An Invitation to Modal Logic: Lecture 5

Philosophy 150

Eric Pacuit

Stanford University
ai.stanford.edu/~epacuit

December 5, 2007

Plan

✓ Motivating Examples

✓ Formalizing the muddy children puzzle, Introduction to Modal Logic

✓ More about truth of modal formulas

✓ Summary so far.
Digression: A small experiment.

12/5: Focus on Epistemic Logic, Dynamics in Logic

12/7: Dynamics in Logic II

Single-Agent Epistemic Logic

Typically, we write KP instead of $\Box P$ when the intended interpretation is "P is known".

Single-Agent Epistemic Logic

Typically, we write KP instead of $\Box P$ when the intended interpretation is "P is known".

$K(P \to Q)$: "Ann knows that P implies Q"

Single-Agent Epistemic Logic

Typically, we write KP instead of $\Box P$ when the intended interpretation is "P is known".

$K(P \to Q)$: "Ann knows that P implies Q"

$KP \vee \neg KP$: "either Ann does or does not know P"

Single-Agent Epistemic Logic

Typically, we write KP instead of $\Box P$ when the intended interpretation is "P is known".

$K(P \to Q)$: "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"

Single-Agent Epistemic Logic

Typically, we write KP instead of $\Box P$ when the intended interpretation is "P is known".

$K(P \to Q)$: "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"

LP: "P is an epistemic possibility"

Single-Agent Epistemic Logic

Typically, we write KP instead of $\Box P$ when the intended interpretation is "P is known".

$K(P \rightarrow Q)$: "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"

LP: "P is an epistemic possibility"

KLP: "Ann knows that she thinks P is possible"

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.

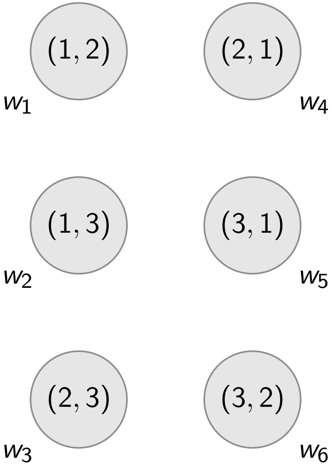
What are the relevant states?

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?

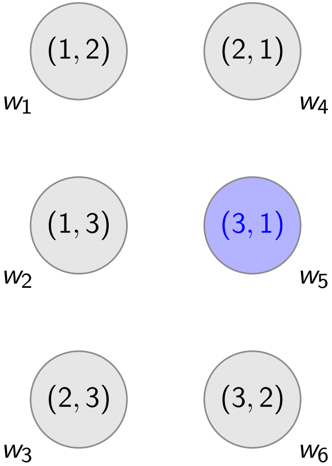


Example

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

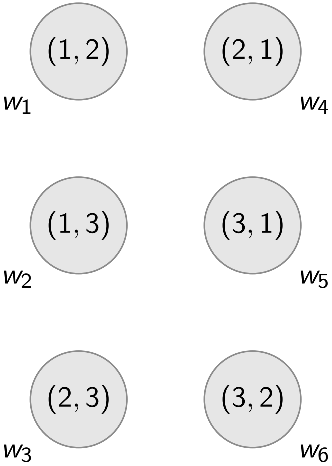


Example

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?

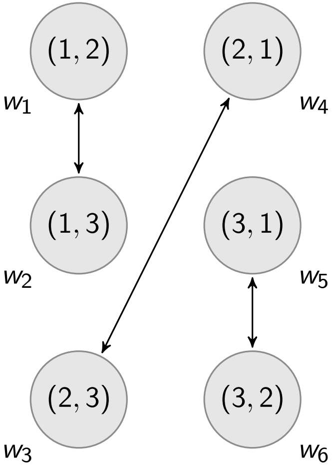


Example

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?

