg\varphi & \text{iff} & \mathcal{M}, w \nvDash \varphi; \\
\mathcal{M}, w \vDash (\varphi \wedge \psi) & \text{iff} & \mathcal{M}, w \vDash \varphi \text{ and } \mathcal{M}, w \vDash \psi;
\end{array}
$$

Idea 1: Knowledge Operators as Restricted Quantifiers

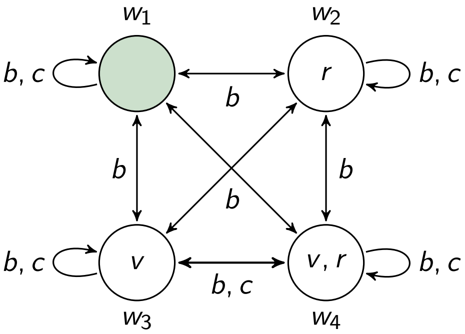
Definition (Truth)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V\rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{EL}}$, we define $\mathcal{M}, w \vDash \varphi$ ("φ is true in $\mathcal{M}$ at w") as follows:

$$
\begin{array}{lll}
\mathcal{M}, w \vDash p & \text{iff} & w \in V(p); \\
\mathcal{M}, w \vDash \neg\varphi & \text{iff} & \mathcal{M}, w \nvDash \varphi; \\
\mathcal{M}, w \vDash (\varphi \wedge \psi) & \text{iff} & \mathcal{M}, w \vDash \varphi \text{ and } \mathcal{M}, w \vDash \psi; \\
\mathcal{M}, w \vDash K_a\varphi & \text{iff} & \forall v \in W\text{: if } wR_av \text{ then } \mathcal{M}, v \vDash \varphi;
\end{array}
$$

Now suppose that agent b doesn't know whether agent c has left Copenhagen for a vacation. (Let v stand for 'c has left Copenhagen on vacation'.) Agent b knows that if c is not on vacation, then c knows whether it's raining in Copenhagen; but if c is on vacation, then c won't bother to follow the weather.

$$K_b(\neg v \to (K_c r \vee K_c \neg r)) \wedge K_b(v \to \neg(K_c r \vee K_c \neg r)).$$



Idea 1: Knowledge Operators as Restricted Quantifiers

Definition (Truth and Validity)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{EL}}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash K_a\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_av \text{ then } \mathcal{M}, v \vDash \varphi;$$

Idea 1: Knowledge Operators as Restricted Quantifiers

Definition (Truth and Validity)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}_{\mathsf{EL}}^{\kappa}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash K_a\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_av \text{ then } \mathcal{M}, v \vDash \varphi;$$

We say φ is globally true in $\mathcal{M}$ iff for all $w \in W$, $\mathcal{M}, w \vDash \varphi$;

Idea 1: Knowledge Operators as Restricted Quantifiers

Definition (Truth and Validity)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{EL}}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash K_a\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_av \text{ then } \mathcal{M}, v \vDash \varphi;$$

We say φ is globally true in $\mathcal{M}$ iff for all $w \in W$, $\mathcal{M}, w \vDash \varphi$;

φ is valid on frame $\mathcal{F}$ iff globally true in all models based on $\mathcal{F}$;

Idea 1: Knowledge Operators as Restricted Quantifiers

Definition (Truth and Validity)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{EL}}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash K_a\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_av \text{ then } \mathcal{M}, v \vDash \varphi;$$

We say φ is globally true in $\mathcal{M}$ iff for all $w \in W$, $\mathcal{M}, w \vDash \varphi$;

φ is valid on frame $\mathcal{F}$ iff globally true in all models based on $\mathcal{F}$;

φ is valid on a class C of models (resp. a class $\mathbb{F}$ of frames) iff globally true in all models in C (resp. based on frames in $\mathbb{F}$).

Idea 1: Knowledge Operators as Restricted Quantifiers

Definition (Truth and Validity)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}_{\mathsf{EL}}^{\kappa}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash K_a\varphi \quad \text{iff} \quad \forall v \in W\colon \text{if } wR_av \text{ then } \mathcal{M}, v \vDash \varphi;$$

We say φ is globally true in $\mathcal{M}$ iff for all $w \in W$, $\mathcal{M}, w \vDash \varphi$;

φ is valid on frame $\mathcal{F}$ iff globally true in all models based on $\mathcal{F}$;

φ is valid on a class C of models (resp. a class $\mathbb{F}$ of frames) iff globally true in all models in C (resp. based on frames in $\mathbb{F}$).

For language $\mathcal{L}$, the $\mathcal{L}$-theory of a class C of models, $\mathsf{Th}_{\mathcal{L}}(\mathsf{C})$ (resp. class $\mathbb{F}$ of frames) is the set of $\varphi \in \mathcal{L}$ valid on C (resp. $\mathbb{F}$).

Correspondence

Proposition

For any frame $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$*:*

T $K_a\varphi \to \varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is reflexive;*

Correspondence

Proposition

For any frame $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$*:*

- T $K_a\varphi \to \varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is reflexive;*
- 4 $K_a \to K_aK_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is transitive;*

Correspondence

Proposition

For any frame $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$*:*

- T $K_a\varphi \rightarrow \varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is reflexive;*
- 4 $K_a \rightarrow K_aK_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is transitive;*
- B $\neg\varphi \rightarrow K_a\neg K_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is symmetric;*

Correspondence

Proposition

For any frame $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$*:*

- T $K_a\varphi \to \varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is reflexive;*
- 4 $K_a \to K_aK_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is transitive;*
- B $\neg\varphi \to K_a\neg K_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is symmetric;*
- 5 $\neg K_a\varphi \to K_a\neg K_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is Euclidean.*

Correspondence

Proposition

For any frame $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathsf{Agt}} \rangle$:

- T $K_a\varphi \rightarrow \varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is reflexive;*
- 4 $K_a \rightarrow K_aK_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is transitive;*
- B $\neg\varphi \rightarrow K_a\neg K_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is symmetric;*
- 5 $\neg K_a\varphi \rightarrow K_a\neg K_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is Euclidean.*

Cf. Theorem 3.5 of

Bjarni Jonsson and Alfred Tarski. 1951. "Boolean Algebras with Operators."

Correspondence

Proposition

For any frame $\mathcal{F} = \langle W, \{R_a\}_{a \in \mathrm{Agt}} \rangle$:

- T $K_a\varphi \to \varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is reflexive;*
- 4 $K_a \to K_aK_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is transitive;*
- B $\neg\varphi \to K_a\neg K_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is symmetric;*
- 5 $\neg K_a\varphi \to K_a\neg K_a\varphi$ *is valid on* $\mathcal{F}$ *iff* R_a *is Euclidean.*

Cf. Theorem 3.5 of

Bjarni Jonsson and Alfred Tarski. 1951. "Boolean Algebras with Operators."

In CS and game theory, it is typically assumed that every R_a is an equivalence relation. We call such models **partition models**.

Complexity

Theorem (Complexity)

1. *model-checking is in linear time (in the size of $\mathcal{M}$ and φ);*
2. *satisfiability in* 1*-agent models is PSPACE-complete;*
3. *satisfiability in* 1*-agent partition models is NP-complete;*
4. *satisfiability in* 2*-agent partition models is PSPACE-complete.*

If we bound the nesting depth of K, satisfiability drops down to NP-complete in most cases; if we further restrict to finitely many atomic sentences, satisfiability drops down to linear time.

Completeness

Theorem

The $\mathcal{L}^{\kappa}_{\mathsf{EL}}$-theory of the class of all κ-agent epistemic models is axiomatized by the system $\mathbf{K}_{\kappa}$:

- *all substitutions of propositional tautologies as axioms;*
- *the K axiom, $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$, for $a \in \mathrm{Agt}$;*
- *modus ponens and necessitation, $\vdash \varphi \Rightarrow \vdash K_a\varphi$, for $a \in \mathrm{Agt}$.*

Completeness

Theorem

The $\mathcal{L}_{\mathsf{EL}}^{\kappa}$-theory of the class of all κ-agent epistemic models is axiomatized by the system $\mathbf{K}_{\kappa}$:

- *all substitutions of propositional tautologies as axioms;*
- *the K axiom, $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$, for $a \in \mathrm{Agt}$;*
- *modus ponens and necessitation, $\vdash \varphi \Rightarrow \vdash K_a\varphi$, for $a \in \mathrm{Agt}$.*

To axiomatize the theories of other model classes, add the axioms corresponding to the defining properties of the class:

Completeness

Theorem

The $\mathcal{L}_{\mathsf{EL}}^{\kappa}$-theory of the class of all κ-agent epistemic models is axiomatized by the system $\mathbf{K}_{\kappa}$:

- *all substitutions of propositional tautologies as axioms;*
- *the K axiom, $K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi)$, for $a \in \mathsf{Agt}$;*
- *modus ponens and necessitation, $\vdash \varphi \Rightarrow \vdash K_a\varphi$, for $a \in \mathsf{Agt}$.*

To axiomatize the theories of other model classes, add the axioms corresponding to the defining properties of the class:

- *for reflexive models, add the T axiom, $K_a\varphi \to \varphi$;*
- *for transitive models, add the 4 axiom, $K_a\varphi \to K_aK_a\varphi$;*
- *for Euclidean models, add the 5 axiom, $\neg K\varphi \to K\neg K\varphi_a$;*
- *etc.*

Idea 2: Common Knowledge as Transitive Closure

The "everyone knows" operator is not the only group knowledge operator of interest. To see this, consider the following example.

Idea 2: Common Knowledge as Transitive Closure

The "everyone knows" operator is not the only group knowledge operator of interest. To see this, consider the following example.

Example (Consecutive Numbers (J.E. Littlewood 1953))

Two agents, b and g, are facing each other. They each see a number on the other's forehead. A perfectly trustworthy source tells them: one of you has $n \in \mathbb{N}$ and the other has $n+1$.

Idea 2: Common Knowledge as Transitive Closure

The "everyone knows" operator is not the only group knowledge operator of interest. To see this, consider the following example.

Example (Consecutive Numbers (J.E. Littlewood 1953))

Two agents, b and g, are facing each other. They each see a number on the other's forehead. A perfectly trustworthy source tells them: one of you has $n \in \mathbb{N}$ and the other has $n+1$.

Let b_m stand for "the number on b's head is m," and let g_m stand for "the number on g's head is m." The truth is: $b_2 \wedge g_3$.

Idea 2: Common Knowledge as Transitive Closure

The "everyone knows" operator is not the only group knowledge operator of interest. To see this, consider the following example.

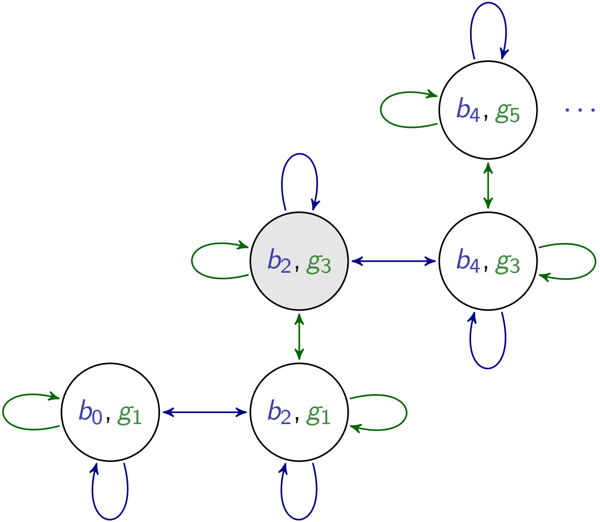
Example (Consecutive Numbers (J.E. Littlewood 1953))

Two agents, b and g, are facing each other. They each see a number on the other's forehead. A perfectly trustworthy source tells them: one of you has $n \in \mathbb{N}$ and the other has $n+1$.

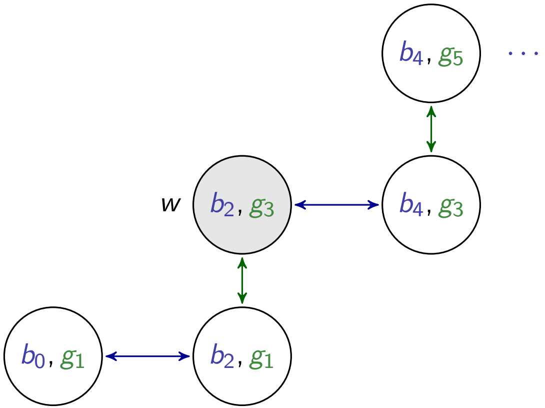
Let b_m stand for "the number on b's head is m," and let g_m stand for "the number on g's head is m." The truth is: $b_2 \wedge g_3$.

What should an epistemic model representing this look like?

