# Reasoning about Knowledge and Beliefs

Lecture 1

Eric Pacuit

University of Maryland, College Park

pacuit.org
epacuit@umd.edu

September 4, 2013

# Practicalities

## Practicalities

- Course website on ELMS
  - Weekly readings will be posted
  - Slides will be posted
  - Online videos and quizzes will be posted
  - Pay attention to the schedule (midterm, canceled classes, etc.)

# Practicalities

- Course website on ELMS
  - Weekly readings will be posted
  - Slides will be posted
  - Online videos and quizzes will be posted
  - Pay attention to the schedule (midterm, canceled classes, etc.)
- Weekly lectures (Monday, Wedensday) + some online lectures

# Practicalities

- Course website on ELMS
  - Weekly readings will be posted
  - Slides will be posted
  - Online videos and quizzes will be posted
  - Pay attention to the schedule (midterm, canceled classes, etc.)
- Weekly lectures (Monday, Wedensday) + some online lectures
- Class canceled **September 16 & 18**

# Practicalities

- Course website on ELMS
  - Weekly readings will be posted
  - Slides will be posted
  - Online videos and quizzes will be posted
  - Pay attention to the schedule (midterm, canceled classes, etc.)
- Weekly lectures (Monday, Wedensday) + some online lectures
- Class canceled **September 16 & 18**
- Office Hours: Tuesdays 1-2 PM
- Office: Skinner 1103A

# Practicalities: Grading

# Practicalities: Grading

1. Participation & short quizzes (30%)
   - Quizzes are available online (multiple choice, true/false, short answers)
   - It is your job to monitor the due dates and take the quizzes!
   - No make-ups!

2. 2-3 Problem Sets (30%)
   - Short essay questions
   - Short proofs

3. Final Exam (40%)
   - In-class exam given during finals week

# Practicalities: Literature

# Practicalities: Literature

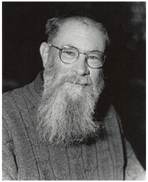
Contemporary research papers published in academic journals and recent books (consult the schedule for details).

# Practicalities: Literature

Contemporary research papers published in academic journals and recent books (consult the schedule for details).

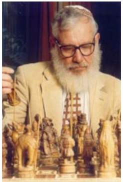

1. W. Holliday, Epistemic Logic and Epistemology, *Handbook of Formal Philosophy*, Springer, forthcoming
2. E. Pacuit, Dynamic Epistemic Logic I: Modeling Knowledge and Belief, *Philosophy Compass*, 2013
3. E. Pacuit, Dynamic Epistemic Logic II: Logics of Information Change, *Philosophy Compass*, 2013
4. R. Sorensen, Epistemic Paradoxes, Stanford Encyclopedia of Philosophy, 2011

# Foundations of Epistemic Logic



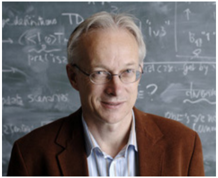
David Lewis

Jakko Hintikka



Robert Aumann

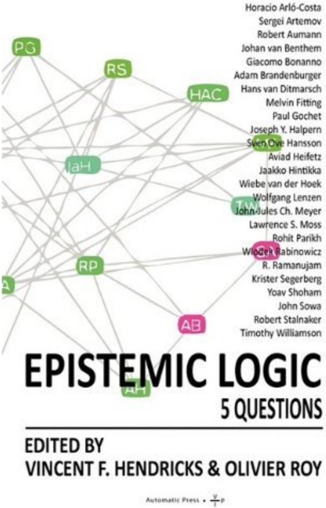
Larry Moss



Johan van Benthem

Alexandru Baltag

# Foundations of Epistemic Logic



# Ten Puzzles and Paradoxes

1. Surprise Exam
2. The Knower
3. Logical Omniscience/Knowledge Closure
4. Lottery Paradox & Preface Paradox
5. Margin of Error Paradox
6. Fitch's Paradox
7. Aumann's Agreeing to Disagree Theorem
8. Brandenburger-Keisler Paradox
9. Absent-Minded Driver
10. Common Knowledge of Rationality and Backwards Induction

## A Word about the Word ‘Paradox’

What is a paradox? Roughly: an apparently good argument from apparently true premises to an apparently false/absurd conclusion.