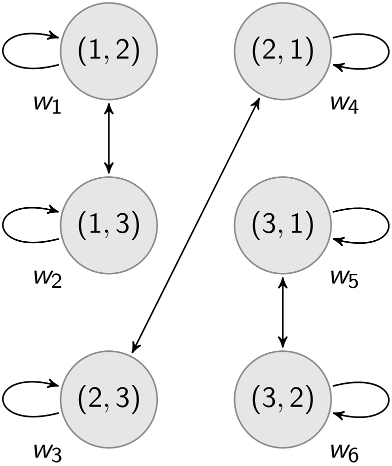


Example

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

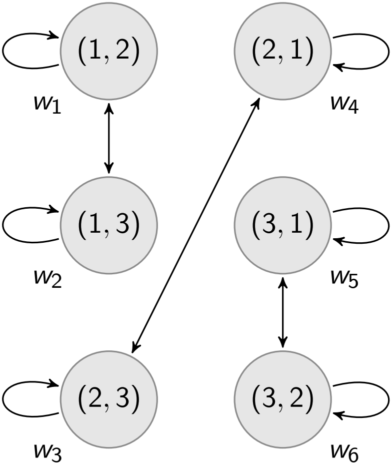
Suppose there are three cards: 1, 2 and 3.

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

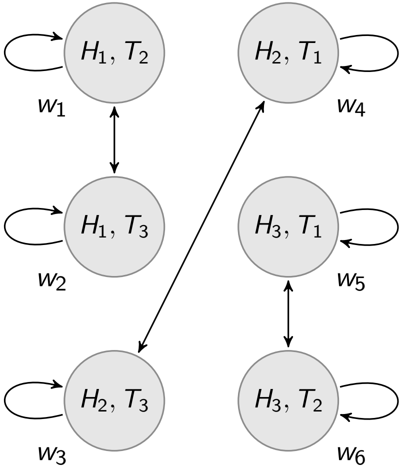
Suppose there are three cards: 1, 2 and 3.

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"

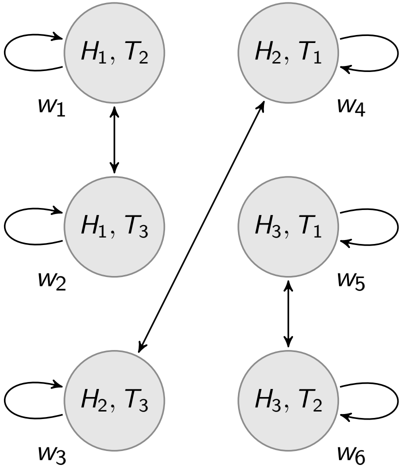
Eg., $V(H_1) = \{w_1, w_2\}$



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.

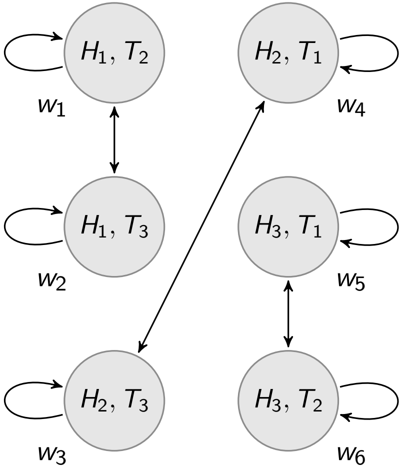


Example

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.

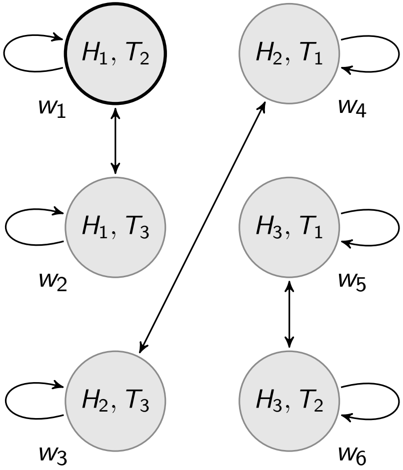


Example

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.

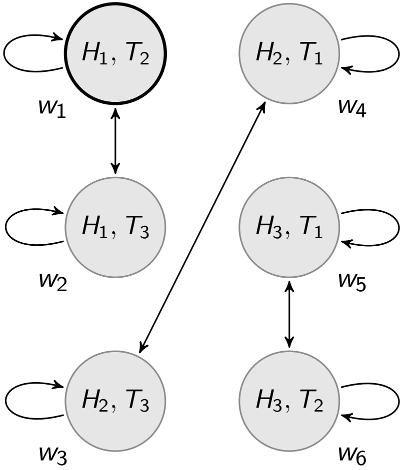


Example

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.

