Introduction to Formal Epistemology

Lecture 1

Eric Pacuit and Rohit Parikh

August 13, 2007

Lecture 1: Introduction, Motivation and Basic Models of Knowledge

Lecture 2: Knowledge in Groups and Group Knowledge

Lecture 3: Reasoning about Knowledge and ....

Lecture 4: Logical Omniscience and Other Problems

Lecture 5: Reasoning about Knowledge in the Context of Social Software

Lecture 1: Introduction, Motivation and Basic Models of Knowledge

Lecture 2: Knowledge in Groups and Group Knowledge

Lecture 3: Reasoning about Knowledge and ....

Lecture 4: Logical Omniscience and Other Problems

Lecture 5: Reasoning about Knowledge in the Context of Social Software

Lecture 1: Introduction, Motivation and Basic Models of Knowledge

Lecture 2: Knowledge in Groups and Group Knowledge

Lecture 3: Reasoning about Knowledge and ....

Lecture 4: Logical Omniscience and Other Problems

Lecture 5: Reasoning about Knowledge in the Context of Social Software

Lecture 1: Introduction, Motivation and Basic Models of Knowledge

Lecture 2: Knowledge in Groups and Group Knowledge

Lecture 3: Reasoning about Knowledge and ....

Lecture 4: Logical Omniscience and Other Problems

Lecture 5: Reasoning about Knowledge in the Context of Social Software

Lecture 1: Introduction, Motivation and Basic Models of Knowledge

Lecture 2: Knowledge in Groups and Group Knowledge

Lecture 3: Reasoning about Knowledge and ....

Lecture 4: Logical Omniscience and Other Problems

Lecture 5: Reasoning about Knowledge in the Context of Social Software

Lecture 1: Introduction, Motivation and Basic Models of Knowledge

Lecture 2: Knowledge in Groups and Group Knowledge

Lecture 3: Reasoning about Knowledge and ....

Lecture 4: Logical Omniscience and Other Problems

Lecture 5: Reasoning about Knowledge in the Context of Social Software

Course Website:

staff.science.uva.nl/~epacuit/formep_esslli.html

Reading Material: The course reader and references therein.

- Introduction and Motivation

- Epistemic Logic

What do we want?

A simple mathematical model that *faithfully represents* (the agents information in) social interactive situations.

What do we want?

A simple mathematical model that *faithfully represents* (the agents information in) social interactive situations.

J. Hintikka. *Knowledge and Belief*. 1962, recently republished.

See references in the notes!

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

$$\varphi \; := \; p \mid \neg\varphi \mid \varphi \wedge \psi \mid K\varphi$$

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

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

- $p \in \mathsf{At}$ is an atomic fact.
  - "It is raining"
  - "The talk is at 2PM"
  - "The card on the table is a 7 of Hearts"

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

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

- $p \in \mathsf{At}$ is an atomic fact.
- The usual propositional language ($\mathcal{L}_0$)

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

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

- $p \in \mathsf{At}$ is an atomic fact.
- The usual propositional language ($\mathcal{L}_0$)
- $K\varphi$ is intended to mean "The agent knows that φ is true".

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

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

- $p \in \mathsf{At}$ is an atomic fact.
- The usual propositional language ($\mathcal{L}_0$)
- $K\varphi$ is intended to mean "The agent knows that φ is true".
- The usual definitions for $\rightarrow, \vee, \leftrightarrow$ apply
- Define $L\varphi$ as $\neg K \neg\varphi$

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

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

$K(p \to q)$: "Ann knows that p implies q"

$Kp \vee \neg Kp$:

$Kp \vee K\neg p$:

$L\varphi$:

$KL\varphi$:

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

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

$K(p \rightarrow q)$: "Ann knows that p implies q"

$Kp \vee \neg Kp$: "either Ann does or does not know p"

$Kp \vee K\neg p$: "Ann knows whether p is true"

$L\varphi$:

$KL\varphi$:

Single-Agent Epistemic Logic: The Language

φ is a formula of Epistemic Logic ($\mathcal{L}$) if it is of the form

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

$K(p \rightarrow q)$: "Ann knows that p implies q"

$Kp \vee \neg Kp$: "either Ann does or does not know p"

$Kp \vee K\neg p$: "Ann knows whether p is true"

$L\varphi$: "φ is an epistemic possibility"