## A Word about the Word ‘Paradox’

What is a paradox? Roughly: an apparently good argument from apparently true premises to an apparently false/absurd conclusion.

Following Quine, there are at least three classes of paradoxes:

# A Word about the Word ‘Paradox’

What is a paradox? Roughly: an apparently good argument from apparently true premises to an apparently false/absurd conclusion.

Following Quine, there are at least three classes of paradoxes:

- Veridical paradox: we’ve decided the conclusion is true despite any initial air of absurdity (e.g., the Barber Paradox).

# A Word about the Word 'Paradox'

What is a paradox? Roughly: an apparently good argument from apparently true premises to an apparently false/absurd conclusion.

Following Quine, there are at least three classes of paradoxes:

- Veridical paradox: we've decided the conclusion is true despite any initial air of absurdity (e.g., the Barber Paradox).
- Falsidical paradox: we've decided the conclusion is false and spotted a fallacy in the argument (e.g., the "proof" of $2 = 1$).

# A Word about the Word 'Paradox'

What is a paradox? Roughly: an apparently good argument from apparently true premises to an apparently false/absurd conclusion.

Following Quine, there are at least three classes of paradoxes:

- Veridical paradox: we've decided the conclusion is true despite any initial air of absurdity (e.g., the Barber Paradox).
- Falsidical paradox: we've decided the conclusion is false and spotted a fallacy in the argument (e.g., the "proof" of $2 = 1$).
- Genuine antinomy: we are unable to decide whether it is veridical or falsidical (e.g., the Liar paradox?).

# A Word about the Word 'Paradox'

What is a paradox? Roughly: an apparently good argument from apparently true premises to an apparently false/absurd conclusion.

Following Quine, there are at least three classes of paradoxes:

- Veridical paradox: we've decided the conclusion is true despite any initial air of absurdity (e.g., the Barber Paradox).
- Falsidical paradox: we've decided the conclusion is false and spotted a fallacy in the argument (e.g., the "proof" of $2 = 1$).
- Genuine antinomy: we are unable to decide whether it is veridical or falsidical (e.g., the Liar paradox?).

W.V. Quine. *The Ways of Paradox*. The Ways of Paradox and Other Essays. Cambridge: Harvard University Press, 1966.

Three introductory examples

# Epistemic Logic

Let $K_a P$ informally mean "agent *a* knows that $P$ (is true)".

# Epistemic Logic

Let $K_a P$ informally mean "agent $a$ knows that $P$ (is true)".

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

# Epistemic Logic

Let $K_a P$ informally mean "agent $a$ knows that $P$ (is true)".

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

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

# Epistemic Logic

Let $K_a P$ informally mean "agent $a$ knows that $P$ (is true)".

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

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

$K_a P \vee K_a \neg P$: "Ann knows whether $P$ is true"

# Epistemic Logic

Let $K_a P$ informally mean "agent $a$ knows that $P$ (is true)".

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

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

$K_a P \vee K_a \neg P$: "Ann knows whether $P$ is true"

$\neg K_a \neg P$: "$P$ is an epistemic possibility for Ann"

# Epistemic Logic

Let $K_a P$ informally mean "agent $a$ knows that $P$ (is true)".

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

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

$K_a P \vee K_a \neg P$: "Ann knows whether $P$ is true"

$\neg K_a \neg P$: "$P$ is an epistemic possibility for Ann"

$K_a K_a P$: "Ann knows that she 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.

## 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.

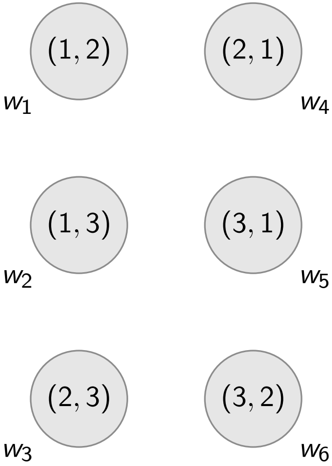
What are the relevant states?