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

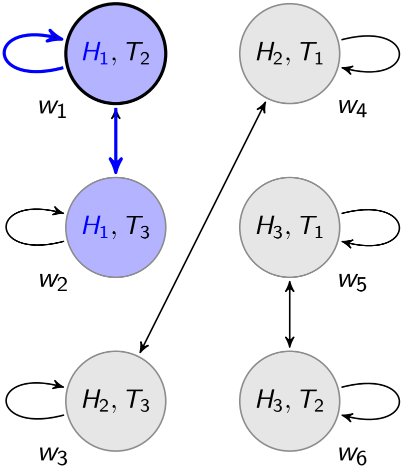


Example

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.

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



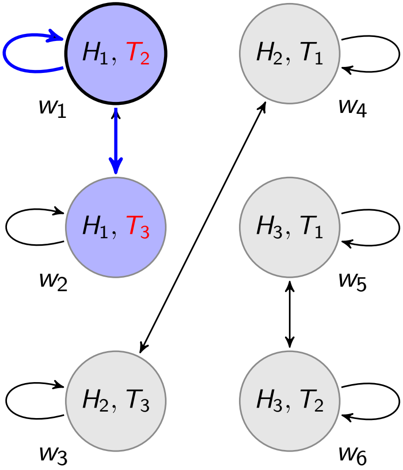
Example

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.

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

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

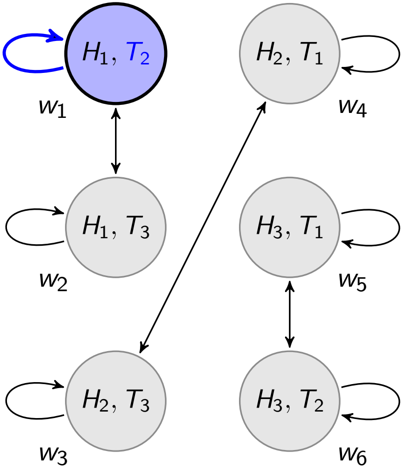


Example

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.

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

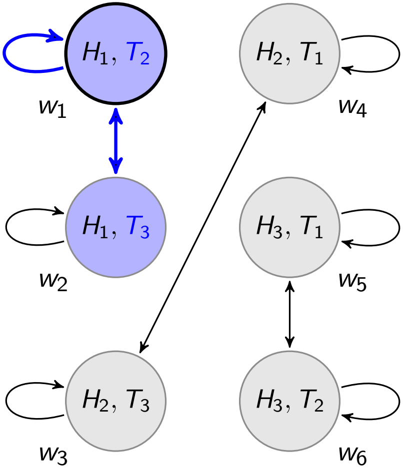


Example

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.

$\mathcal{M}, w_1 \models K(T_2 \vee T_3)$



Some Questions

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

Some Questions

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

For two states w and v, say wRv provided "w and v are indistinguishable according to Ann's information". What properties should R satisfy?

Some Questions

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

For two states w and v, say wRv provided "w and v are indistinguishable according to Ann's information". What properties should R satisfy?

What idealizations have we made?

Modal Formula	Property	Philosophical Assumption

Modal Formula	Property	Philosophical Assumption
$K(P \rightarrow Q) \rightarrow (KP \rightarrow KQ)$	—	Logical Omniscience

Modal Formula	Property	Philosophical Assumption
$K(P \to Q) \to (KP \to KQ)$	—	Logical Omniscience
$KP \to P$	Reflexive	Truth

Modal Formula	Property	Philosophical Assumption
$K(P \rightarrow Q) \rightarrow (KP \rightarrow KQ)$	—	Logical Omniscience
$KP \rightarrow P$	Reflexive	Truth
$KP \rightarrow KKP$	Transitive	Positive Introspection

Modal Formula	Property	Philosophical Assumption
$K(P \to Q) \to (KP \to KQ)$	—	Logical Omniscience
$KP \to P$	Reflexive	Truth
$KP \to KKP$	Transitive	Positive Introspection
$\neg KP \to K\neg KP$	Euclidean	Negative Introspection

Modal Formula	Property	Philosophical Assumption
$K(P \to Q) \to (KP \to KQ)$	—	Logical Omniscience
$KP \to P$	Reflexive	Truth
$KP \to KKP$	Transitive	Positive Introspection
$\neg KP \to K\neg KP$	Euclidean	Negative Introspection
$\neg K\bot$	Serial	Consistency

Multiagent Epistemic Logic

Many of the examples we are interested in involve more than one agent!