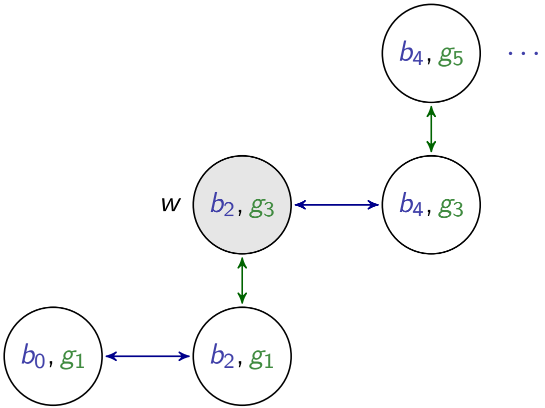
Here is an epistemic model representing the example:



Here is an epistemic model representing the example:

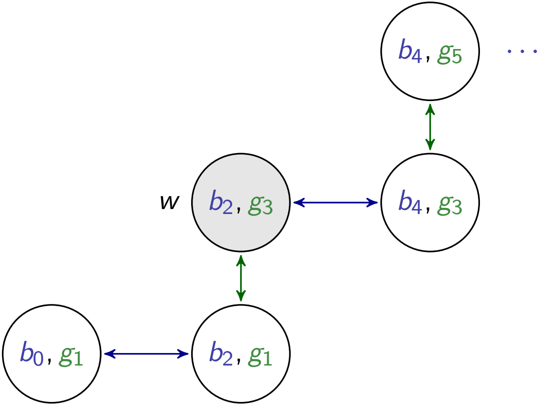


Here is an epistemic model representing the example:



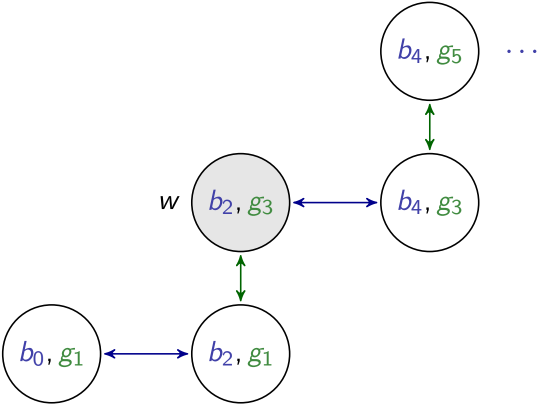
$\mathcal{M}, w \vDash K_b g_3$;

Here is an epistemic model representing the example:



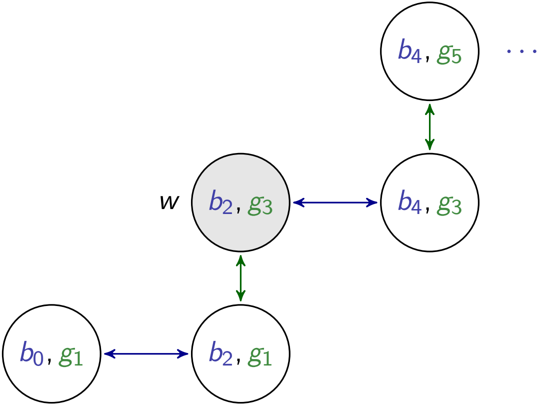
$\mathcal{M}, w \vDash K_b g_3$; $\mathcal{M}, w \vDash K_b(b_2 \vee b_4)$;

Here is an epistemic model representing the example:



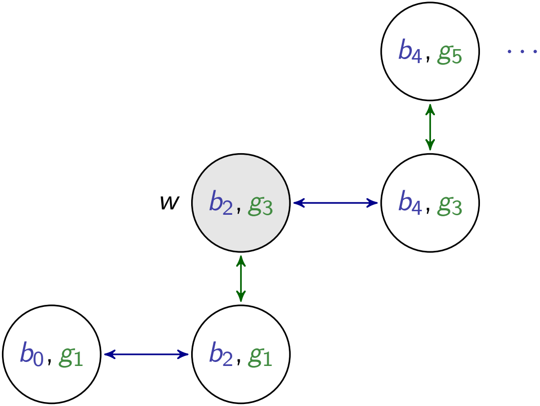
$\mathcal{M}, w \vDash K_b g_3$; $\mathcal{M}, w \vDash K_b(b_2 \vee b_4)$; $\mathcal{M}, w \vDash K_b K_g(g_1 \vee g_3 \vee g_5)$;

Here is an epistemic model representing the example:



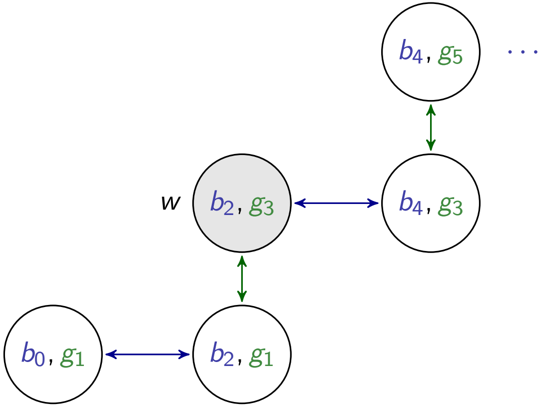
$\mathcal{M}, w \vDash K_b g_3$; $\mathcal{M}, w \vDash K_b(b_2 \vee b_4)$; $\mathcal{M}, w \vDash K_b K_g(g_1 \vee g_3 \vee g_5)$;
$\mathcal{M}, w \vDash K_b K_g K_b(g_1 \vee g_3 \vee g_5)$; and so on...

Here is an epistemic model representing the example:



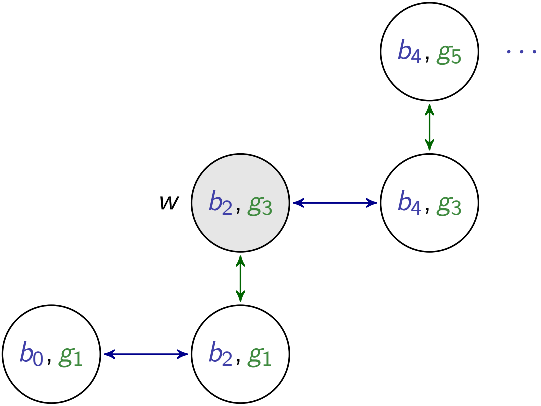
$\mathcal{M}, w \vDash \hat{K}_g g_1 \wedge \hat{K}_g g_3$;

Here is an epistemic model representing the example:



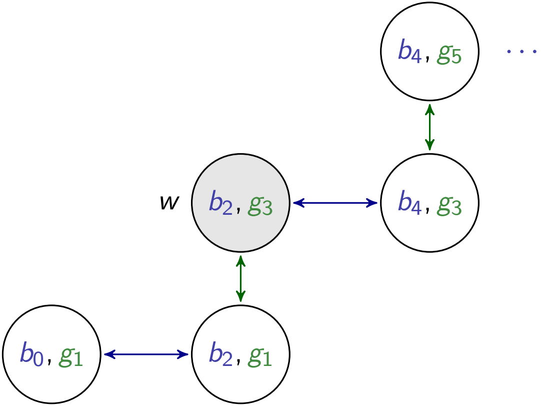
$\mathcal{M}, w \vDash \hat{K}_g g_1 \wedge \hat{K}_g g_3$; $\mathcal{M}, w \vDash \hat{K}_g \hat{K}_b K_g g_1$;

Here is an epistemic model representing the example:



$\mathcal{M}, w \vDash \hat{K}_g g_1 \wedge \hat{K}_g g_3$; $\mathcal{M}, w \vDash \hat{K}_g \hat{K}_b K_g g_1$; and interestingly,
$\mathcal{M}, w \vDash E_{\{b,g\}} \neg g_5 \wedge \neg E_{\{b,g\}} E_{\{b,g\}} \neg g_5$.

Here is an epistemic model representing the example:



$\mathcal{M}, w \vDash \hat{K}_g g_1 \wedge \hat{K}_g g_3$; $\mathcal{M}, w \vDash \hat{K}_g \hat{K}_b K_g g_1$; and interestingly, $\mathcal{M}, w \vDash E_{\{b,g\}} \neg g_5 \wedge \neg E_{\{b,g\}} E_{\{b,g\}} \neg g_5$. Not common knowledge.

Definition (Basic Epistemic Languages)

Given countable sets $\mathsf{At} = \{p, q, r, \dots\}$ and $\mathsf{Agt} = \{a, b, c, \dots\}$ with $|\mathsf{Agt}| = \kappa$, the epistemic language $\mathcal{L}^{\kappa}_{\mathsf{EL}}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

For the epistemic language with common knowledge, $\mathcal{L}^{\kappa}_{\mathsf{EL\text{-}C}}$,

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C_A\varphi,$$

where $A \subseteq \mathsf{Agt}$.

Definition (Basic Epistemic Languages)

Given countable sets $\mathsf{At} = \{p, q, r, \ldots\}$ and $\mathsf{Agt} = \{a, b, c, \ldots\}$ with $|\mathsf{Agt}| = \kappa$, the epistemic language $\mathcal{L}^{\kappa}_{\mathsf{EL}}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

For the epistemic language with common knowledge, $\mathcal{L}^{\kappa}_{\mathsf{EL\text{-}C}}$,

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C_A\varphi,$$

where $A \subseteq \mathsf{Agt}$. We read $C_A\varphi$ as "**it is common knowledge among the members of** A **that** φ."

Idea 2: Common Knowledge as Transitive Closure

Definition

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{EL\text{-}C}}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash K_a\varphi \quad \text{iff} \quad \forall v \in W\colon wR_av \text{ implies } \mathcal{M}, v \vDash \varphi;$$

$$\mathcal{M}, w \vDash C_A\varphi \quad \text{iff} \quad \forall v \in W\colon wR_A^{+}v \text{ implies } \mathcal{M}, v \vDash \varphi,$$

where R_A^{+} is the transitive closure of $\bigcup_{a \in A} R_a$.

Idea 2: Common Knowledge as Transitive Closure

Definition

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V \rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{EL\text{-}C}}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

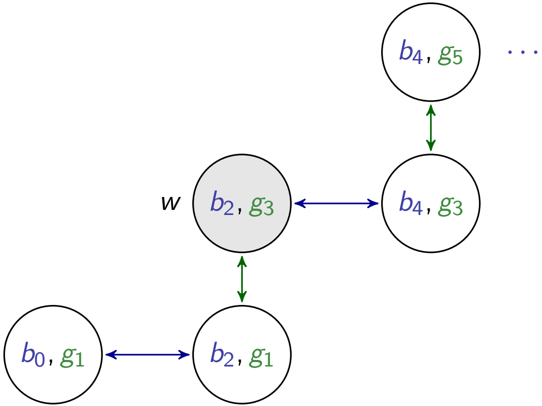
$$\mathcal{M}, w \vDash K_a\varphi \quad \text{iff} \quad \forall v \in W\colon wR_av \text{ implies } \mathcal{M}, v \vDash \varphi;$$

$$\mathcal{M}, w \vDash C_A\varphi \quad \text{iff} \quad \forall v \in W\colon wR^+_Av \text{ implies } \mathcal{M}, v \vDash \varphi,$$

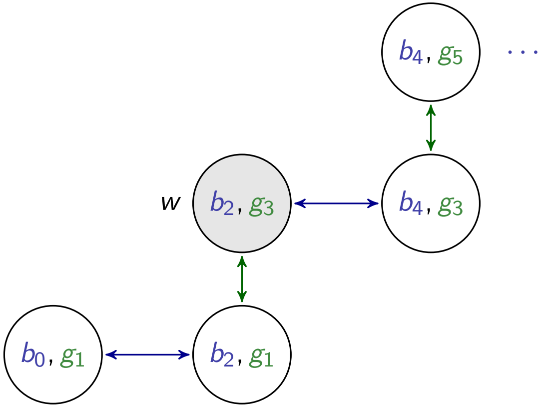
where R^+_A is the transitive closure of $\bigcup_{a \in A} R_a$.

In other words, $C_A\varphi$ is true at w iff any sequence of steps from w along any agents' relations ends in a state where φ is true.

Example: Consecutive Numbers

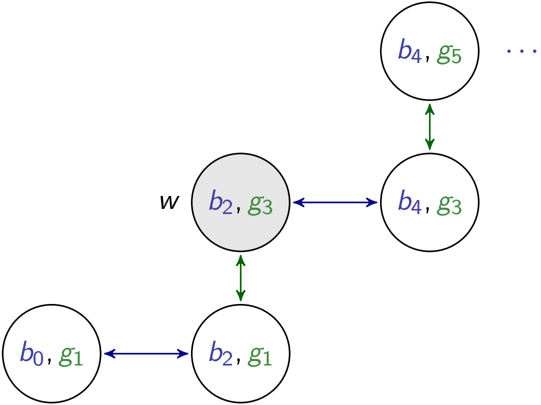


Example: Consecutive Numbers



Recall: $\mathcal{M}, w \vDash E_{\{b,g\}}\neg g_5 \wedge \neg E_{\{b,g\}}E_{\{b,g\}}\neg g_5$.