# 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?

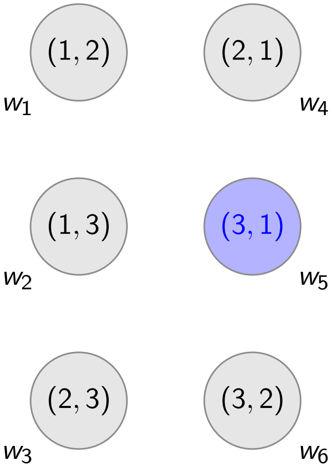


## Example

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

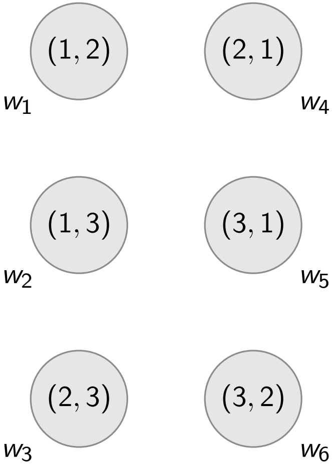


# Example

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?

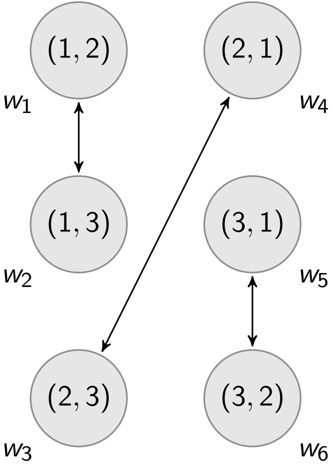


## Example

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?

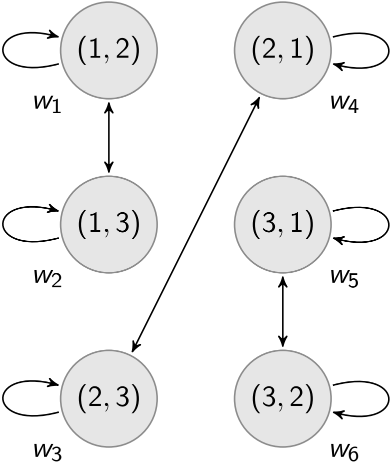


# Example

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

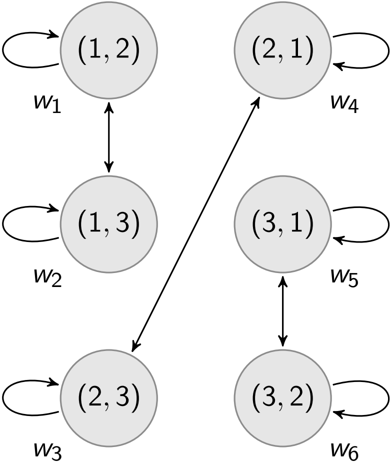
Suppose there are three cards: 1, 2 and 3.

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

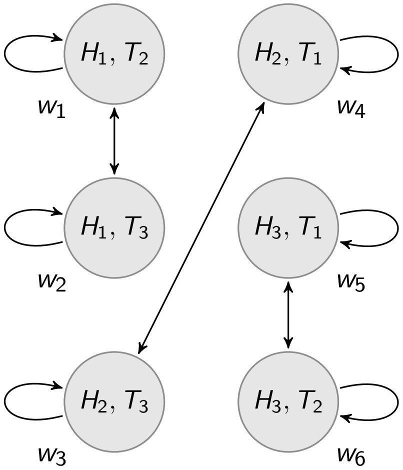
Suppose there are three cards: 1, 2 and 3.