$KL\varphi$: "Ann knows that she thinks φ is possible"

Single-Agent Epistemic Logic: Kripke Models

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

Single-Agent Epistemic Logic: Kripke Models

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

- $W \neq \emptyset$ is the set of all relevant situations (states of affairs, possible worlds)

Single-Agent Epistemic Logic: Kripke Models

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

- $W \neq \emptyset$ is the set of all relevant situations (states of affairs, possible worlds)
- $R \subseteq W \times W$ represents the information of the agent: wRv provided "w and v are epistemically indistinguishable"

Single-Agent Epistemic Logic: Kripke Models

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

- $W \neq \emptyset$ is the set of all relevant situations (states of affairs, possible worlds)
- $R \subseteq W \times W$ represents the information of the agent: wRv provided "w and v are epistemically indistinguishable"
- $V : \mathrm{At} \rightarrow \wp(W)$ is a valuation function assigning propositional variables to worlds

Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

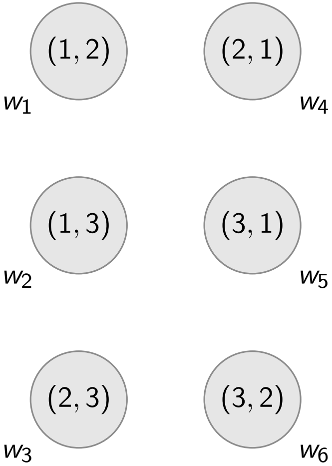
What are the relevant states?

Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

What are the relevant states?

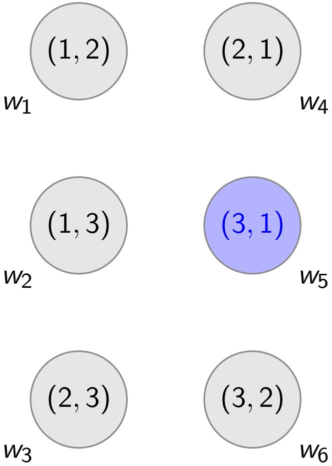


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

Ann receives card 3 and card 1 is put on the table

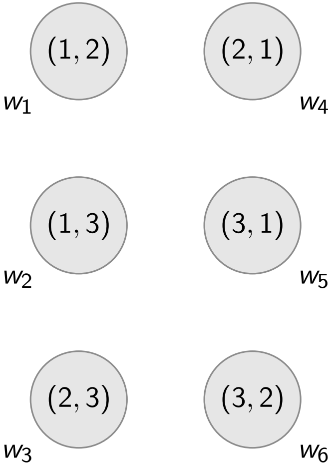


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

What information does Ann have?

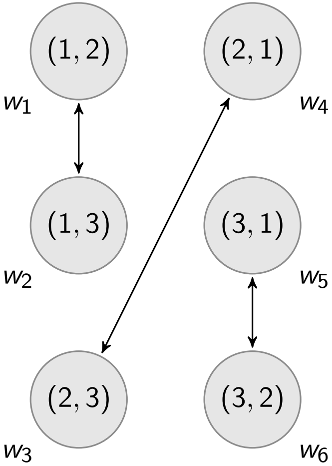


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

What information does Ann have?

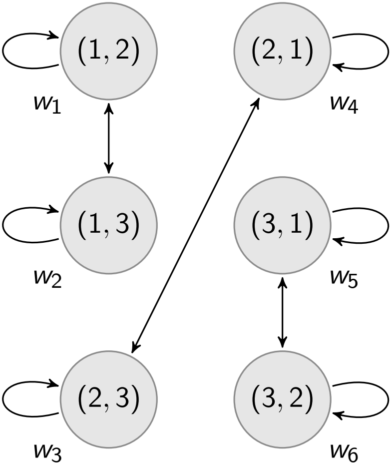


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

What information does Ann have?