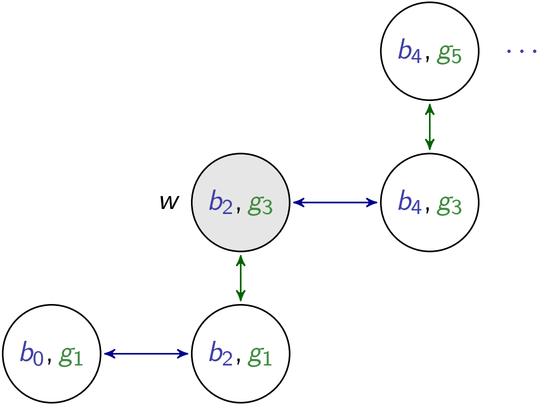
Example: Consecutive Numbers



Recall: $\mathcal{M}, w \vDash E_{\{b,g\}} \neg g_5 \wedge \neg E_{\{b,g\}} E_{\{b,g\}} \neg g_5$.

Pattern: $\mathcal{M}, w \vDash E^3_{\{b,g\}} \neg g_6 \wedge \neg E^4_{\{b,g\}} \neg g_6$.

Example: Consecutive Numbers

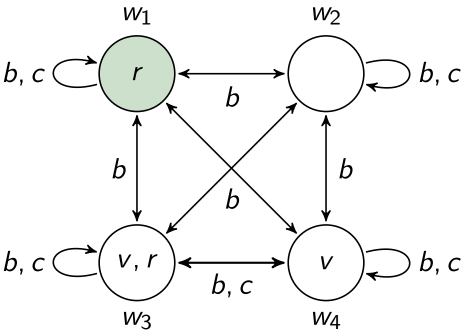


Surprising result: $\mathcal{M}, w \vDash K_b g_3 \wedge K_g b_2 \wedge \neg C_{\{b,g\}} \neg b_{1000}$.

It's not common knowledge that b doesn't have 1000 on his head!

By contrast, in our earlier example with Berkeley and Copenhagen, it is common knowledge that if c is on vacation, then he doesn't know whether or not it's raining in Copenhagen:

$$\mathcal{M}, w_1 \vDash C_{\{b,c\}}(v \to \neg(K_c r \vee K_c \neg r)).$$



Idea 2: Common Knowledge as Transitive Closure

Definition

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V\rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{EL\text{-}C}}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash C_A\varphi \quad \text{iff} \quad \forall v \in W\colon wR^+_A v \text{ implies } \mathcal{M}, v \vDash \varphi,$$

where R^+_A is the transitive closure of $\bigcup_{a \in G} R_a$.

Observation

Given the infinitary character of common knowledge, it's no surprise that with $\mathcal{L}_{\mathsf{EL\text{-}C}}$ we lose compactness:

$$\{E^n_A p \mid n \in \mathbb{N}\} \cup \{\neg C_A p\}$$

is finitely-satisfiable but not satisfiable.

Completeness and Complexity

Theorem

The $\mathcal{L}_{\mathsf{EL\text{-}C}}$-theory of the class of all (resp. reflexive, preorder, partition, etc.) models is axiomatized by $\mathbf{K}_\kappa$ *(resp.* $\mathbf{T}_\kappa$, $\mathbf{S4}_\kappa$, $\mathbf{S5}_\kappa$, *etc.) plus the following schemas for all* $A \subseteq \mathrm{Agt}$*:*

- *Fixed-Point Axiom:* $C_A\varphi \leftrightarrow (\varphi \wedge E_A C_A \varphi)$

Completeness and Complexity

Theorem

The $\mathcal{L}_{\text{EL-C}}$-theory of the class of all (resp. reflexive, preorder, partition, etc.) models is axiomatized by $\mathbf{K}_\kappa$ *(resp.* $\mathbf{T}_\kappa$, $\mathbf{S4}_\kappa$, $\mathbf{S5}_\kappa$, *etc.) plus the following schemas for all* $A \subseteq \text{Agt}$*:*

- *Fixed-Point Axiom:* $C_A\varphi \leftrightarrow (\varphi \wedge E_A C_A\varphi)$
- *Induction Axiom:* $(\varphi \wedge C_A(\varphi \rightarrow E_A\varphi)) \rightarrow C_A\varphi$

Completeness and Complexity

Theorem

The $\mathcal{L}_{\text{EL-C}}$-theory of the class of all (resp. reflexive, preorder, partition, etc.) models is axiomatized by $\mathbf{K}_\kappa$ *(resp.* $\mathbf{T}_\kappa$, $\mathbf{S4}_\kappa$, $\mathbf{S5}_\kappa$, *etc.) plus the following schemas for all* $A \subseteq \text{Agt}$*:*

- *Fixed-Point Axiom:* $C_A\varphi \leftrightarrow (\varphi \wedge E_A C_A \varphi)$
- *Induction Axiom:* $(\varphi \wedge C_A(\varphi \to E_A\varphi)) \to C_A\varphi$

Complexity of checking satisfiability goes to EXPTIME-complete with common knowledge.

Idea 3: Information Update as Model Restriction

Example (Muddle Children Puzzle)

Three children—r, g, and b—have returned after playing outside. Their mother says to them, "At least one of you has mud on your forehead. Step forward now if you know whether you are dirty."

Idea 3: Information Update as Model Restriction

Example (Muddle Children Puzzle)

Three children—r, g, and b—have returned after playing outside. Their mother says to them, "At least one of you has mud on your forehead. Step forward now if you know whether you are dirty."

In fact, g and b are dirty, but no one can see his own forehead.

Idea 3: Information Update as Model Restriction

Example (Muddle Children Puzzle)

Three children—r, g, and b—have returned after playing outside. Their mother says to them, “At least one of you has mud on your forehead. Step forward now if you know whether you are dirty.”

In fact, g and b are dirty, but no one can see his own forehead.

None of the children step forward. So their mother says again, “Step forward now if you know whether you are dirty.”

Idea 3: Information Update as Model Restriction

Example (Muddle Children Puzzle)

Three children—r, g, and b—have returned after playing outside. Their mother says to them, "At least one of you has mud on your forehead. Step forward now if you know whether you are dirty."

In fact, g and b are dirty, but no one can see his own forehead.

None of the children step forward. So their mother says again, "Step forward now if you know whether you are dirty."

This time g and b both step forward, but r does not.

Idea 3: Information Update as Model Restriction

Example (Muddle Children Puzzle)

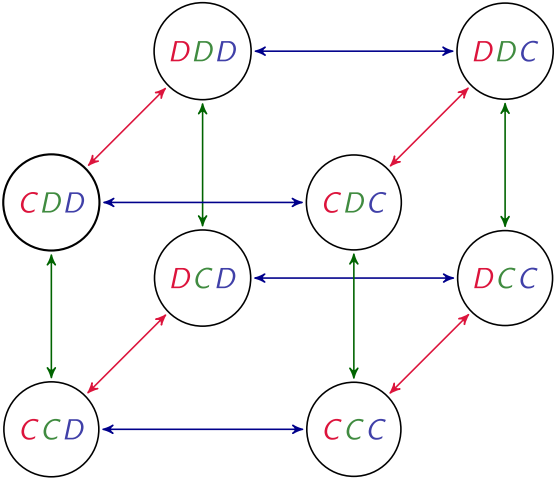
Three children—r, g, and b—have returned after playing outside. Their mother says to them, “At least one of you has mud on your forehead. Step forward now if you know whether you are dirty.”

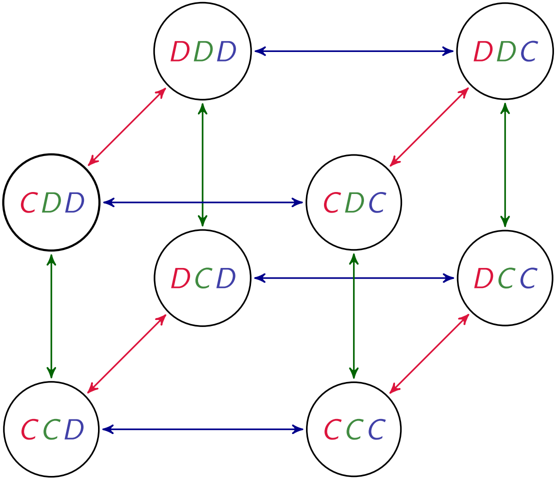
In fact, g and b are dirty, but no one can see his own forehead.

None of the children step forward. So their mother says again, “Step forward now if you know whether you are dirty.”

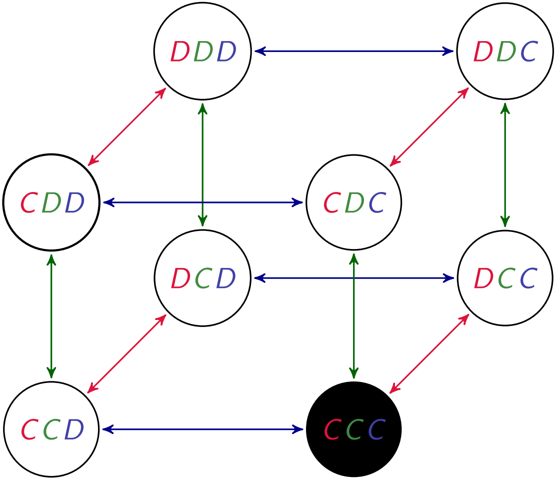
This time g and b both step forward, but r does not.

So their mother says again. “Step forward now if you know whether you are dirty.” Finally r steps forward.

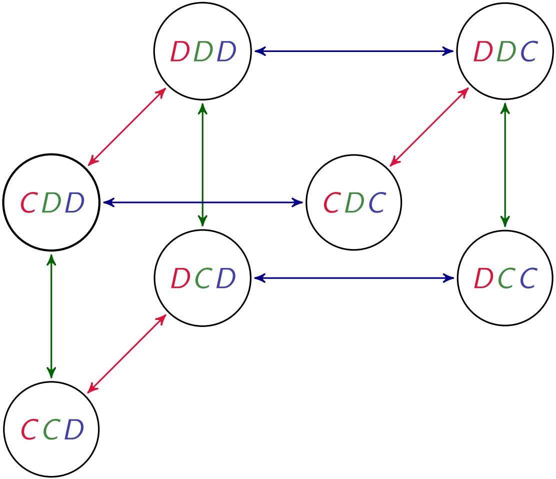

DDD
DDC
CDD
CDC
DCD
DCC
CCD
CCC



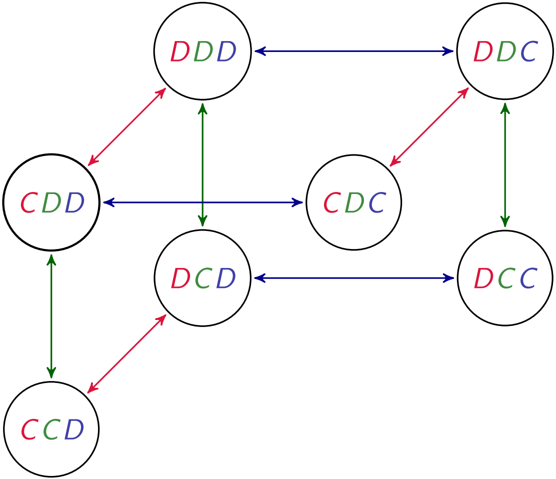
"At least one of you is dirty." $D_r \vee D_g \vee D_b$. What happens?



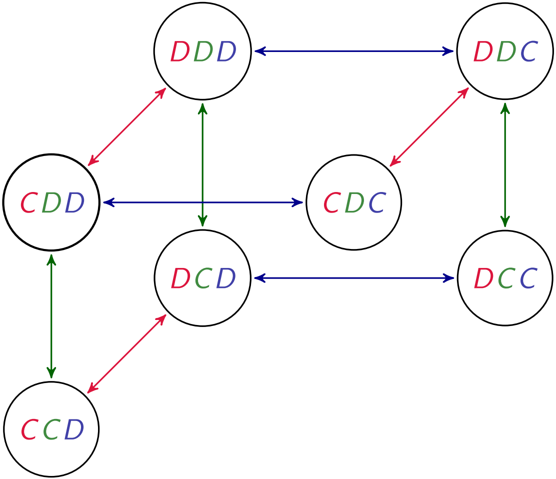
"At least one of you is dirty." $D_r \vee D_g \vee D_b$. What happens?



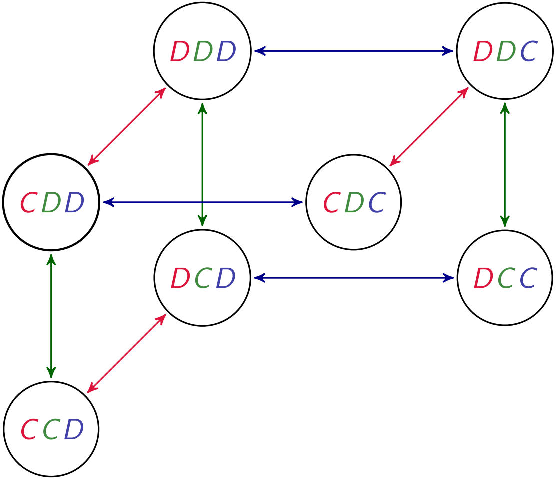
"At least one of you is dirty." $D_r \vee D_g \vee D_b$. What happens?