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"

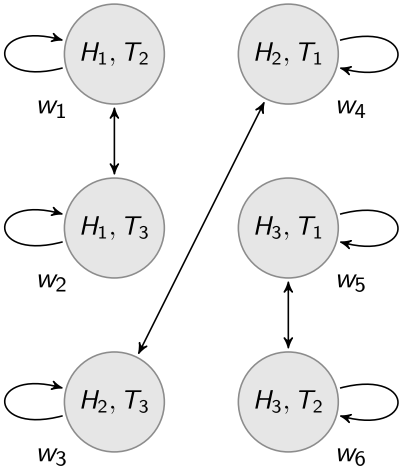
Eg., $V(H_1) = \{w_1, w_2\}$



# 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.

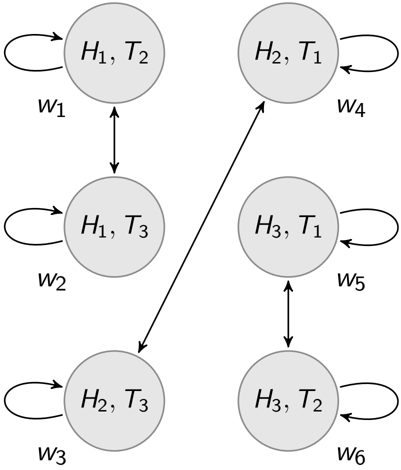


# Example

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.

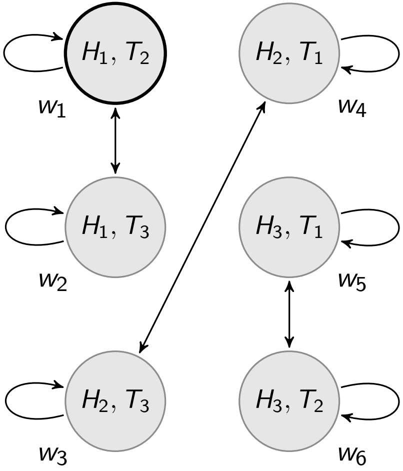


## Example

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.

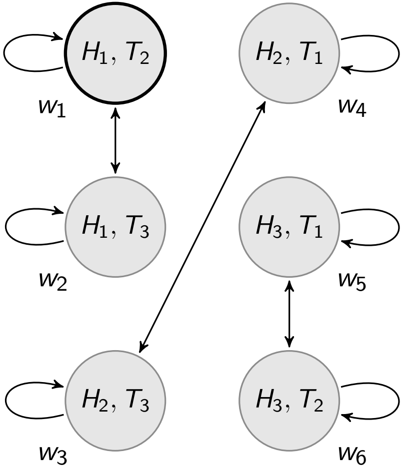


# Example

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.

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

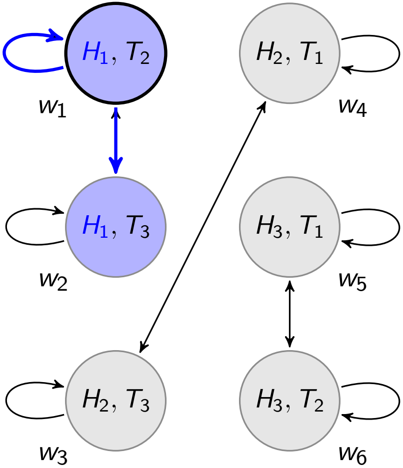


# Example

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.

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



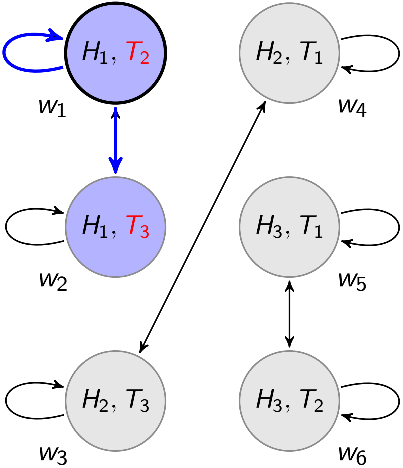
## Example

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.

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

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

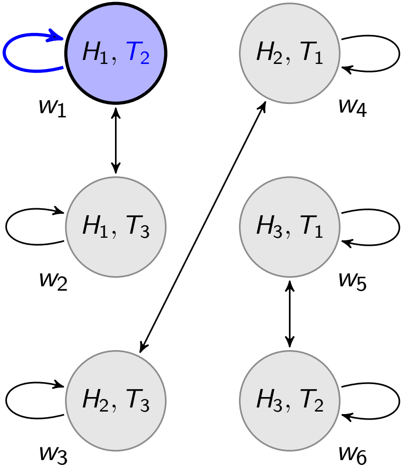


# Example

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.