Example

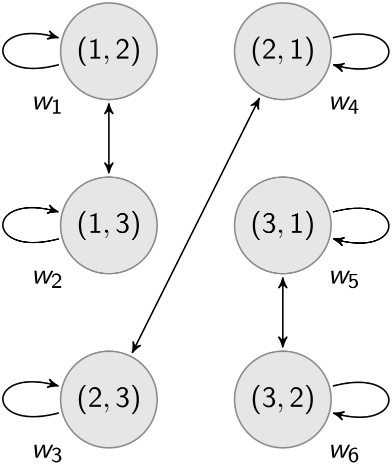
Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

Suppose H_i is intended to mean "Ann has card i"

T_i is intended to mean "card i is on the table"

Eg., $V(H_1) = \{w_1, w_2\}$



Example

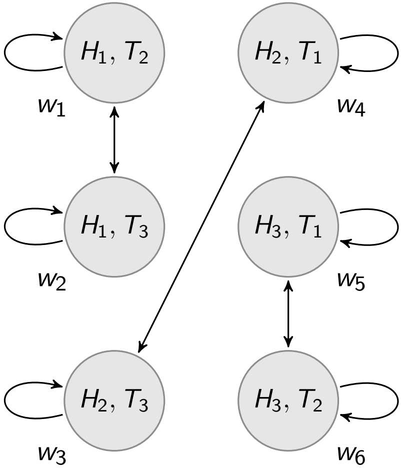
Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

Suppose H_i is intended to mean "Ann has card i"

T_i is intended to mean "card i is on the table"

Eg., $V(H_1) = \{w_1, w_2\}$



Single Agent Epistemic Logic: Truth in a Model

Given $\varphi \in \mathcal{L}$, a Kripke model $\mathcal{M} = \langle W, R, V \rangle$ and $w \in W$

$\mathcal{M}, w \models \varphi$ means "in $\mathcal{M}$, if the actual state is w, then φ is true"

Single Agent Epistemic Logic: Truth in a Model

Given $\varphi \in \mathcal{L}$, a Kripke model $\mathcal{M} = \langle W, R, V \rangle$ and $w \in W$

$\mathcal{M}, w \models \varphi$ is defined as follows:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$ (with $p \in \mathsf{At}$)
- $\mathcal{M}, w \models \neg\varphi$ if $\mathcal{M}, w \not\models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ if $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- $\mathcal{M}, w \models K\varphi$ if for each $v \in W$, if wRv, then $\mathcal{M}, v \models \varphi$

Single Agent Epistemic Logic: Truth in a Model

Given $\varphi \in \mathcal{L}$, a Kripke model $\mathcal{M} = \langle W, R, V \rangle$ and $w \in W$

$\mathcal{M}, w \models \varphi$ is defined as follows:

- ✓ $\mathcal{M}, w \models p$ iff $w \in V(p)$ (with $p \in \text{At}$)
- $\mathcal{M}, w \models \neg\varphi$ if $\mathcal{M}, w \not\models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ if $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- $\mathcal{M}, w \models K\varphi$ if for each $v \in W$, if wRv, then $\mathcal{M}, v \models \varphi$

Single Agent Epistemic Logic: Truth in a Model

Given $\varphi \in \mathcal{L}$, a Kripke model $\mathcal{M} = \langle W, R, V \rangle$ and $w \in W$

$\mathcal{M}, w \models \varphi$ is defined as follows:

- ✓ $\mathcal{M}, w \models p$ iff $w \in V(p)$ (with $p \in \mathrm{At}$)
- ✓ $\mathcal{M}, w \models \neg\varphi$ if $\mathcal{M}, w \not\models \varphi$
- ✓ $\mathcal{M}, w \models \varphi \wedge \psi$ if $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- $\mathcal{M}, w \models K\varphi$ if for each $v \in W$, if wRv, then $\mathcal{M}, v \models \varphi$

Single Agent Epistemic Logic: Truth in a Model

Given $\varphi \in \mathcal{L}$, a Kripke model $\mathcal{M} = \langle W, R, V \rangle$ and $w \in W$

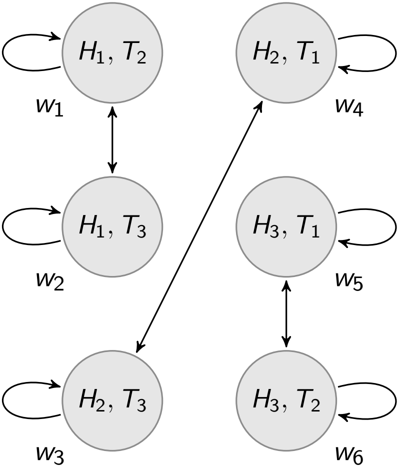
$\mathcal{M}, w \models \varphi$ is defined as follows:

- ✓ $\mathcal{M}, w \models p$ iff $w \in V(p)$ (with $p \in \text{At}$)
- ✓ $\mathcal{M}, w \models \neg\varphi$ if $\mathcal{M}, w \not\models \varphi$
- ✓ $\mathcal{M}, w \models \varphi \wedge \psi$ if $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- ✓ $\mathcal{M}, w \models K\varphi$ if for each $v \in W$, if wRv, then $\mathcal{M}, v \models \varphi$

Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

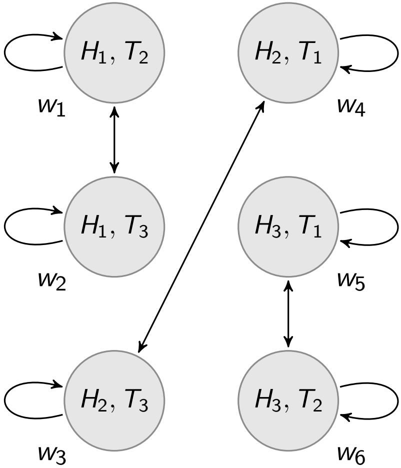


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

Suppose that Ann receives card 1 and card 2 is on the table.

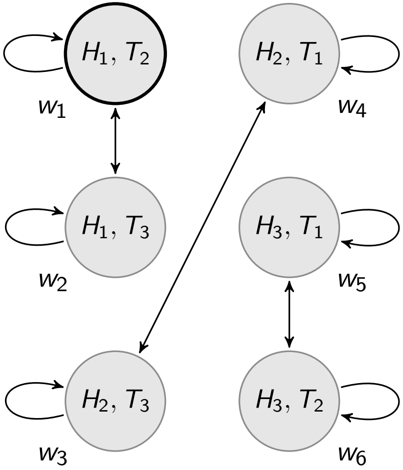


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

Suppose that Ann receives card 1 and card 2 is on the table.

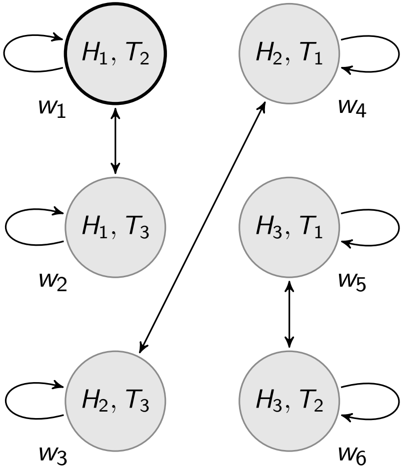


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

$\mathcal{M}, w_1 \models KH_1$

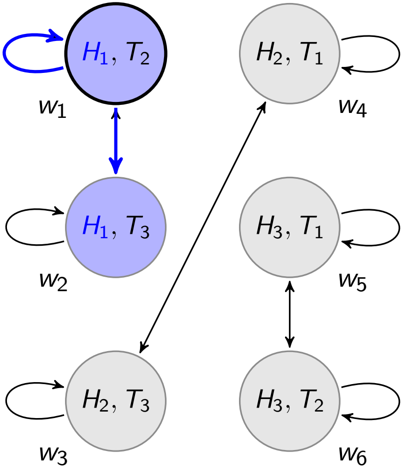


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

$\mathcal{M}, w_1 \models K H_1$



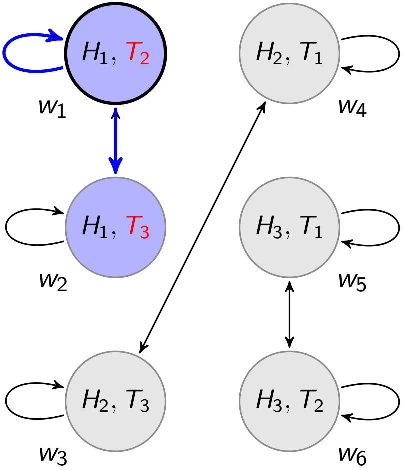
Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

$\mathcal{M}, w_1 \models K H_1$

$\mathcal{M}, w_1 \models K \neg T_1$

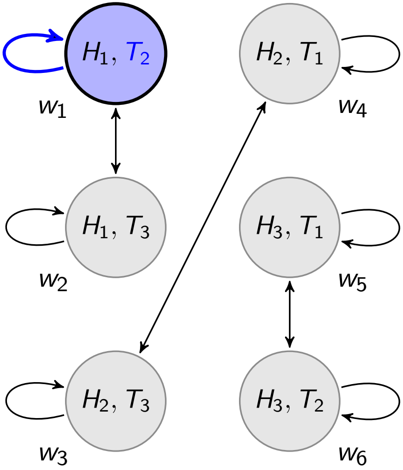


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

$\mathcal{M}, w_1 \models LT_2$



Some Questions

Should we make additional assumptions about R (i.e., reflexive, transitive, etc.)