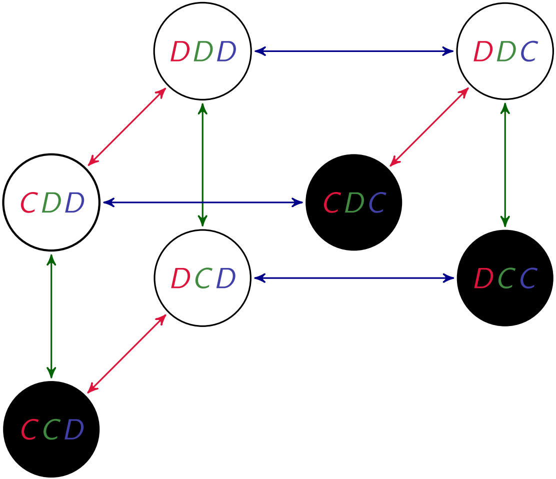
"If you know whether you're dirty, step forward."



"If you know whether you're dirty, step forward." No one does.
$\neg(K_r D_r \vee K_r \neg D_r) \wedge \neg(K_g D_g \vee K_g \neg D_g) \wedge \neg(K_b D_b \vee K_b \neg D_b)$.

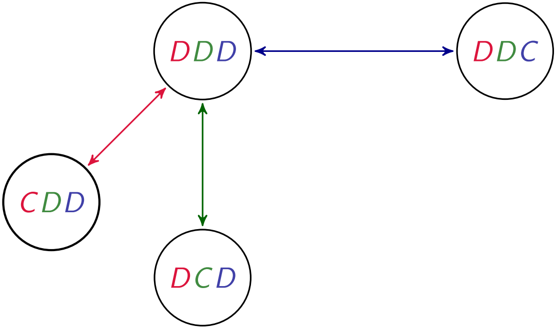


$\neg(K_r D_r \vee K_r \neg D_r) \wedge \neg(K_g D_g \vee K_g \neg D_g) \wedge \neg(K_b D_b \vee K_b \neg D_b)$.
What happens after this is announced?

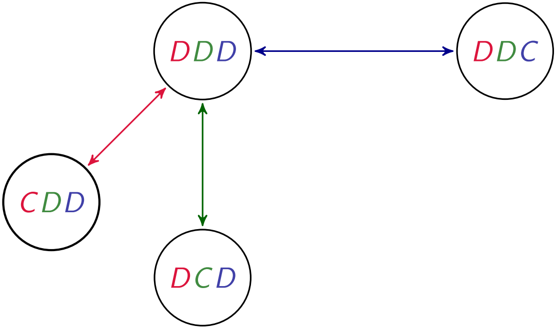


$\neg(K_r D_r \vee K_r \neg D_r) \wedge \neg(K_g D_g \vee K_g \neg D_g) \wedge \neg(K_b D_b \vee K_b \neg D_b)$.

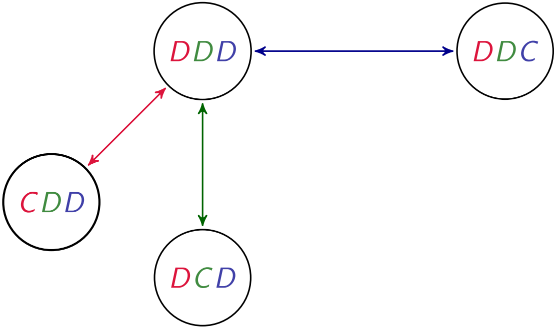
What happens after this? Three more possibilities are eliminated.



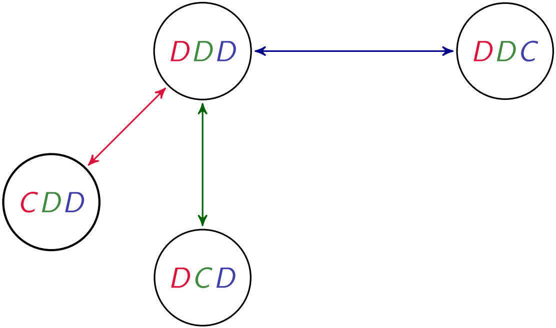
"If you know whether you're dirty, step forward."



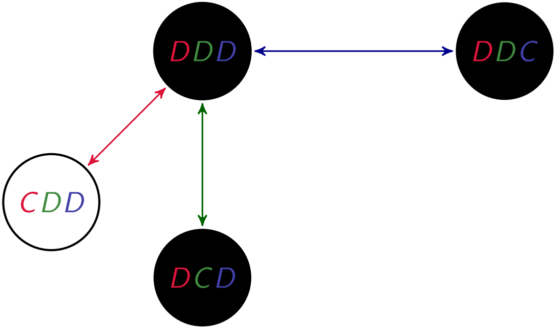
"If you know whether you're dirty, step forward." Blue and green (but not red) step forward.



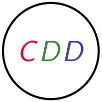
"If you know whether you're dirty, step forward." Blue and green (but not red) step forward. $(K_b D_b \vee K_b \neg D_b) \wedge (K_g D_g \vee K_g \neg D_g)$.



Blue and green (but not red) step forward. $(K_b D_b \vee K_b \neg D_b) \wedge (K_g D_g \vee K_g \neg D_g)$. What happens after this is announced?



$(K_b D_b \vee K_b \neg D_b) \wedge (K_g D_g \vee K_g \neg D_g)$. What happens after this?
The remaining non-actual possibilities are eliminated.



$(K_b D_b \vee K_b \neg D_b) \wedge (K_g D_g \vee K_g \neg D_g)$. What happens after this?
The remaining non-actual possibilities are eliminated.

One of the big ideas of **dynamic** epistemic logic is to add to our formal language operators that can describe the kinds of model changes that we just saw for the Muddy Children puzzle.

One of the big ideas of **dynamic** epistemic logic is to add to our formal language operators that can describe the kinds of model changes that we just saw for the Muddy Children puzzle.

Definition (Basic Epistemic Languages)

Given sets $\mathsf{At} = \{p, q, r, \dots\}$ and $\mathsf{Agt} = \{a, b, c, \dots\}$ with $|\mathsf{Agt}| = \kappa$, the language of public announcement logic, $\mathcal{L}^{\kappa}_{\mathsf{PAL}}$, is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi.$$

One of the big ideas of **dynamic** epistemic logic is to add to our formal language operators that can describe the kinds of model changes that we just saw for the Muddy Children puzzle.

Definition (Basic Epistemic Languages)

Given sets $\mathsf{At} = \{p, q, r, \ldots\}$ and $\mathsf{Agt} = \{a, b, c, \ldots\}$ with $|\mathsf{Agt}| = \kappa$, the language of public announcement logic, $\mathcal{L}^{\kappa}_{\mathsf{PAL}}$, is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi.$$

Read $[\varphi]\psi$ as "after (every) true announcement of φ, ψ."

One of the big ideas of **dynamic** epistemic logic is to add to our formal language operators that can describe the kinds of model changes that we just saw for the Muddy Children puzzle.

Definition (Basic Epistemic Languages)

Given sets $\text{At} = \{p, q, r, \dots\}$ and $\text{Agt} = \{a, b, c, \dots\}$ with $|\text{Agt}| = \kappa$, the language of public announcement logic, $\mathcal{L}^{\kappa}_{\text{PAL}}$, is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi.$$

Read $[\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle\varphi\rangle\psi := \neg[\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

One of the big ideas of **dynamic** epistemic logic is to add to our formal language operators that can describe the kinds of model changes that we just saw for the Muddy Children puzzle.

Definition (Basic Epistemic Languages)

Given sets $\mathsf{At} = \{p, q, r, \dots\}$ and $\mathsf{Agt} = \{a, b, c, \dots\}$ with $|\mathsf{Agt}| = \kappa$, the language of public announcement logic, $\mathcal{L}^{\kappa}_{\mathsf{PAL}}$, is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi.$$

Read $[\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle\varphi\rangle\psi := \neg[\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

One can also add common knowledge C_A to the language.

Definition (Basic Epistemic Languages)

Given sets $\mathsf{At} = \{p, q, r, \dots\}$ and $\mathsf{Agt} = \{a, b, c, \dots\}$ with $|\mathsf{Agt}| = \kappa$, the epistemic language $\mathcal{L}^{\kappa}_{\mathsf{EL}}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

For the epistemic language with common knowledge, $\mathcal{L}^{\kappa}_{\mathsf{EL\text{-}C}}$,

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid C\varphi.$$

For the language of public announcement logic, $\mathcal{L}^{\kappa}_{\mathsf{PAL}}$,

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi.$$

For public announcement with common knowledge, $\mathcal{L}^{\kappa}_{\mathsf{PAL\text{-}C}}$,

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi \mid [\varphi]\varphi \mid C\varphi.$$

Idea 3: Information Update as Model Restriction

Definition (Truth cont.)

Given a κ-agent model $\mathcal{M} = \langle W, \{R_a\}_{a \in \mathsf{Agt}}, V\rangle$ with $w \in W$ and $\varphi \in \mathcal{L}^{\kappa}_{\mathsf{PAL}}$, we define $\mathcal{M}, w \vDash \varphi$ as follows (other cases as before):

$$\mathcal{M}, w \vDash [\varphi]\psi \quad \text{iff} \quad \mathcal{M}, w \vDash \varphi \text{ implies } \mathcal{M}_{|\varphi}, w \vDash \psi,$$

where $\mathcal{M}_{|\varphi} = \langle W_{|\varphi}, \{R_{a_{|\varphi}}\}_{a \in \mathsf{Agt}}, V_{|\varphi}\rangle$ is the submodel such that

$$W_{|\varphi} = \{v \in W \mid \mathcal{M}, v \vDash \varphi\};$$
$$\forall a \in \mathsf{Agt}\colon R_{a_{|\varphi}} = R_a \cap (W_{|\varphi} \times W_{|\varphi});$$
$$\forall p \in \mathsf{At}\colon V_{|\varphi}(p) = V(p) \cap W_{|\varphi}.$$

Public Announcement Logic

Read $[\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle\varphi\rangle\psi := \neg[\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

Let's compare. The truth clause for the dynamic operator $[\varphi]$ is:

- $\mathcal{M}, w \vDash [\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

Public Announcement Logic

Read $[\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle\varphi\rangle\psi := \neg[\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

Let's compare. The truth clause for the dynamic operator $[\varphi]$ is:

- $\mathcal{M}, w \vDash [\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

So if φ is false, $[\varphi]\psi$ is vacuously true.

Public Announcement Logic

Read $[\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle\varphi\rangle\psi := \neg[\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

Let's compare. The truth clause for the dynamic operator $[\varphi]$ is:

- $\mathcal{M}, w \vDash [\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

So if φ is false, $[\varphi]\psi$ is vacuously true. Here is the $\langle\varphi\rangle$ clause:

- $\mathcal{M}, w \vDash \langle\varphi\rangle\psi$ iff $\mathcal{M}, w \vDash \varphi$ *and* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

Public Announcement Logic

Read $[\varphi]\psi$ as "after (every) true announcement of φ, ψ."

Read $\langle\varphi\rangle\psi := \neg[\varphi]\neg\psi$ as "after a true announcement of φ, ψ."

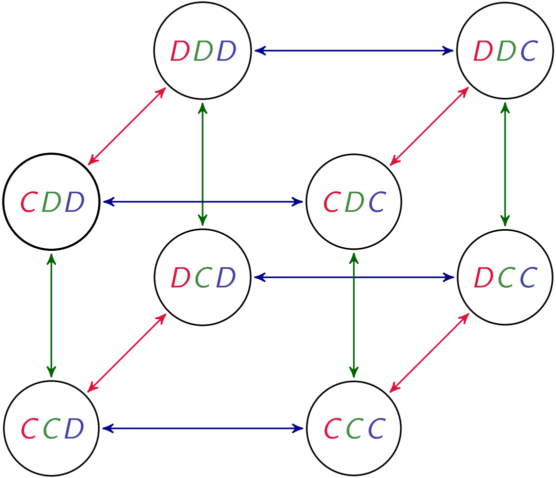
Let's compare. The truth clause for the dynamic operator $[\varphi]$ is:

- $\mathcal{M}, w \vDash [\varphi]\psi$ iff $\mathcal{M}, w \vDash \varphi$ *implies* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

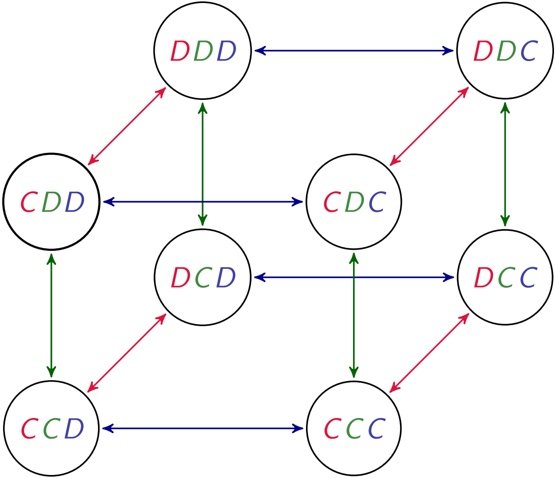
So if φ is false, $[\varphi]\psi$ is vacuously true. Here is the $\langle\varphi\rangle$ clause:

- $\mathcal{M}, w \vDash \langle\varphi\rangle\psi$ iff $\mathcal{M}, w \vDash \varphi$ *and* $\mathcal{M}_{|\varphi}, w \vDash \psi$.