$\mathcal{M}, w_1 \models \neg K_a \neg T_2$

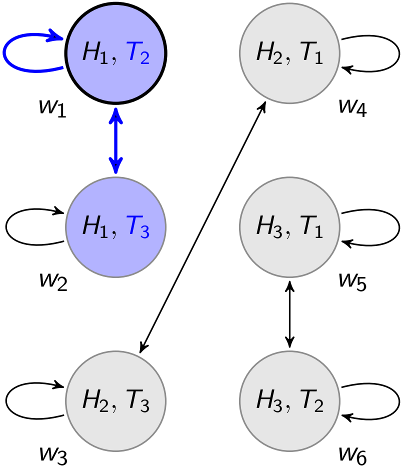


## Example

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.

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



# 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 $P$"

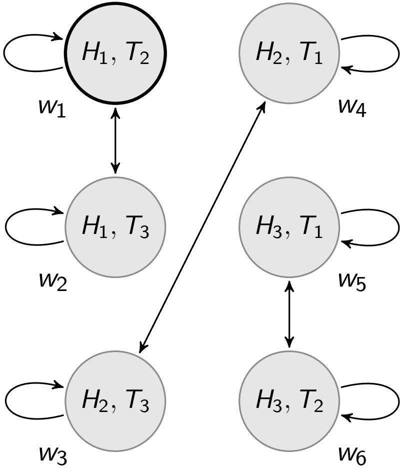
- $K_a K_b \varphi$: "Ann knows that Bob knows $\varphi$"
- $K_a(K_b\varphi \vee K_b\neg\varphi)$: "Ann knows that Bob knows whether $\varphi$
- $\neg K_b K_a K_b(\varphi)$: "Bob does not know that Ann knows that Bob knows that $\varphi$"

# 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.

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

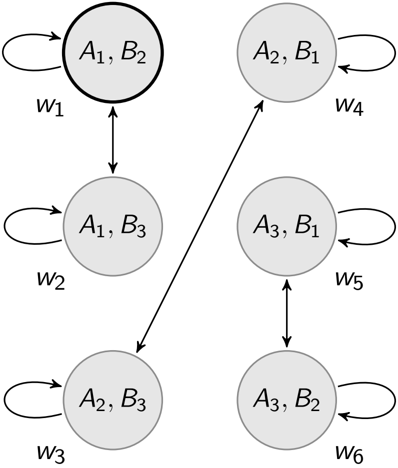


# Example

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.

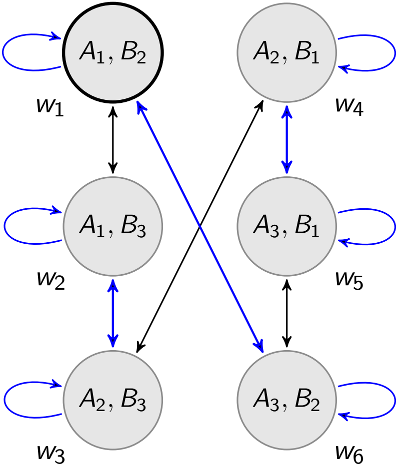


# Example

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.

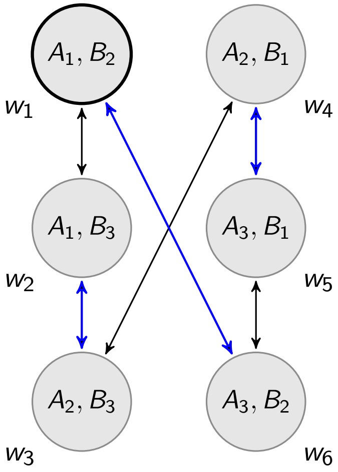


# Example

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.



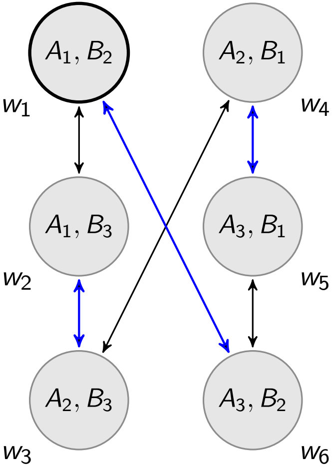
# Example

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.