Multiagent Epistemic Logic

Many of the examples we are interested in involve more than one agent!

$K_A P$ means "Ann knows P"

$K_B P$ means "Bob knows P"

Multiagent Epistemic Logic

Many of the examples we are interested in involve more than one agent!

$K_A P$ means "Ann knows P"

$K_B P$ means "Bob knows P"

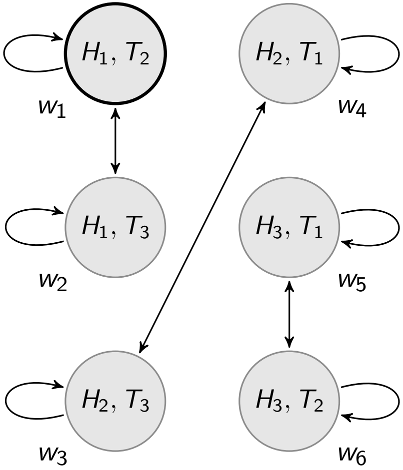
- $K_A K_B P$: "Ann knows that Bob knows P"
- $K_A(K_B P \vee K_B \neg P)$: "Ann knows that Bob knows whether P
- $\neg K_B K_A K_B(P)$: "Bob does not know that Ann knows that Bob knows that P"

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.

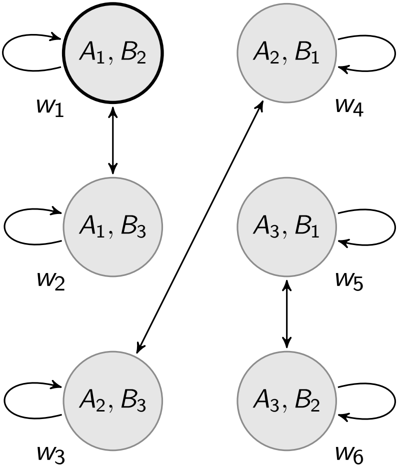


Example

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.

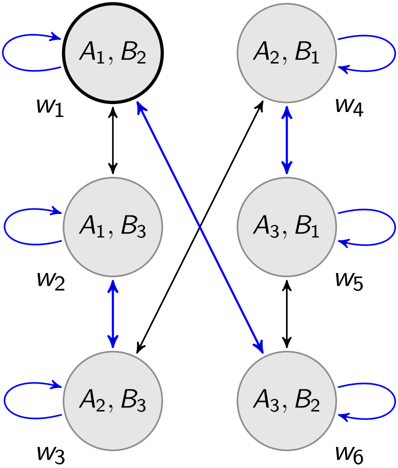


Example

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.

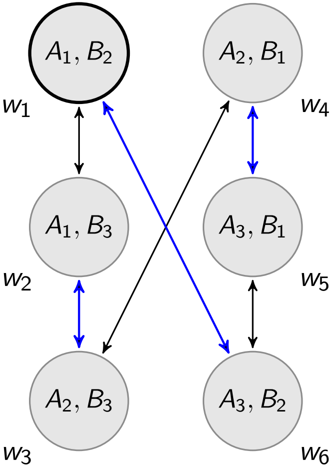


Example

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.



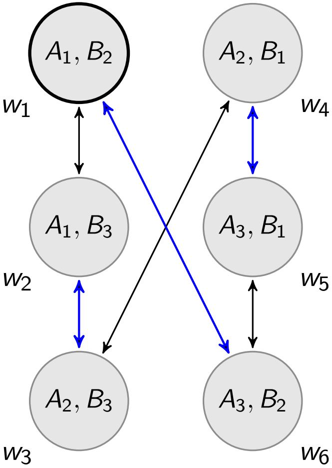
Example

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.