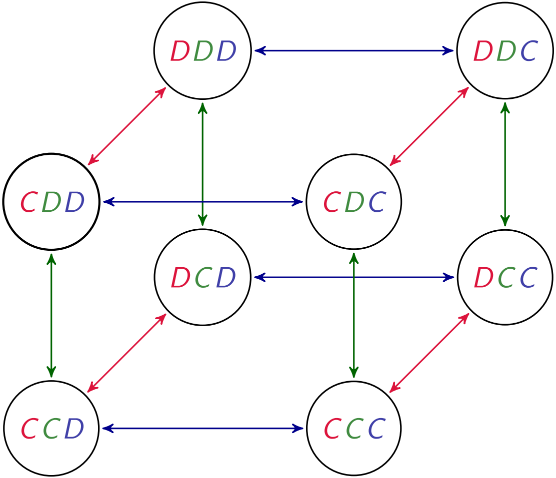
Big Idea: we evaluate $[\varphi]\psi$ and $\langle\varphi\rangle\psi$ not by looking at *other worlds in the same model*, but rather by looking at a new model.



Where $A_1 := D_r \vee D_g \vee D_b$, check $\mathcal{M}, CCD \vDash \langle A_1 \rangle K_b D_b$.

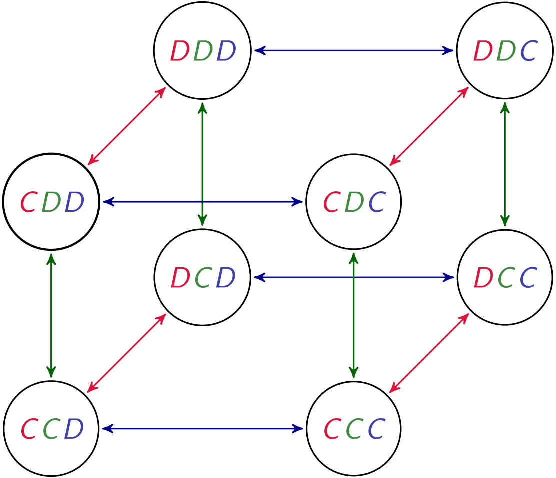


Where $A_1 := D_r \vee D_g \vee D_b$, check $\mathcal{M}, CCD \vDash \langle A_1 \rangle K_b D_b$.
A_1 is Mom's announcement: "At least one of you is dirty."



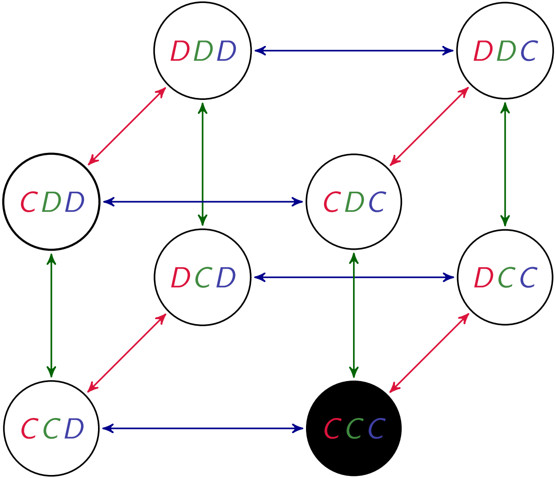
Where $A_1 := D_r \vee D_g \vee D_b$, check $\mathcal{M}, CCD \vDash \langle A_1 \rangle K_b D_b$.

First, observe that $\mathcal{M}, CCD \vDash D_r \vee D_g \vee D_b$.



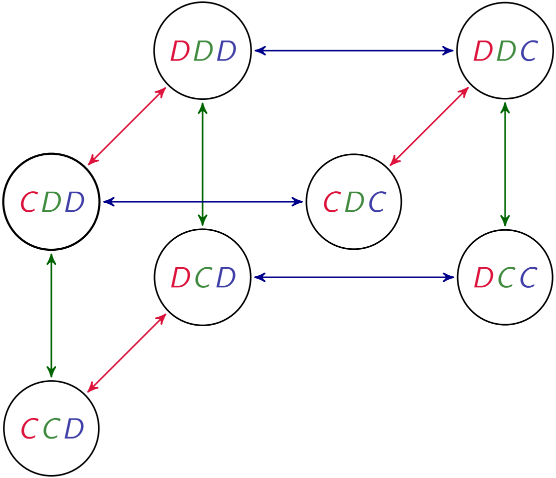
Where $A_1 := D_r \vee D_g \vee D_b$, check $\mathcal{M}, CCD \vDash \langle A_1 \rangle K_b D_b$.

Second, delete all worlds where $D_r \vee D_g \vee D_b$ is false.



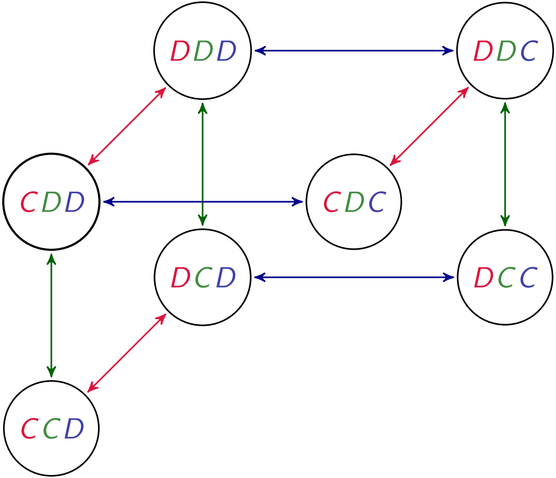
Where $A_1 := D_r \vee D_g \vee D_b$, check $\mathcal{M}, CCD \vDash \langle A_1 \rangle K_b D_b$.

Second, delete all worlds where $D_r \vee D_g \vee D_b$ is false.

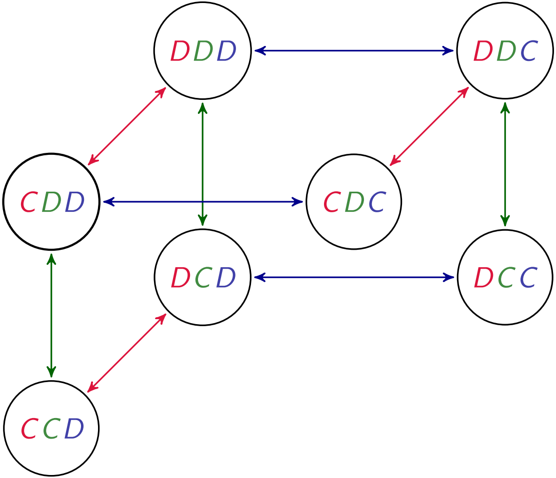


Where $A_1 := D_r \vee D_g \vee D_b$, check $\mathcal{M}, CCD \vDash \langle A_1 \rangle K_b D_b$.

Finally, observe that $\mathcal{M}_{|A_1}, CCD \vDash K_b D_b$.

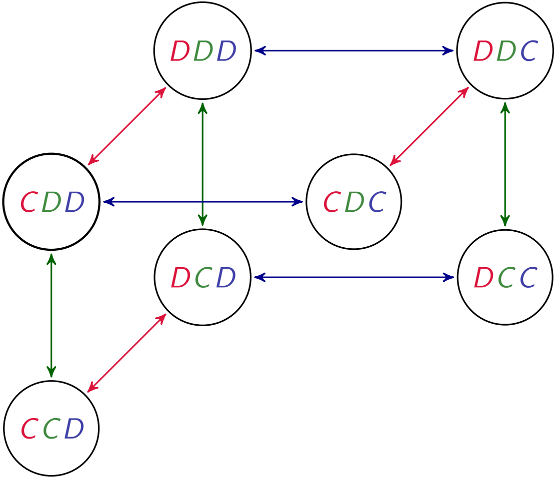


Where $A_2 := \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$, representing that no one steps forward, check $\mathcal{M}_{|A_1}, CDD \vDash \langle A_2 \rangle (K_g D_g \wedge K_b D_b)$.



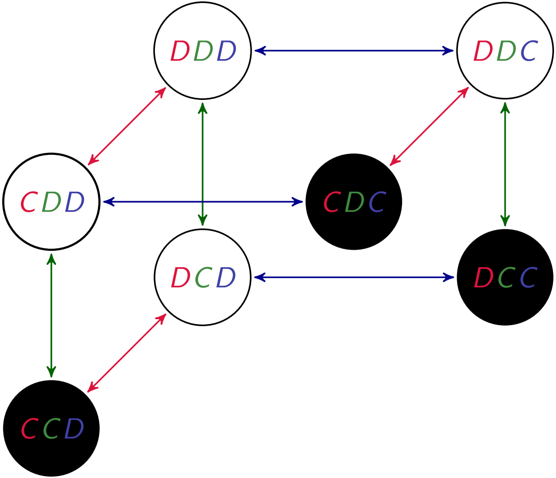
$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b)$.

First, note $\mathcal{M}_{|A_1}, CDD \vDash \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$.



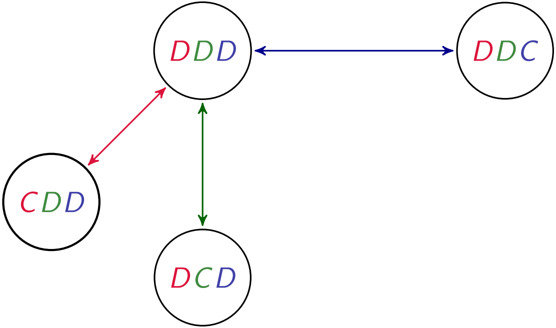
$$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b).$$

Second, delete all worlds where $\bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$ is false.



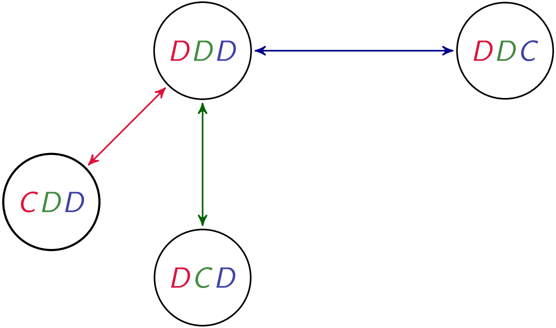
$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b)$.

Second, delete all worlds where $\bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$ is false.

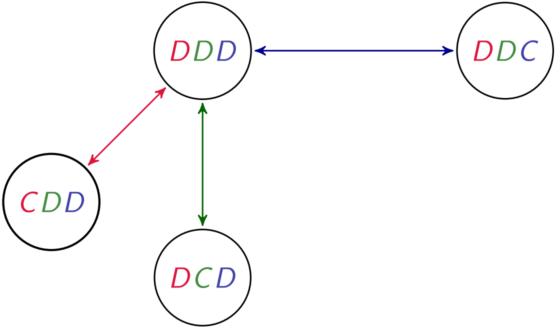


$$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b).$$

Finally, observe $\mathcal{M}_{|A_1| \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)}, CDD \vDash K_g D_g \wedge K_b D_b$.

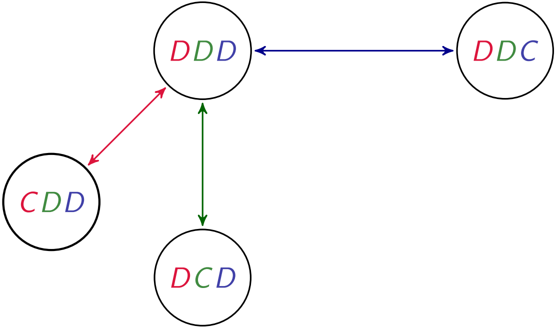


Check $\mathcal{M}_{|A_1|A_2}, CDD \vDash \langle \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a) \rangle K_r \neg D_r$.



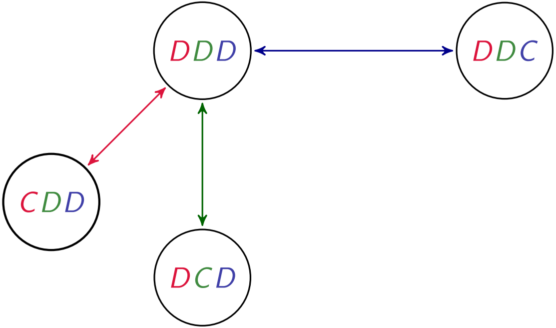
Check $\mathcal{M}_{|A_1|A_2}, CDD \vDash \langle \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a) \rangle K_r \neg D_r$.

This represents the info. given when g and b step forward after Mom says *again*, "If you know whether you're dirty, step forward."



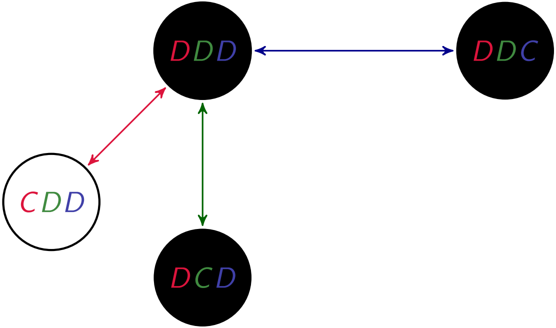
Check $\mathcal{M}_{|A_1|A_2}, CDD \vDash \langle \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a) \rangle K_r \neg D_r$.