$\mathcal{M}, w_1 \models K_b(K_a A_1 \vee K_a \neg A_1)$



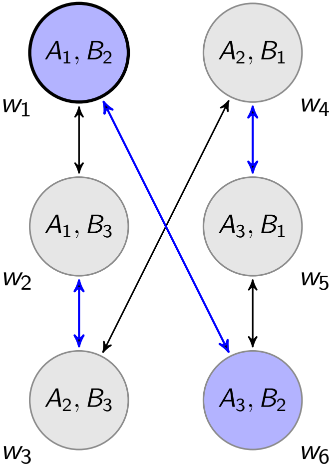
## Example

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.

$\mathcal{M}, w_1 \models K_b(K_a A_1 \vee K_a \neg A_1)$



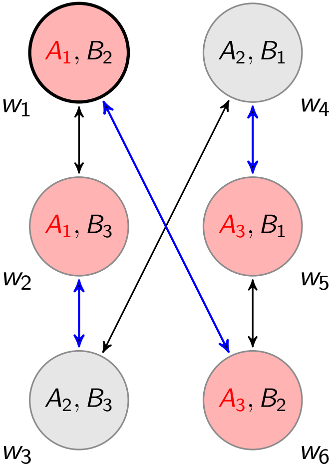
# Example

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.

$\mathcal{M}, w_1 \models K_b(K_a A_1 \vee K_a \neg A_1)$



# College Park and Amsterdam

Let $K_c$ stand for **agent** $c$ **knows that** and $K_a$ stand for **agent** $a$ **knows that**. Suppose agent $c$, who lives in College Park, knows that agent $a$ lives in Amsterdam. Let $r$ stand for 'it's raining in Amsterdam'. Although $c$ doesn't know whether it's raining in Amsterdam, $c$ knows that $a$ knows whether it's raining there:

## College Park and Amsterdam

Let $K_c$ stand for **agent** $c$ **knows that** and $K_a$ stand for **agent** $a$ **knows that**. Suppose agent $c$, who lives in College Park, knows that agent $a$ lives in Amsterdam. Let $r$ stand for 'it's raining in Amsterdam'. Although $c$ doesn't know whether it's raining in Amsterdam, $c$ knows that $a$ knows whether it's raining there:

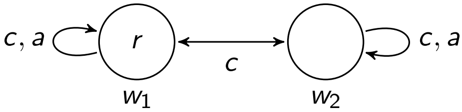
$$\neg(K_c r \lor K_c \neg r) \land K_c(K_a r \lor K_a \neg r).$$

# College Park and Amsterdam

Let $K_c$ stand for **agent** $c$ **knows that** and $K_a$ stand for **agent** $a$ **knows that**. Suppose agent $c$, who lives in College Park, knows that agent $a$ lives in Amsterdam. Let $r$ stand for 'it's raining in Amsterdam'. Although $c$ doesn't know whether it's raining in Amsterdam, $c$ knows that $a$ knows whether it's raining there:

$$\neg(K_c r \lor K_c \neg r) \land K_c(K_a r \lor K_a \neg r).$$

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

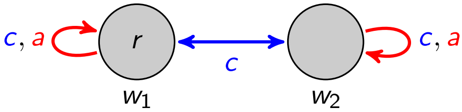


# College Park and Amsterdam

Let $K_c$ stand for **agent** $c$ **knows that** and $K_a$ stand for **agent** $a$ **knows that**. Suppose agent $c$, who lives in College Park, knows that agent $a$ lives in Amsterdam. Let $r$ stand for 'it's raining in Amsterdam'. Although $c$ doesn't know whether it's raining in Amsterdam, $c$ knows that $a$ knows whether it's raining there:

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

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

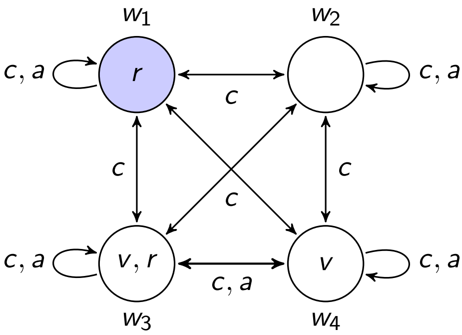


Now suppose that agent $c$ doesn't know whether agent $a$ has left Amsterdam for a vacation. (Let $v$ stand for '$a$ has left Amsterdam on vacation'.) Agent $c$ knows that if $a$ is not on vacation, then $a$ knows whether it's raining in Amsterdam; but if $a$ is on vacation, then $a$ won't bother to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

$$K_c(\neg v \to (K_a r \lor K_a \neg r)) \land K_c(v \to \neg(K_a r \lor K_a \neg r)).$$

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

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



The Muddy Children Puzzle

Three children are outside playing. Two of them get mud on their forehead. They cannot see or feel the mud on their own foreheads, but can see who is dirty.

Three children are outside playing. Two of them get mud on their forehead. They cannot see or feel the mud on their own foreheads, but can see who is dirty.

Their mother enters the room and says "At least one of you have mud on your forehead".

Three children are outside playing. Two of them get mud on their forehead. They cannot see or feel the mud on their own foreheads, but can see who is dirty.

Their mother enters the room and says "At least one of you have mud on your forehead".

Then the children are repeatedly asked "do you know if you have mud on your forehead?"

Three children are outside playing. Two of them get mud on their forehead. They cannot see or feel the mud on their own foreheads, but can see who is dirty.

Their mother enters the room and says "At least one of you have mud on your forehead".

Then the children are repeatedly asked "do you know if you have mud on your forehead?"

What happens?

Three children are outside playing. Two of them get mud on their forehead. They cannot see or feel the mud on their own foreheads, but can see who is dirty.

Their mother enters the room and says "At least one of you have mud on your forehead".

Then the children are repeatedly asked "do you know if you have mud on your forehead?"

What happens?

**Claim:** After first question, the children answer "I don't know",

Three children are outside playing. Two of them get mud on their forehead. They cannot see or feel the mud on their own foreheads, but can see who is dirty.

Their mother enters the room and says "At least one of you have mud on your forehead".

Then the children are repeatedly asked "do you know if you have mud on your forehead?"

What happens?

**Claim:** After first question, the children answer "I don't know", after the second question the muddy children answer "I have mud on my forehead!" (but the clean child is still in the dark).

Three children are outside playing. Two of them get mud on their forehead. They cannot see or feel the mud on their own foreheads, but can see who is dirty.

Their mother enters the room and says "At least one of you have mud on your forehead".

Then the children are repeatedly asked "do you know if you have mud on your forehead?"

What happens?

**Claim:** After first question, the children answer "I don't know", after the second question the muddy children answer "I have mud on my forehead!" (but the clean child is still in the dark). Then the clean child says, "Oh, I must be clean."

# Muddy Children

Assume:

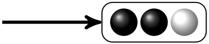
- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.

# Muddy Children

Assume:

- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.

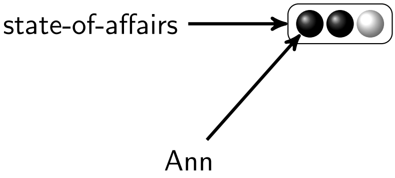
state-of-affairs



# Muddy Children

Assume:

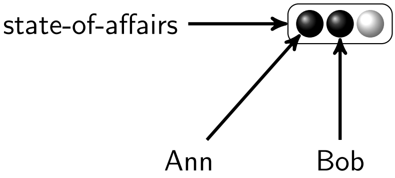
- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.



# Muddy Children

Assume:

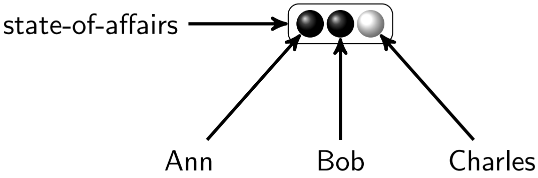
- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.