$w_1 \models K_B(K_A A_1 \vee K_A \neg A_1)$



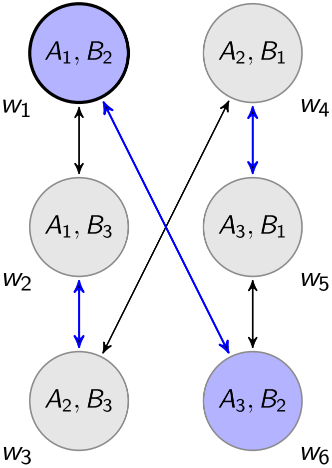
Example

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.

$w_1 \models K_B(K_A A_1 \vee K_A \neg A_1)$



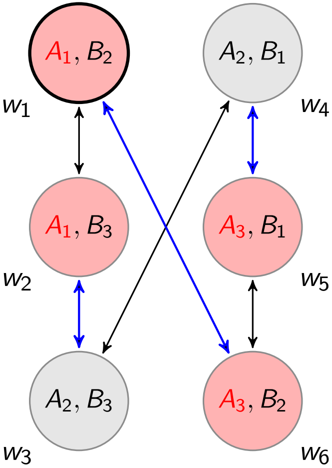
Example

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.

$w_1 \models K_B(K_A A_1 \vee K_A \neg A_1)$



Two issues

Suppose we want to be completely formal in our solution to the muddy children puzzle, what is Epistemic Logic missing?

1. Group knowledge (all the children know there is at least one muddy child, they all know this fact, they all know that they know this fact, etc.).
2. Public announcements (various statements are publicly announced in the course of the puzzle).

Group Knowledge

$K_A P$: "Ann knows that P"

Group Knowledge

$K_A P$: "Ann knows that P"

$K_B P$: "Bob knows that P"

Group Knowledge

$K_A P$: "Ann knows that P"

$K_B P$: "Bob knows that P"

$K_A K_B P$: "Ann knows that Bob knows that P"

Group Knowledge

$K_A P$: "Ann knows that P"

$K_B P$: "Bob knows that P"

$K_A K_B P$: "Ann knows that Bob knows that P"

$K_A P \wedge K_B P$: "Every one knows P".

Group Knowledge

$K_A P$: "Ann knows that P"

$K_B P$: "Bob knows that P"

$K_A K_B P$: "Ann knows that Bob knows that P"

$K_A P \wedge K_B P$: "Every one knows P". let $EP := K_A P \wedge K_B P$

Group Knowledge

$K_A P$: "Ann knows that P"

$K_B P$: "Bob knows that P"

$K_A K_B P$: "Ann knows that Bob knows that P"

$K_A P \wedge K_B P$: "Every one knows P". let $EP := K_A P \wedge K_B P$

$K_A EP$: "Ann knows that everyone knows that P".

Group Knowledge

$K_A P$: "Ann knows that P"

$K_B P$: "Bob knows that P"

$K_A K_B P$: "Ann knows that Bob knows that P"

$K_A P \wedge K_B P$: "Every one knows P". let $EP := K_A P \wedge K_B P$

$K_A EP$: "Ann knows that everyone knows that P".

EEP: "Everyone knows that everyone knows that P".

Group Knowledge

$K_A P$: "Ann knows that P"

$K_B P$: "Bob knows that P"

$K_A K_B P$: "Ann knows that Bob knows that P"

$K_A P \wedge K_B P$: "Every one knows P". let $EP := K_A P \wedge K_B P$

$K_A EP$: "Ann knows that everyone knows that P".

EEP: "Everyone knows that everyone knows that P".

$EEEP$: "Everyone knows that everyone knows that everyone knows that P."

Common Knowledge

CP: "It is common knowledge that P"

Common Knowledge

CP: "It is common knowledge that P" — "Everyone knows that everyone knows that everyone knows that $\cdots$ P".

Common Knowledge

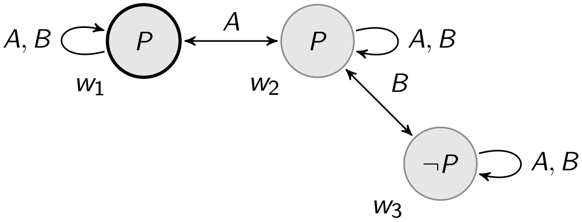
CP: "It is common knowledge that P" — "Everyone knows that everyone knows that everyone knows that $\cdots$ P".