First, observe $\mathcal{M}_{|A_1|A_2}, CDD \vDash \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a)$.



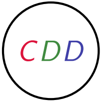
Check $\mathcal{M}_{|A_1|A_2}, CDD \vDash \langle \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a) \rangle K_r \neg D_r$.

Second, delete all worlds where $\bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a)$ is false.



Check $\mathcal{M}_{|A_1|A_2}, CDD \vDash \langle \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a) \rangle K_r \neg D_r$.

Second, delete all worlds where $\bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a)$ is false.



Check $\mathcal{M}_{|A_1|A_2}, CDD \vDash \langle \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a) \rangle K_r \neg D_r$.

Finally, observe $\mathcal{M}_{|A_1|A_2| \bigwedge_{a \in \{g,b\}} (K_a D_a \vee K_a \neg D_a)}, CDD \vDash K_r \neg D_r$.

Completeness and Complexity

Theorem (PAL Axiomatization (Plaza 1989))

The $\mathcal{L}^{\kappa}_{\mathsf{PAL}}$-theory of the class of all (resp. reflexive, preorder, partition) models is axiomatized by $\mathbf{K}_\omega$ (resp. $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$) plus:

i.	(replacement)	$\dfrac{\psi \leftrightarrow \chi}{\varphi(\psi/p) \leftrightarrow \varphi(\chi/p)}$
ii.	(atomic reduction)	$\langle\varphi\rangle p \leftrightarrow (\varphi \wedge p)$
iii.	(negation reduction)	$\langle\varphi\rangle\neg\psi \leftrightarrow (\varphi \wedge \neg\langle\varphi\rangle\psi)$
iv.	(conjunction reduction)	$\langle\varphi\rangle(\psi \wedge \chi) \leftrightarrow (\langle\varphi\rangle\psi \wedge \langle\varphi\rangle\chi)$
v.	(diamond reduction)	$\langle\varphi\rangle\hat{K}_a\psi \leftrightarrow (\varphi \wedge \hat{K}_a\langle\varphi\rangle\psi)$.

Complexity of satisfiability for PAL is the same as for EL.

With the background of basic epistemic logic in place, let's turn to mathematical themes arising in recent work in epistemic logic:

- Preservation Theorems
- Uniform Substitution
- Relations to FOL

Preservation of Truth (in Inquiry)

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Preservation of Truth (in Inquiry)

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Given our epistemic understanding of the models, this kind of preservation takes on a new meaning: once true, φ remains true henceforth in inquiry, no matter what new knowledge is acquired.

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Definition (Universal Fragment)

The universal fragment of $\mathcal{L}^{\kappa}_{\mathsf{EL}}$ is generated by:

$$\varphi ::= p \mid \neg p \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid K_a \varphi$$

for $p \in \mathsf{At}$ and $a \in \mathsf{Agt}$.

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Definition (Universal Fragment)

The universal fragment of $\mathcal{L}_{\mathsf{EL}}^{\kappa}$ is generated by:

$$\varphi ::= p \mid \neg p \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid K_a \varphi$$

for $p \in \mathsf{At}$ and $a \in \mathsf{Agt}$.

Theorem (Modal Łoś-Tarski)

A formula $\varphi \in \mathcal{L}_{\mathsf{EL}}^{\kappa}$ is preserved under submodels **iff** *it is equivalent to a formula in the universal fragment of $\mathcal{L}_{\mathsf{EL}}^{\kappa}$.*

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Definition (Universal Fragment)

The universal fragment of $\mathcal{L}^{\kappa}_{\text{PAL-C}}$ is generated by:

$$\varphi ::= p \mid \neg p \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid K_a\varphi \mid C_A\varphi \mid [\neg\varphi]\varphi$$

for $p \in \text{At}$ and $a \in \text{Agt}$.

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Definition (Universal Fragment)

The universal fragment of $\mathcal{L}^{\kappa}_{\text{PAL-C}}$ is generated by:

$$\varphi ::= p \mid \neg p \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid K_a\varphi \mid C_A\varphi \mid [\neg\varphi]\varphi$$

for $p \in \text{At}$ and $a \in \text{Agt}$.

Theorem (van Ditmarsch, Kooi)

A formula $\varphi \in \mathcal{L}^{\kappa}_{\text{PAL-C}}$ is preserved under submodels **if** *it is equivalent to a formula in the universal fragment of $\mathcal{L}^{\kappa}_{\text{PAL-C}}$.*

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Definition (Universal Fragment)

The universal fragment of $\mathcal{L}^{\kappa}_{\text{PAL-C}}$ is generated by:

$$\varphi ::= p \mid \neg p \mid (\varphi \wedge \varphi) \mid (\varphi \vee \varphi) \mid K_a\varphi \mid C_A\varphi \mid [\neg\varphi]\varphi$$

for $p \in \text{At}$ and $a \in \text{Agt}$.

Theorem (van Ditmarsch, Kooi)

A formula $\varphi \in \mathcal{L}^{\kappa}_{\text{PAL-C}}$ is preserved under submodels **if** *it is equivalent to a formula in the universal fragment of $\mathcal{L}^{\kappa}_{\text{PAL-C}}$.*

Open Problem 1: does 'only if' hold as well?

Definition (Preservation Under Submodels)

A formula φ is preserved under submodels iff for any pointed model $\mathcal{M}, w$ and $\mathcal{M}' \subseteq \mathcal{M}$: $\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}', w \vDash \varphi$.

Open Problem 1: Which formulas of $\mathcal{L}_{\text{EL-C}}$ and $\mathcal{L}_{\text{PAL-C}}$ (with common knowledge), syntactically characterized, are preserved?

(A place to start might be the proof of Łoś-Tarski for the μ-calculus (D'Agostino, 1997), of which $\mathcal{L}_{\text{EL-C}}$ is a fragment.)

Definition (Successful Formulas)

A formula φ is successful iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Definition (Successful Formulas)

A formula φ is successful iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Successful formulas are preserved under "self-defined" submodels.

Definition (Successful Formulas)

A formula φ is successful iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Successful formulas are preserved under "self-defined" submodels.

Not all formulas are successful. Some formulas are *unsuccessful*.

Definition (Successful Formulas)

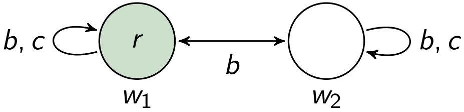
A formula φ is successful iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Successful formulas are preserved under "self-defined" submodels.

Not all formulas are successful. Some formulas are *unsuccessful*.

The simplest example: the Copenhagen agent c calls the Berkeley agent b on the phone and says "You don't know it's raining in Copenhagen, but it's raining in Copenhagen" ($\neg K_b r \wedge r$).



Definition (Successful Formulas)

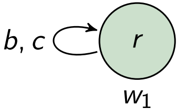
A formula φ is successful iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

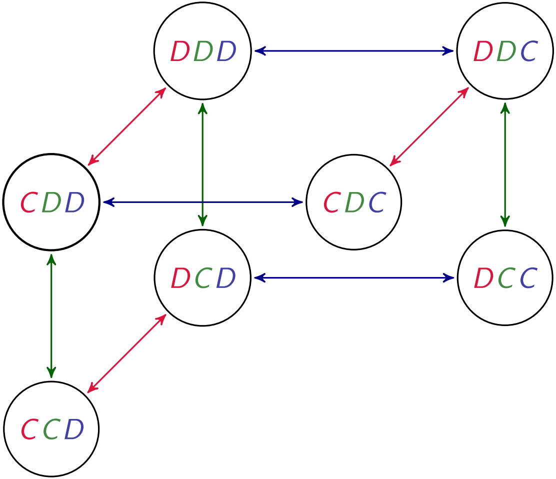
Successful formulas are preserved under "self-defined" submodels.

Not all formulas are successful. Some formulas are *unsuccessful*.

The simplest example: the Copenhagen agent c calls the Berkeley agent b on the phone and says "You don't know it's raining in Copenhagen, but it's raining in Copenhagen" ($\neg K_b r \wedge r$).

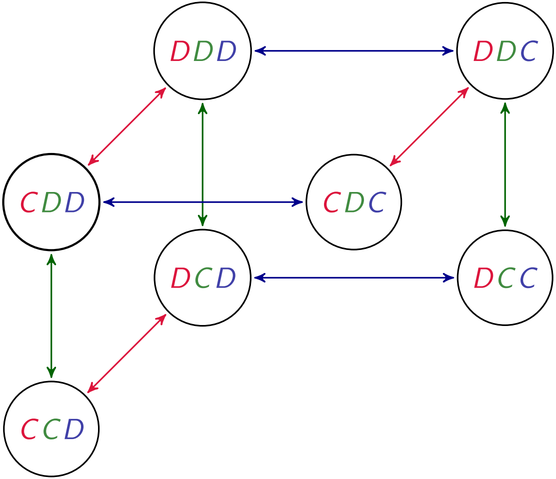


Let's go back to the first time that no students step forward. This is like their announcing, "We don't know whether we are dirty."



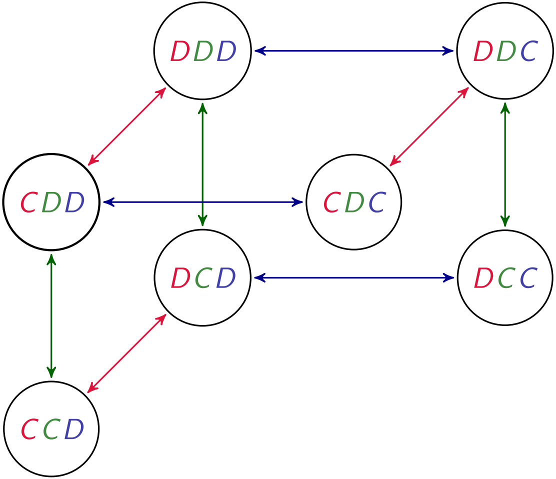
$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b)$.

This is like their saying "We don't know whether we're dirty."



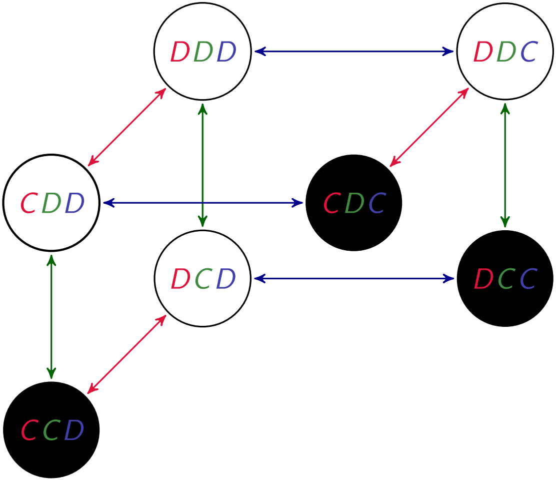
$$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b).$$

First, note $\mathcal{M}, CDD \vDash \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$.



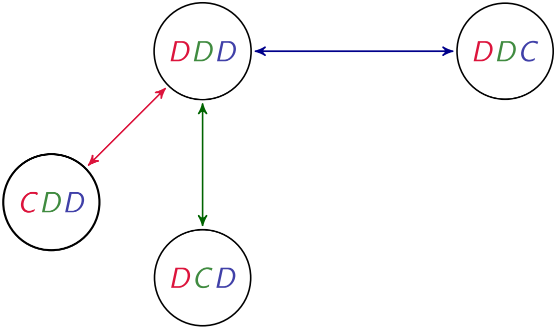
$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b).$

Second, delete all worlds where $\bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$ is false.



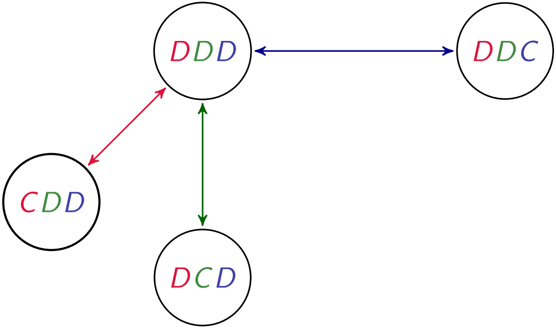
$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a\in\{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)\rangle (K_g D_g \wedge K_b D_b)$.

Second, delete all worlds where $\bigwedge_{a\in\{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$ is false.



$$\mathcal{M}_{|A_1}, CDD \vDash \langle \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a) \rangle (K_g D_g \wedge K_b D_b).$$

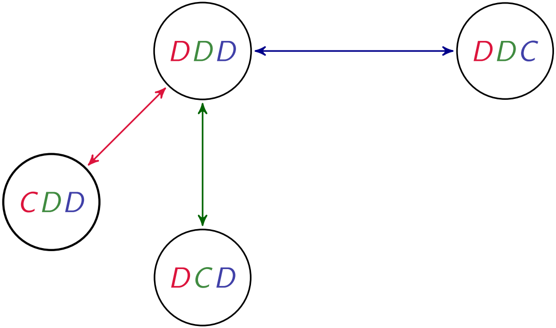
Finally, observe $\mathcal{M}_{|A_1| \bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)}, CDD \vDash K_g D_g \wedge K_b D_b$.