What idealizations have we made?

Some Notation

A Kripke Frame is a tuple $\langle W, R\rangle$ where $R \subseteq W \times W$.

φ is valid in a Kripke model $\mathcal{M}$ if $\mathcal{M}, w \models \varphi$ for all states w (we write $\mathcal{M} \models \varphi$).

φ is valid on a Kripke frame $\mathcal{F}$ if $\mathcal{M} \models \varphi$ for all models $\mathcal{M}$ based on $\mathcal{F}$.

Logical Omniscience

Fact: φ is valid then $K\varphi$ is valid

Logical Omniscience

Fact: $K\varphi \wedge K\psi \rightarrow K(\varphi \wedge \psi)$ is valid on all Kripke frames

Logical Omniscience

Fact: If $\varphi \rightarrow \psi$ is valid then $K\varphi \rightarrow K\psi$ is valid

Logical Omniscience

Fact: $K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$ is valid on all Kripke frames.

Logical Omniscience

Fact: $\varphi \leftrightarrow \psi$ is valid then $K\varphi \leftrightarrow K\psi$ is valid

Correspondence

Definition

A model formula φ corresponds to a property P (of a relation in a Kripke frame) provided

$$\mathcal{F} \models \varphi \text{ iff } \mathcal{F} \text{ has } P$$

Modal Formula	Corresponding Property
$K\varphi \rightarrow \varphi$	Reflexive
$K\varphi \rightarrow KK\varphi$	Transitive
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean
$\varphi \rightarrow KL\varphi$	Symmetric
$\neg K\bot$	Serial

Correspondence

Definition

A model formula φ corresponds to a property P (of a relation in a Kripke frame) provided

$$\mathcal{F} \models \varphi \text{ iff } \mathcal{F} \text{ has } P$$

Modal Formula	Corresponding Property
$K\varphi \to \varphi$	Reflexive
$K\varphi \to KK\varphi$	Transitive
$\neg K\varphi \to K\neg K\varphi$	Euclidean
$\varphi \to KL\varphi$	Symmetric
$\neg K\bot$	Serial

Correspondence

Definition

A model formula φ corresponds to a property P (of a relation in a Kripke frame) provided

$$\mathcal{F} \models \varphi \text{ iff } \mathcal{F} \text{ has } P$$

Modal Formula	Corresponding Property
$K\varphi \to \varphi$	Reflexive
$K\varphi \to KK\varphi$	Transitive
$\neg K\varphi \to K\neg K\varphi$	Euclidean
$\varphi \to KL\varphi$	Symmetric
$\neg K\bot$	Serial

Correspondence

Definition

A model formula φ corresponds to a property P (of a relation in a Kripke frame) provided

$$\mathcal{F} \models \varphi \text{ iff } \mathcal{F} \text{ has } P$$

Modal Formula	Corresponding Property
$K\varphi \to \varphi$	Reflexive
$K\varphi \to KK\varphi$	Transitive
$\neg K\varphi \to K\neg K\varphi$	Euclidean
$\varphi \to KL\varphi$	Symmetric
$\neg K\bot$	Serial

Correspondence

Definition

A model formula φ corresponds to a property P (of a relation in a Kripke frame) provided

$$\mathcal{F} \models \varphi \text{ iff } \mathcal{F} \text{ has } P$$

Modal Formula	Corresponding Property
$K\varphi \to \varphi$	Reflexive
$K\varphi \to KK\varphi$	Transitive
$\neg K\varphi \to K\neg K\varphi$	Euclidean
$\varphi \to KL\varphi$	Symmetric
$\neg K\bot$	Serial