Is *common knowledge* different from *everyone knows*?

Common Knowledge

CP: "It is common knowledge that P" — "Everyone knows that everyone knows that everyone knows that $\cdots$ P".

Is *common knowledge* different from *everyone knows*?

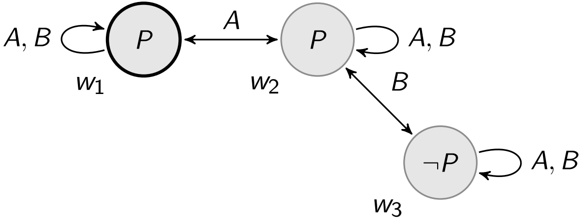


$$w_1 \models EP \wedge \neg CP$$

Common Knowledge

CP: "It is common knowledge that P" — "Everyone knows that everyone knows that everyone knows that $\cdots$ P".

Is *common knowledge* different from *everyone knows*?

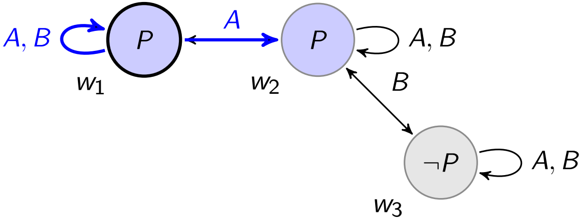


$$w_1 \models EP \wedge \neg CP$$

Common Knowledge

CP: "It is common knowledge that P" — "Everyone knows that everyone knows that everyone knows that $\cdots$ P".

Is *common knowledge* different from *everyone knows*?

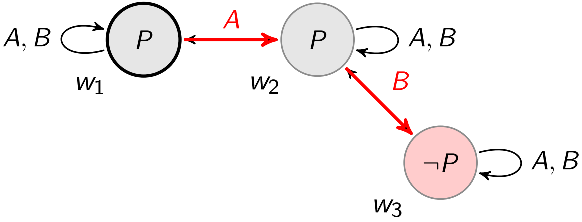


$$w_1 \models EP \wedge \neg CP$$

Common Knowledge

CP: "It is common knowledge that P" — "Everyone knows that everyone knows that everyone knows that $\cdots$ P".

Is *common knowledge* different from *everyone knows*?



$$w_1 \models EP \wedge \neg CP$$

The Generals Puzzle

Imagine two two allied generals, G_1 and G_2, standing on two mountain summits, with their enemy in the valley between them. Both G_1 and G_2 know that if they attack at the same time, then they will defeat the enemy, but if only one attacks then he will certainly loose the battle.

The Generals Puzzle

Imagine two two allied generals, G_1 and G_2, standing on two mountain summits, with their enemy in the valley between them. Both G_1 and G_2 know that if they attack at the same time, then they will defeat the enemy, but if only one attacks then he will certainly loose the battle.

- G_1 sends a message $m=$"Let's attack at 8 AM"; however it is not guaranteed that the message will arrive.

The Generals Puzzle

Imagine two two allied generals, G_1 and G_2, standing on two mountain summits, with their enemy in the valley between them. Both G_1 and G_2 know that if they attack at the same time, then they will defeat the enemy, but if only one attacks then he will certainly loose the battle.

- G_1 sends a message $m=$"Let's attack at 8 AM"; however it is not guaranteed that the message will arrive.
- Suppose G_2 receives message m. Should they attack?

The Generals Puzzle

Imagine two two allied generals, G_1 and G_2, standing on two mountain summits, with their enemy in the valley between them. Both G_1 and G_2 know that if they attack at the same time, then they will defeat the enemy, but if only one attacks then he will certainly loose the battle.

- G_1 sends a message m= "Let's attack at 8 AM"; however it is not guaranteed that the message will arrive.
- Suppose G_2 receives message m. Should they attack?
- No! (G_1 thinks that 'perhaps G_2 did not receive m.'). So, G_2 sends a message m'= "OK, let's attack at 8AM".

The Generals Puzzle

Imagine two two allied generals, G_1 and G_2, standing on two mountain summits, with their enemy in the valley between them. Both G_1 and G_2 know that if they attack at the same time, then they will defeat the enemy, but if only one attacks then he will certainly loose the battle.