Finally, observe $\mathcal{M}_{|A_1|_{\bigwedge_{a\in\{r,g,b\}}\neg(K_aD_a\vee K_a\neg D_a)}}, CDD \vDash K_gD_g \wedge K_bD_b$.

Hence where $\varphi := \bigwedge_{a\in\{r,g,b\}} \neg(K_aD_a \vee K_a\neg D_a)$,

$$\mathcal{M}_{|A_1}, CDD \vDash \langle\varphi\rangle\neg\varphi.$$



Finally, observe $\mathcal{M}_{|A_1|_{\bigwedge_{a\in\{r,g,b\}}\neg(K_aD_a\vee K_a\neg D_a)}}, CDD \vDash K_gD_g \wedge K_bD_b$.

Hence where $\varphi := \bigwedge_{a\in\{r,g,b\}} \neg(K_aD_a \vee K_a\neg D_a)$,

$$\mathcal{M}_{|A_1}, CDD \vDash \langle\varphi\rangle\neg\varphi.$$

The childrens' **true** announcement that no one knows makes it **false** that no one knows! This is a so-called *unsuccessful* update.

Definition (Successful Formulas)

A formula φ is successful iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Successful formulas are preserved under "self-defined" submodels.

Not all formulas are successful. Some formulas are *unsuccessful*.

Example (Unsuccessful Formulas)

We've seen two examples so far:

- $\neg K_b r \wedge r$
- $\bigwedge_{a \in \{r,g,b\}} \neg(K_a D_a \vee K_a \neg D_a)$

Definition (Successful Formulas)

A formula φ is successful over a class of models C iff for all pointed models $\mathcal{M}, w$ with $\mathcal{M}$ in C,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Question

Which formulas of $\mathcal{L}_{\mathsf{EL}}^{\mathcal{K}}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Definition (Successful Formulas)

A formula φ is successful over a class of models C iff for all pointed models $\mathcal{M}, w$ with $\mathcal{M}$ in C,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Question

Which formulas of $\mathcal{L}^{\kappa}_{\mathsf{EL}}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Proposition (Complexity of the Success Problem)

The problem of deciding whether a formula $\varphi \in \mathcal{L}^{1}_{\mathsf{EL}}$ is successful over the class of partition models is co-NP complete.

Definition (Successful Formulas)

A formula φ is successful over a class of models C iff for all pointed models $\mathcal{M}, w$ with $\mathcal{M}$ in C,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \vDash \varphi.$$

Question

Which formulas of $\mathcal{L}_{\mathsf{EL}}^{\kappa}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Proposition (Complexity of the Success Problem)

The problem of deciding whether a formula $\varphi \in \mathcal{L}_{\mathsf{EL}}^{1}$ is successful over the class of partition models is co-NP complete.

So we cannot expect a very simple syntactic characterization.

Question

Which formulas of $\mathcal{L}^{\mathcal{K}}_{\mathsf{EL}}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Question

Which formulas of $\mathcal{L}_{\mathsf{EL}}^{\mathcal{K}}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Q2 of van Benthem's "Open Problems in Logical Dynamics."

Question

Which formulas of $\mathcal{L}^{\kappa}_{\mathsf{EL}}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Q2 of van Benthem's "Open Problems in Logical Dynamics."

For the case $\kappa = 1$ with partition models, this was answered in:

Wesley Holliday and Thomas Icard. 2010.
"Moorean Phenomena in Epistemic Logic." *Advances in Modal Logic* **8**.

Question

Which formulas of $\mathcal{L}^{\kappa}_{\mathsf{EL}}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Q2 of van Benthem's "Open Problems in Logical Dynamics."

For the case $\kappa = 1$ with partition models, this was answered in:

Wesley Holliday and Thomas Icard. 2010.
"Moorean Phenomena in Epistemic Logic." *Advances in Modal Logic* **8**.

In essence, we show that all *un*successful formulas involve variations on the Moorean theme of "you don't know it, but *p*."

Question

Which formulas of $\mathcal{L}^{\kappa}_{\mathsf{EL}}$, syntactically characterized, are successful over the class of all models? The class of partition models?

Q2 of van Benthem's "Open Problems in Logical Dynamics."

For the case $\kappa = 1$ with partition models, this was answered in:

Wesley Holliday and Thomas Icard. 2010.

"Moorean Phenomena in Epistemic Logic." *Advances in Modal Logic* **8**.

In essence, we show that all *un*successful formulas involve variations on the Moorean theme of "you don't know it, but *p*."

Open Problem 2: answer the question for $\kappa > 1$ or for non-partition models (cf. Saraf and Sourabh for some leads).

Other Interesting Formula Classes

Definition (Self-Refuting)

A formula φ is self-refuting iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \nvDash \varphi.$$

Other Interesting Formula Classes

Definition (Self-Refuting)

A formula φ is self-refuting iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \nvDash \varphi.$$

Definition (Eventually Self-Refuting)

Define $\mathcal{M}_{|^n\varphi}$ recursively by $\mathcal{M}_{|^0\varphi} = \mathcal{M}$, $\mathcal{M}_{|^{n+1}\varphi} = \left(\mathcal{M}_{|^n\varphi}\right)_{|\varphi}$.
φ is eventually self-refuting iff for all pointed models $\mathcal{M}, w$:

$$\mathcal{M}, w \vDash \varphi \Rightarrow \exists n : \mathcal{M}_{|^n\varphi}, w \nvDash \varphi.$$

Other Interesting Formula Classes

Definition (Self-Refuting)

A formula φ is self-refuting iff for all pointed models $\mathcal{M}, w$,

$$\mathcal{M}, w \vDash \varphi \Rightarrow \mathcal{M}_{|\varphi}, w \nvDash \varphi.$$

Definition (Eventually Self-Refuting)

Define $\mathcal{M}_{|^n\varphi}$ recursively by $\mathcal{M}_{|^0\varphi} = \mathcal{M}$, $\mathcal{M}_{|^{n+1}\varphi} = \left(\mathcal{M}_{|^n\varphi}\right)_{|\varphi}$.
φ is eventually self-refuting iff for all pointed models $\mathcal{M}, w$:

$$\mathcal{M}, w \vDash \varphi \Rightarrow \exists n : \mathcal{M}_{|^n\varphi}, w \nvDash \varphi.$$

Definition (Always Learnable)

φ is always learnable (for a) iff for any pointed model $\mathcal{M}, w$:

$$\mathcal{M}, w \vDash \varphi \Rightarrow \exists \psi : \mathcal{M}_{|\psi}, w \vDash K_a\varphi.$$

Substitution Closure

A striking feature of epistemic logic with operators for information update is that it is not closed under uniform substitution: e.g., $[p]p$ is valid, but $[\neg K_b r \wedge r](\neg K_b r \wedge r)$ is not valid, as we saw.

Definition (Substitution Core & Schematic Validity)

A substitution is any function $\sigma\colon \mathsf{At} \to \mathcal{L}$; and we define $(\cdot)^\sigma\colon \mathcal{L} \to \mathcal{L}$ as the extension such that $(\varphi)^\sigma$ is obtained from φ by replacing, for all $p \in \mathsf{At}(\varphi)$, each occurrence of p in φ by $\sigma(p)$.

Definition (Substitution Core & Schematic Validity)

A substitution is any function $\sigma\colon \mathsf{At} \to \mathcal{L}$; and we define $(\cdot)^\sigma\colon \mathcal{L} \to \mathcal{L}$ as the extension such that $(\varphi)^\sigma$ is obtained from φ by replacing, for all $p \in \mathsf{At}(\varphi)$, each occurrence of p in φ by $\sigma(p)$.

The substitution core of logic L is the set

$$\{\varphi \in \mathcal{L}_{\mathsf{L}} \mid (\varphi)^\sigma \in \mathsf{L} \text{ for all substitutions } \sigma\}.$$

Formulas in the substitution core are schematically valid.

Example (Valid but not Schematically Valid)

Where C is the class of all models, formulas that are in $\mathrm{Th}_{\mathcal{L}^{\kappa}_{\mathrm{PAL}}}(C)$ but are not in the substitution core of $\mathrm{Th}_{\mathcal{L}^{\kappa}_{\mathrm{PAL}}}(C)$ include ($\kappa \geq 1$):

$$[p]p \qquad K_a p \to [p]K_a p$$
$$[p]K_a p \qquad K_a p \to [p](p \to K_a p)$$
$$[p](p \to K_a p) \qquad K_a(p \to q) \to (\langle q\rangle K_a r \to \langle p\rangle K_a r)$$
$$[p \wedge \neg K_a p]\neg(p \wedge \neg K_a p) \qquad (\langle p\rangle K_a r \wedge \langle q\rangle K_a r) \to \langle p \vee q\rangle K_a r.$$

Example (Valid but not Schematically Valid)

Where C is the class of all models, formulas that are in $\mathsf{Th}_{\mathcal{L}^{\kappa}_{\mathsf{PAL}}}(C)$ but are not in the substitution core of $\mathsf{Th}_{\mathcal{L}^{\kappa}_{\mathsf{PAL}}}(C)$ include ($\kappa \geq 1$):

$[p]p$	$K_a p \to [p]K_a p$
$[p]K_a p$	$K_a p \to [p](p \to K_a p)$
$[p](p \to K_a p)$	$K_a(p \to q) \to (\langle q \rangle K_a r \to \langle p \rangle K_a r)$
$[p \wedge \neg K_a p]\neg(p \wedge \neg K_a p)$	$(\langle p \rangle K_a r \wedge \langle q \rangle K_a r) \to \langle p \vee q \rangle K_a r.$

So what are the uniform principles of information dynamics?

Question

Is the substitution core of PAL axiomatizable?

Q1 in van Benthem's "Open Problems in Logical Dynamics."

Question

Is the substitution core of PAL axiomatizable?

Q1 in van Benthem's "Open Problems in Logical Dynamics."

The answer is given in:

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. 2012.

"A Uniform Logic of Information Dynamics." *Advances in Modal Logic* **9**.

Theorem (Axiomatization)

For the class C of all (resp. reflexive, preorder, partition) ω-agent models, the substitution core of $Th_{\mathcal{L}^{\omega}_{\mathsf{PAL}}}(C)$ is axiomatizable.

Question

Is the substitution core of PAL axiomatizable?

Q1 in van Benthem's "Open Problems in Logical Dynamics."

The answer is given in:

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. 2012.

"A Uniform Logic of Information Dynamics." *Advances in Modal Logic* **9**.

Theorem (Axiomatization)

For the class C of all (resp. reflexive, preorder, partition) ω-agent models, the substitution core of $Th_{\mathcal{L}^{\omega}_{\mathsf{PAL}}}(C)$ is axiomatizable.

Upshot: we have a logic for reasoning about information change in which the atomic sentences p can be genuine propositional variables, standing in for any propositions, even epistemic ones.

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$
3. (extensionality) $\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$
3. (extensionality) $\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$
4. (distribution) $[p](q \to r) \to ([p]q \to [p]r)$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$
3. (extensionality) $\dfrac{\varphi \leftrightarrow \psi}{\langle \varphi \rangle p \leftrightarrow \langle \psi \rangle p}$
4. (distribution) $[p](q \to r) \to ([p]q \to [p]r)$
5. (p-seriality) $p \to \langle p \rangle \top$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$
3. (extensionality) $\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$
4. (distribution) $[p](q \to r) \to ([p]q \to [p]r)$
5. (p-seriality) $p \to \langle p\rangle\top$
6. (truthfulness) $\langle p\rangle\top \to p$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$
3. (extensionality) $\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$
4. (distribution) $[p](q \to r) \to ([p]q \to [p]r)$
5. (p-seriality) $p \to \langle p\rangle\top$
6. (truthfulness) $\langle p\rangle\top \to p$
7. ($\top$-reflexivity) $p \to \langle\top\rangle p$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$
3. (extensionality) $\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$
4. (distribution) $[p](q \to r) \to ([p]q \to [p]r)$
5. (p-seriality) $p \to \langle p\rangle\top$
6. (truthfulness) $\langle p\rangle\top \to p$
7. ($\top$-reflexivity) $p \to \langle\top\rangle p$
8. (functionality) $\langle p\rangle q \to [p]q$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1. (uniformity) $\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2. (necessitation) $\dfrac{\varphi}{[p]\varphi}$
3. (extensionality) $\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$
4. (distribution) $[p](q \to r) \to ([p]q \to [p]r)$
5. (p-seriality) $p \to \langle p\rangle\top$
6. (truthfulness) $\langle p\rangle\top \to p$
7. ($\top$-reflexivity) $p \to \langle\top\rangle p$
8. (functionality) $\langle p\rangle q \to [p]q$
9. (pa-commutativity) $\langle p\rangle\hat{K}_a q \to \hat{K}_a\langle p\rangle q$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1.	(uniformity)	$\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2.	(necessitation)	$\dfrac{\varphi}{[p]\varphi}$
3.	(extensionality)	$\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$
4.	(distribution)	$[p](q \to r) \to ([p]q \to [p]r)$
5.	(p-seriality)	$p \to \langle p\rangle\top$
6.	(truthfulness)	$\langle p\rangle\top \to p$
7.	($\top$-reflexivity)	$p \to \langle\top\rangle p$
8.	(functionality)	$\langle p\rangle q \to [p]q$
9.	(pa-commutativity)	$\langle p\rangle\hat{K}_a q \to \hat{K}_a\langle p\rangle q$
10.	(ap-commutativity)	$\hat{K}_a\langle p\rangle q \to [p]\hat{K}_a q$