Correspondence

Definition

A model formula φ corresponds to a property P (of a relation in a Kripke frame) provided

$$\mathcal{F} \models \varphi \text{ iff } \mathcal{F} \text{ has } P$$

Modal Formula	Corresponding Property
$K\varphi \rightarrow \varphi$	Reflexive
$K\varphi \rightarrow KK\varphi$	Transitive
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean
$\varphi \rightarrow KL\varphi$	Symmetric
$\neg K\bot$	Serial

Correspondence

Definition

A model formula φ corresponds to a property P (of a relation in a Kripke frame) provided

$$\mathcal{F} \models \varphi \text{ iff } \mathcal{F} \text{ has } P$$

Modal Formula	Corresponding Property
$K\varphi \rightarrow \varphi$	Reflexive
$K\varphi \rightarrow KK\varphi$	Transitive
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean
$\varphi \rightarrow KL\varphi$	Symmetric
$\neg K\bot$	Serial

Modal Formula	Property	Philosophical Assumption
$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$	—	Logical Omniscience
$K\varphi \rightarrow \varphi$	Reflexive	Truth
$K\varphi \rightarrow KK\varphi$	Transitive	Positive Introspection
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean	Negative Introspection
$\neg K\bot$	Serial	Consistency

Modal Formula	Property	Philosophical Assumption
$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$	—	Logical Omniscience
$K\varphi \rightarrow \varphi$	Reflexive	Truth
$K\varphi \rightarrow KK\varphi$	Transitive	Positive Introspection
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean	Negative Introspection
$\neg K\bot$	Serial	Consistency

Modal Formula	Property	Philosophical Assumption
$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$	—	Logical Omniscience
$K\varphi \rightarrow \varphi$	Reflexive	Truth
$K\varphi \rightarrow KK\varphi$	Transitive	Positive Introspection
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean	Negative Introspection
$\neg K\bot$	Serial	Consistency

Modal Formula	Property	Philosophical Assumption
$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$	—	Logical Omniscience
$K\varphi \rightarrow \varphi$	Reflexive	Truth
$K\varphi \rightarrow KK\varphi$	Transitive	Positive Introspection
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean	Negative Introspection
$\neg K\bot$	Serial	Consistency

Modal Formula	Property	Philosophical Assumption
$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$	—	Logical Omniscience
$K\varphi \rightarrow \varphi$	Reflexive	Truth
$K\varphi \rightarrow KK\varphi$	Transitive	Positive Introspection
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean	Negative Introspection
$\neg K\bot$	Serial	Consistency

Modal Formula	Property	Philosophical Assumption
$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$	—	Logical Omniscience
$K\varphi \rightarrow \varphi$	Reflexive	Truth
$K\varphi \rightarrow KK\varphi$	Transitive	Positive Introspection
$\neg K\varphi \rightarrow K\neg K\varphi$	Euclidean	Negative Introspection
$\neg K\bot$	Serial	Consistency

The Logic **S5**

The logic **S5** contains the following axioms and rules:

Pc	Axiomatization of Propositional Calculus
K	$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$
T	$K\varphi \rightarrow \varphi$
4	$K\varphi \rightarrow KK\varphi$
5	$\neg K\varphi \rightarrow K\neg K\varphi$
MP	$\dfrac{\varphi \qquad \varphi \rightarrow \psi}{\psi}$
Nec	$\dfrac{\varphi}{K\psi}$

Theorem

S5 *is sound and strongly complete with respect to the class of Kripke frames with equivalence relations.*

The Logic **S5**

The logic **S5** contains the following axioms and rules:

Pc	Axiomatization of Propositional Calculus
K	$K(\varphi \rightarrow \psi) \rightarrow (K\varphi \rightarrow K\psi)$
T	$K\varphi \rightarrow \varphi$
4	$K\varphi \rightarrow KK\varphi$
5	$\neg K\varphi \rightarrow K\neg K\varphi$
MP	$\dfrac{\varphi \qquad \varphi \rightarrow \psi}{\psi}$
Nec	$\dfrac{\varphi}{K\psi}$

Theorem

S5 *is sound and strongly complete with respect to the class of Kripke frames with equivalence relations.*

Multi-agent Epistemic Logic

The Language: $\varphi \ := \ p \mid \neg\varphi \mid \varphi \wedge \psi \mid K\varphi$