- G_1 sends a message m= "Let's attack at 8 AM"; however it is not guaranteed that the message will arrive.
- Suppose G_2 receives message m. Should they attack?
- No! (G_1 thinks that 'perhaps G_2 did not receive m.'). So, G_2 sends a message m'= "OK, let's attack at 8AM".
- Suppose G_1 receives message m'. Should they attack?

The Generals Puzzle

Imagine two two allied generals, G_1 and G_2, standing on two mountain summits, with their enemy in the valley between them. Both G_1 and G_2 know that if they attack at the same time, then they will defeat the enemy, but if only one attacks then he will certainly loose the battle.

- G_1 sends a message m= "Let's attack at 8 AM"; however it is not guaranteed that the message will arrive.
- Suppose G_2 receives message m. Should they attack?
- No! (G_1 thinks that 'perhaps G_2 did not receive m.'). So, G_2 sends a message m'= "OK, let's attack at 8AM".
- Suppose G_1 receives message m'. Should they attack?
- No! (G_2 thinks that 'perhaps G_1 did not receive m' and G_1 knows this).

The Generals Puzzle

Imagine two two allied generals, G_1 and G_2, standing on two mountain summits, with their enemy in the valley between them. Both G_1 and G_2 know that if they attack at the same time, then they will defeat the enemy, but if only one attacks then he will certainly loose the battle.

- G_1 sends a message m= "Let's attack at 8 AM"; however it is not guaranteed that the message will arrive.
- Suppose G_2 receives message m. Should they attack?
- No! (G_1 thinks that 'perhaps G_2 did not receive m.'). So, G_2 sends a message m'= "OK, let's attack at 8AM".
- Suppose G_1 receives message m'. Should they attack?
- No! (G_2 thinks that 'perhaps G_1 did not receive m' and G_1 knows this).
- So G_1 sends a message m'', ...

Spreading Gossip

Suppose that there are three friends, Ann, Bob and Charles, and Ann learns a interesting piece of news (P). If each of the friends are at home, how many calls are needed to create common knowledge that P?

Public Announcements

$\langle !P \rangle Q$ is intended to mean "after publicly announcing P, Q is true".

Public Announcements

$\langle !P \rangle Q$ is intended to mean "after publicly announcing P, Q is true".

$\langle !P \rangle K_A P$: "After P is announced, Ann knows that P"

Public Announcements

$\langle !P \rangle Q$ is intended to mean "after publicly announcing P, Q is true".

$\langle !P \rangle K_A P$: "After P is announced, Ann knows that P"

$\langle !(K_A P \vee K_A \neg P) \rangle K_B K_A P$: "After it is announced that Ann knows whether P is true, then Bob knows that Ann knows that P."

Public Announcements

$\langle !P \rangle Q$ is intended to mean "after publicly announcing P, Q is true".

$\langle !P \rangle K_A P$: "After P is announced, Ann knows that P"

$\langle !(K_A P \vee K_A \neg P) \rangle K_B K_A P$: "After it is announced that Ann knows whether P is true, then Bob knows that Ann knows that P."

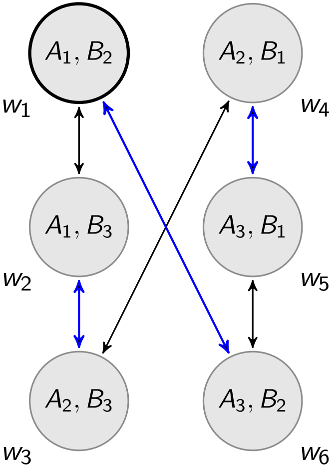
$\langle !(\neg(K_A \vee K_A \neg P)) \rangle K_B P$: "After it is announced that A does not know whether P, then B knows P."

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.



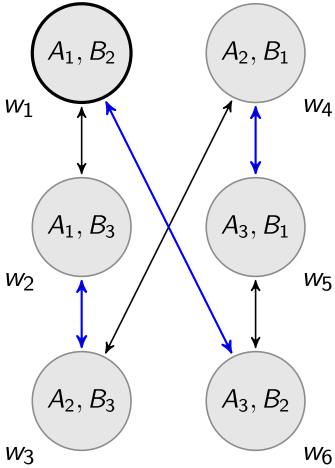
Example

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.