The axiomatization is obtained by adding to the normal modal logic $\mathbf{K}_\omega$ (or $\mathbf{T}_\omega$, $\mathbf{S4}_\omega$, $\mathbf{S5}_\omega$ depending on the model class):

1.	(uniformity)	$\dfrac{\varphi}{(\varphi)^\sigma}$ for any substitution σ
2.	(necessitation)	$\dfrac{\varphi}{[p]\varphi}$
3.	(extensionality)	$\dfrac{\varphi \leftrightarrow \psi}{\langle\varphi\rangle p \leftrightarrow \langle\psi\rangle p}$
4.	(distribution)	$[p](q \to r) \to ([p]q \to [p]r)$
5.	(p-seriality)	$p \to \langle p\rangle\top$
6.	(truthfulness)	$\langle p\rangle\top \to p$
7.	($\top$-reflexivity)	$p \to \langle\top\rangle p$
8.	(functionality)	$\langle p\rangle q \to [p]q$
9.	(pa-commutativity)	$\langle p\rangle\hat{K}_a q \to \hat{K}_a\langle p\rangle q$
10.	(ap-commutativity)	$\hat{K}_a\langle p\rangle q \to [p]\hat{K}_a q$
11.	(composition)	$\langle p\rangle\langle q\rangle r \leftrightarrow \langle\langle p\rangle q\rangle r$

Theorem (Axiomatization)

For the class C of all (resp. reflexive, preorder, partition) ω-agent models, the substitution core of $Th_{\mathcal{L}^{\omega}_{\mathsf{PAL}}}(C)$ is axiomatizable.

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. 2012.

"A Uniform Logic of Information Dynamics." *Advances in Modal Logic* **9**.

Theorem (Axiomatization)

For the class C of all (resp. reflexive, preorder, partition) ω-agent models, the substitution core of $Th_{\mathcal{L}^{\omega}_{\mathsf{PAL}}}(C)$ is axiomatizable.

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. 2012.
"A Uniform Logic of Information Dynamics." *Advances in Modal Logic* **9**.

Theorem (Decidability)

For the class C of all (resp. reflexive, preorder, partition) ω-agent models, the substitution core of $Th_{\mathcal{L}^{\omega}_{\mathsf{PAL}}}(C)$ is decidable.

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. Forthcoming.
"Information Dynamics and Uniform Substitution." *Synthese*.

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. 2011.
"Schematic Validity in Dynamic Epistemic Logic: Decidability." *LNAI*.

Theorem (Axiomatization)

For the class C of all (resp. reflexive, preorder, partition) ω-agent models, the substitution core of $Th_{\mathcal{L}^{\omega}_{\mathsf{PAL}}}(C)$ is axiomatizable.

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. 2012.

"A Uniform Logic of Information Dynamics." *Advances in Modal Logic* **9**.

Open Problem 3: Does this extend to $\mathsf{Th}_{\mathcal{L}^{\omega}_{\mathsf{PAL-C}}}(C)$? Is there a Uniform Logic of Information Dynamics & Common Knowledge?

Theorem (Axiomatization)

For the class C of all (resp. reflexive, preorder, partition) ω-agent models, the substitution core of $Th_{\mathcal{L}^{\omega}_{\mathsf{PAL}}}(C)$ is axiomatizable.

Wesley Holliday, Tomohiro Hoshi, and Thomas Icard. 2012.

"A Uniform Logic of Information Dynamics." *Advances in Modal Logic* **9**.

Open Problem 3: Does this extend to $\mathsf{Th}_{\mathcal{L}^{\omega}_{\mathsf{PAL-C}}}(C)$? Is there a Uniform Logic of Information Dynamics & Common Knowledge?

(The obstacle is that given the failure of compactness, we must use a finite version of the Henkin-style canonical model construction to prove completeness; but the use of the standard Fisher-Ladner closure to do so proves difficult in this case...)

Transferring the Questions to FOL

In "Open Problems in Logical Dynamics," van Benthem poses the questions we have been discussing for FOL as well.

Open Problem 4: Characterize the formulas φ of FOL with exactly one free variable x such that for any FO structure $\mathfrak{A}$ and variable assignment s, if $\mathfrak{A}, s \vDash \varphi$, then $\mathfrak{A}_\varphi, s \vDash \varphi$, where $\mathfrak{A}_\varphi$ is the substructure of $\mathfrak{A}$ with domain $\{d \in |\mathfrak{A}| \mid \mathfrak{A}, s_{[x:=d]} \vDash \varphi\}$.

Transferring the Questions to FOL

In "Open Problems in Logical Dynamics," van Benthem poses the questions we have been discussing for FOL as well.

Open Problem 5: Extend the language of FOL with relativization operators $(\cdot)^\varphi$ for all φ with exactly one free variable x, such that

$$\mathfrak{A}, s \vDash (\psi)^\varphi \Leftrightarrow \text{if } \forall y \in FV(\psi)\, \mathfrak{A}, s_{[x:=s(y)]} \vDash \varphi, \text{ then } \mathfrak{A}_\varphi, s \vDash \psi,$$

where $FV(\psi)$ is the set of free variables in ψ.

What is the complete logic of FOL with relativization?

Transferring the Questions to FOL

In “Open Problems in Logical Dynamics,” van Benthem poses the questions we have been discussing for FOL as well.

Open Problem 5: Extend the language of FOL with relativization operators $(\cdot)^{\varphi}$ for all φ with exactly one free variable x, such that

$$\mathfrak{A}, s \vDash (\psi)^{\varphi} \Leftrightarrow \text{if } \forall y \in FV(\psi)\, \mathfrak{A}, s_{[x:=s(y)]} \vDash \varphi, \text{ then } \mathfrak{A}_{\varphi}, s \vDash \psi,$$

where $FV(\psi)$ is the set of free variables in ψ.

What is the complete logic of FOL with relativization?

Since this logic is not closed under uniform substitution (e.g., $Px \rightarrow (Px)^{Qz}$ is valid, but $(Px \wedge \exists y \neg Py) \rightarrow (Px \wedge \exists y \neg Py)^{Qz}$ is not), is the substitution core of this logic axiomatizable?

Translating EL into FOL

Can we transfer our results for modal logic to fragments of FOL?

Translating EL into FOL

Can we transfer our results for modal logic to fragments of FOL?

Definition (Propositional Epistemic Language)

Given sets $\text{At} = \{p, q, r, \dots\}$ and $\text{Agt} = \{a, b, c, \dots\}$ with $|\text{Agt}| = \kappa$, the epistemic language $\mathcal{L}^{\kappa}_{\text{EL}}$ is generated by

$$\varphi ::= p \mid \neg\varphi \mid (\varphi \wedge \varphi) \mid K_a\varphi.$$

Consider a first-order language that contains, for every $p \in \text{At}$ and $a \in \text{Agt}$, a one-place predicate $\mathbf{P}$ and a two-place predicate $\mathbf{R}_a$.

Translating EL into FOL

Definition (Standard Translation)

Fix two variables x and y and define functions ST_x and ST_y:

- $ST_x(p) = \mathbf{P}x$;
- $ST_x(\neg\varphi) = \neg ST_x(\varphi)$;
- $ST_x(\varphi \wedge \psi) = ST_x(\varphi) \wedge ST_x(\psi)$;
- $ST_x(K_a\varphi) = \forall y(\mathbf{R}_a xy \to ST_y(\varphi))$

- $ST_y(p) = \mathbf{P}y$;
- $ST_y(\neg\varphi) = \neg ST_y(\varphi)$;
- $ST_y(\varphi \wedge \psi) = ST_y(\varphi) \wedge ST_y(\psi)$;
- $ST_y(K_a\varphi) = \forall y(\mathbf{R}_a xy \to ST_x(\varphi))$

Translating EL into FOL

Definition (Standard Translation)

Fix two variables x and y and define functions ST_x and ST_y:

- $ST_x(p) = \mathbf{P}x$;
- $ST_x(\neg\varphi) = \neg ST_x(\varphi)$;
- $ST_x(\varphi \wedge \psi) = ST_x(\varphi) \wedge ST_x(\psi)$;
- $ST_x(K_a\varphi) = \forall y(\mathbf{R}_a xy \rightarrow ST_y(\varphi))$;
- and similarly for ST_y.

Any modal model $\mathcal{A}$ can be viewed as a first-order structure $\mathfrak{A}$ such that for any $\varphi \in \mathcal{L}_{\mathrm{EL}}^{\kappa}$, variable x, and assignment s,

$$\mathcal{A}, w \vDash \varphi \text{ iff } \mathfrak{A}, s_{[x:=w]} \vDash ST_x(\varphi)$$

Under the translation, the modal language lives in the two-variable fragment of FOL. It also lives in the "guarded" fragment of FOL, since all the quantifiers have guards. The satisfiability/validity problem for formulas in both of these fragment is decidable.

Under the translation, the modal language lives in the two-variable fragment of FOL. It also lives in the "guarded" fragment of FOL, since all the quantifiers have guards. The satisfiability/validity problem for formulas in both of these fragment is decidable.

Van Benthem's Characterization Theorem gives a *semantic characterization* of the FOL formulas (with one free variable) that are equivalent to the standard translation of some modal formula: they are the FOL formulas that are invariant for bisimulation.

Monadic First-Order Logic

In the other direction, any formula φ of Monadic First-Order Logic with only x free (and not in the scope of any quantifier) can be translated into a modal formula φ' such that for any MFO structure $\mathfrak{A}$ and assignment s, when we view $\mathfrak{A}$ as a single-agent epistemic partition model $\mathcal{A}$ (with R_a the universal relation),

$$\mathfrak{A}, s \vDash \varphi \text{ iff } \mathcal{A}, s(x) \vDash \varphi'.$$

Monadic First-Order Logic

In the other direction, any formula φ of Monadic First-Order Logic with only x free (and not in the scope of any quantifier) can be translated into a modal formula φ' such that for any MFO structure $\mathfrak{A}$ and assignment s, when we view $\mathfrak{A}$ as a single-agent epistemic partition model $\mathcal{A}$ (with R_a the universal relation),

$$\mathfrak{A}, s \vDash \varphi \text{ iff } \mathcal{A}, s(x) \vDash \varphi'.$$

Since we have solved the success problem and substitution core problem for single-agent epistemic logic over partition models, we can transfer these results to the modal-like fragment of MFOL.

Monadic First-Order Logic

In the other direction, any formula φ of Monadic First-Order Logic with only x free (and not in the scope of any quantifier) can be translated into a modal formula φ' such that for any MFO structure $\mathfrak{A}$ and assignment s, when we view $\mathfrak{A}$ as a single-agent epistemic partition model $\mathcal{A}$ (with R_a the universal relation),

$$\mathfrak{A}, s \vDash \varphi \text{ iff } \mathcal{A}, s(x) \vDash \varphi'.$$

Since we have solved the success problem and substitution core problem for single-agent epistemic logic over partition models, we can transfer these results to the modal-like fragment of MFOL.

But it is a long way from here to answers for full FOL.

Directions from Here

Going beyond the basic epistemic logic we have covered today:

- Quantified epistemic logic ($K_a \exists x Spy(x)$ vs. $\exists x K_a Spy(x)$);
- "Softer" belief revision instead of "hard" knowledge update;
- Modeling private communication, eavesdropping, etc.;
- Epistemic logic and probability (qualitative and quantitative);
- Philosophical applications of EL, mentioned at the beginning;
- and more...

Directions from Here

Going beyond the basic epistemic logic we have covered today:

- Quantified epistemic logic ($K_a \exists x Spy(x)$ vs. $\exists x K_a Spy(x)$);
- "Softer" belief revision instead of "hard" knowledge update;
- Modeling private communication, eavesdropping, etc.;
- Epistemic logic and probability (qualitative and quantitative);
- Philosophical applications of EL, mentioned at the beginning;
- and more...

Two sources for learning about recents research directions are:

- *Advances in Modal Logic*, biennial conference and book series.
- *European Summer School for Logic, Language, and Information*.

Thank you!

Sum and Product

A says to S and P: "I have picked $x, y \in \mathbb{N}$ such that $1 < x < y$ and $x + y \leq 100$. Shortly I will whisper their sum $s = x + y$ to S only, and their product $p = x \cdot y$ to P only." A acts accordingly. The following conversation between S and P then takes place:

1. P: "I don't know the numbers."
2. S: "I knew that you didn't know them."
3. P: "Now I know the numbers."
4. S: "Now I know them too."

Question: what are x and y? (There is a unique answer.)