Kripke Models: $\mathcal{M} = \langle W, R, V \rangle$ and $w \in W$

Truth: $\mathcal{M}, w \models \varphi$ is defined as follows:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$ (with $p \in \text{At}$)
- $\mathcal{M}, w \models \neg\varphi$ if $\mathcal{M}, w \not\models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ if $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- $\mathcal{M}, w \models K\varphi$ if for each $v \in W$, if wRv, then $\mathcal{M}, v \models \varphi$

Multi-agent Epistemic Logic

The Language: $\varphi \; := \; p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_i\varphi$ with $i \in \mathcal{A}$

Kripke Models: $\mathcal{M} = \langle W, \{R_i\}_{i \in \mathcal{A}}, V \rangle$ and $w \in W$

Truth: $\mathcal{M}, w \models \varphi$ is defined as follows:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$ (with $p \in \text{At}$)
- $\mathcal{M}, w \models \neg\varphi$ if $\mathcal{M}, w \not\models \varphi$
- $\mathcal{M}, w \models \varphi \wedge \psi$ if $\mathcal{M}, w \models \varphi$ and $\mathcal{M}, w \models \psi$
- $\mathcal{M}, w \models K_i\varphi$ if for each $v \in W$, if wR_iv, then $\mathcal{M}, v \models \varphi$

Multi-agent Epistemic Logic

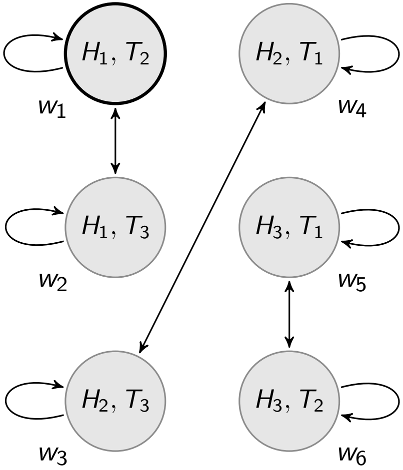
- $K_A K_B \varphi$: "Ann knows that Bob knows φ"
- $K_A(K_B \varphi \vee K_B \neg\varphi)$: "Ann knows that Bob knows whether φ
- $\neg K_B K_A K_B(\varphi)$: "Bob does not know that Ann knows that Bob knows that φ"

Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, one of the cards is placed face down on the table and the third card is put back in the deck.

Suppose that Ann receives card 1 and card 2 is on the table.

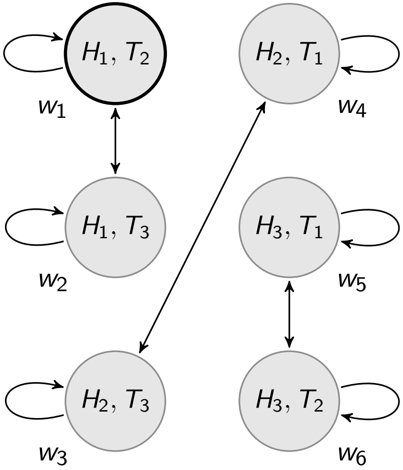


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, Bob is given one of the cards and the third card is put back in the deck.

Suppose that Ann receives card 1 and Bob receives card 2.

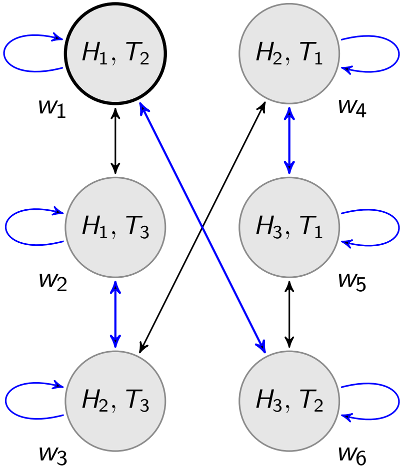


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, Bob is given one of the cards and the third card is put back in the deck.

Suppose that Ann receives card 1 and Bob receives card 2.

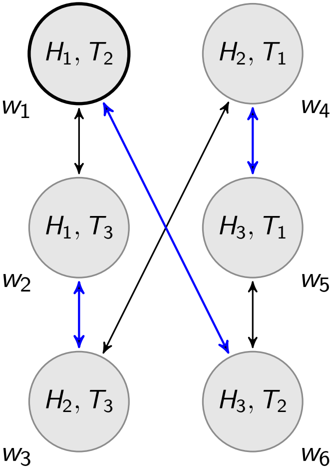


Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, Bob is given one of the cards and the third card is put back in the deck.