$w_1 \models \langle !(\neg A_3 \wedge \neg B_3) \rangle K_A B_2$



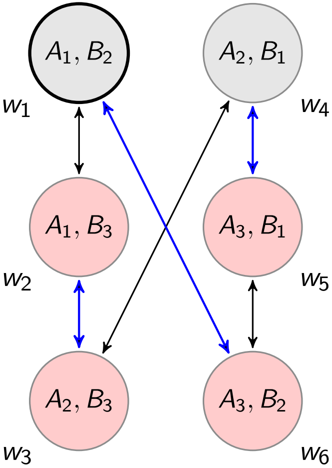
Example

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.

$w_1 \models \langle !(\neg A_3 \wedge \neg B_3) \rangle K_A B_2$



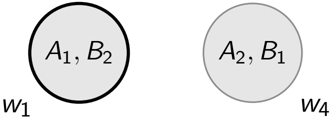
Example

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.

$w_1 \models \langle !(\neg A_3 \wedge \neg B_3) \rangle K_A B_2$



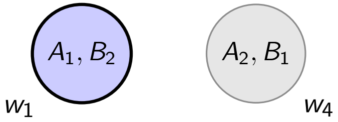
Example

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.

$w_1 \models \langle !(\neg A_3 \wedge \neg B_3)\rangle K_A B_2$



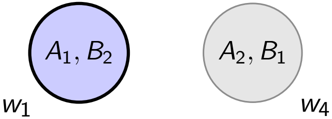
Example

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.

$w_1 \models \langle !(\neg A_3 \wedge \neg B_3) \rangle C(A_1 \wedge B_2)$



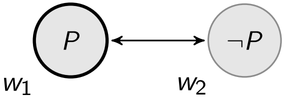
Question: If a true state P is publicly announced, does it become common knowledge? does it become known by everyone?

Is $\langle !Q \rangle KQ$ always true?

Question: If a true state P is publicly announced, does it become common knowledge? does it become known by everyone?

Is $\langle !Q \rangle KQ$ always true?

"You don't know it, but you have a bug on your shoulder."

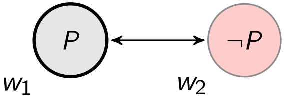


$$w_1 \not\models \langle !(\neg KP \wedge P)\rangle K(\neg KP \wedge P)$$

Question: If a true state P is publicly announced, does it become common knowledge? does it become known by everyone?

Is $\langle !Q \rangle KQ$ always true?

"You don't know it, but you have a bug on your shoulder."

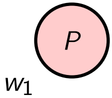


$$w_1 \not\models \langle !(\neg KP \wedge P)\rangle K(\neg KP \wedge P)$$

Question: If a true state P is publicly announced, does it become common knowledge? does it become known by everyone?

Is $\langle !Q \rangle KQ$ always true?

"You don't know it, but you have a bug on your shoulder."



$$w_1 \not\models \langle !(\neg KP \wedge P) \rangle K(\neg KP \wedge P)$$

Something to Think About

1. Suppose Ann and Bob both know that two numbers n and $n+1$ will be chosen and placed on their foreheads. They will be able to see the other player's number, but not their own. Say 3 is written on Ann's forehead and 4 is written on Bob's forehead. Draw a Kripke structure that represents this situation (it is infinite). Is it common knowledge that the numbers are less than 1000? What happens if the agents start (truthfully) announcing "I don't know my number."?

Something to Think About

1. Suppose Ann and Bob both know that two numbers n and $n+1$ will be chosen and placed on their foreheads. They will be able to see the other player's number, but not their own. Say 3 is written on Ann's forehead and 4 is written on Bob's forehead. Draw a Kripke structure that represents this situation (it is infinite). Is it common knowledge that the numbers are less than 1000? What happens if the agents start (truthfully) announcing "I don't know my number."?

2. **Russian Cards Problem**: From a deck of seven cards Ann and Bob each receive three cards and Charles the remaining card. How can Ann and Bob openly inform each other about their cards, without informing Charles who holds which card?

Next lecture: Dynamics in logic.

Questions?

Email: epacuit@stanford.edu
Website: ai.stanford.edu/~epacuit
Office: Gates 258