Suppose that Ann receives card 1 and Bob receives card 2.



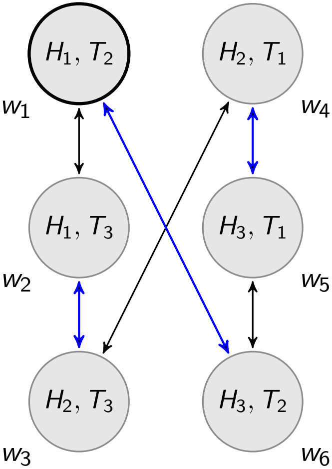
Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, Bob is given one of the cards and the third card is put back in the deck.

Suppose that Ann receives card 1 and Bob receives card 2.

$\mathcal{M}, w \models K_B(K_A H_1 \vee K_A \neg H_1)$



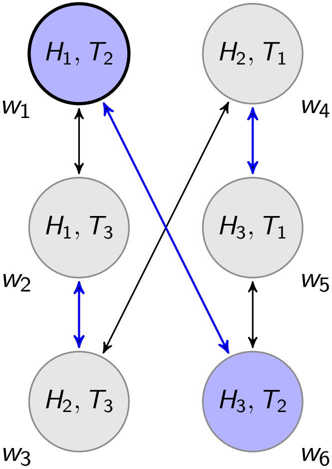
Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, Bob is given one of the cards and the third card is put back in the deck.

Suppose that Ann receives card 1 and Bob receives card 2.

$\mathcal{M}, w \models K_B(K_A H_1 \vee K_A \neg H_1)$



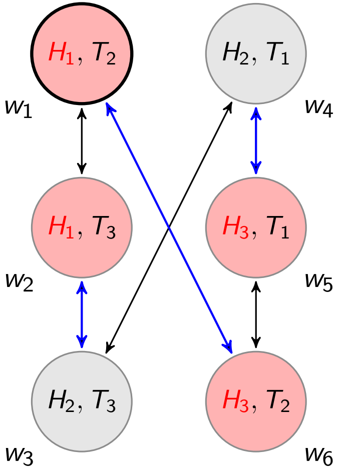
Example

Suppose there are three cards: 1, 2 and 3.

Ann is dealt one of the cards, Bob is given one of the cards and the third card is put back in the deck.

Suppose that Ann receives card 1 and Bob receives card 2.

$\mathcal{M}, w \models K_B(K_A H_1 \vee K_A \neg H_1)$



Example

Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

There is a very simple procedure to solve Ann's problem: *have a (trusted) friend tell Bob the time and subject of her talk.*

Is this procedure correct?

Example

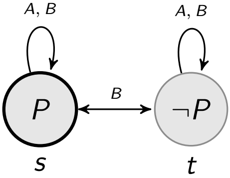
Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

There is a very simple procedure to solve Ann's problem: *have a (trusted) friend tell Bob the time and subject of her talk.*

Is this procedure correct? Yes, if

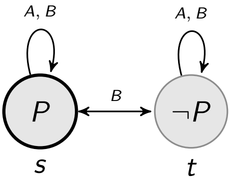
1. Ann knows about the talk.
2. Bob knows about the talk.
3. Ann knows that Bob knows about the talk.
4. Bob *does not* know that Ann knows that he knows about the talk.
5. *And nothing else.*

Example



P means "The talk is at 2PM".

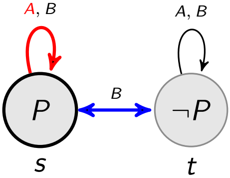
Example



P means "The talk is at 2PM".

$\mathcal{M}, s \models K_A P \wedge \neg K_B P$

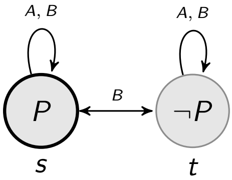
Example

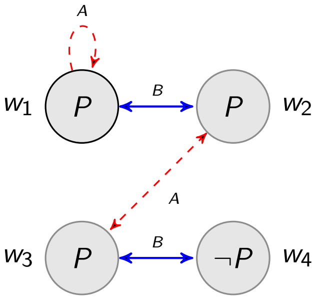


P means "The talk is at 2PM".

$\mathcal{M}, s \models K_A P \wedge \neg K_B P$

Example





Thank you!