## Muddy Children

Assume:

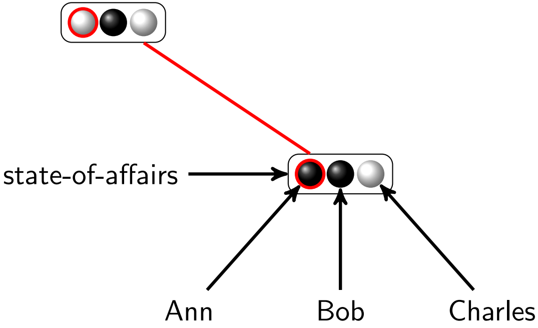
- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.



# Muddy Children

Assume:

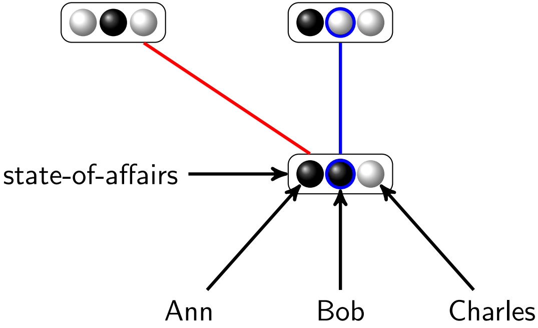
- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.



# Muddy Children

Assume:

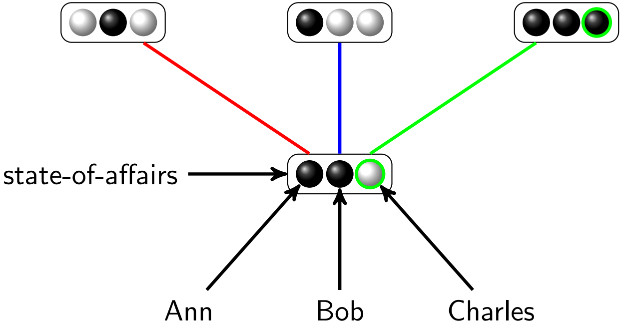
- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.



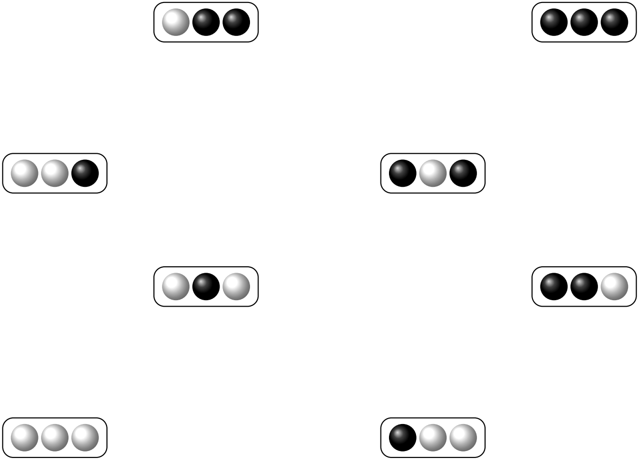
# Muddy Children

Assume:

- There are three children: Ann, Bob and Charles.
- (Only) Ann and Bob have mud on their forehead.

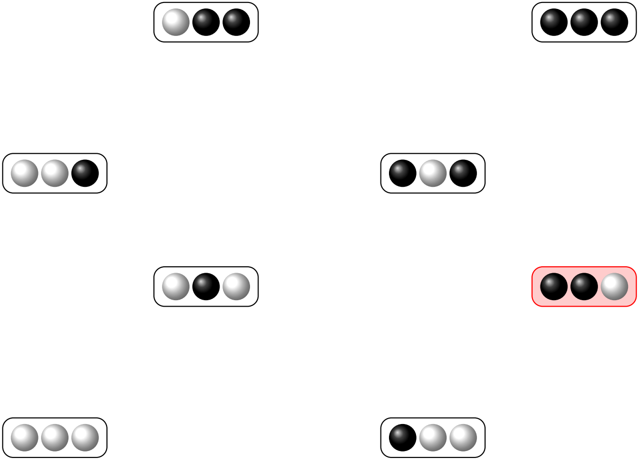


# Muddy Children



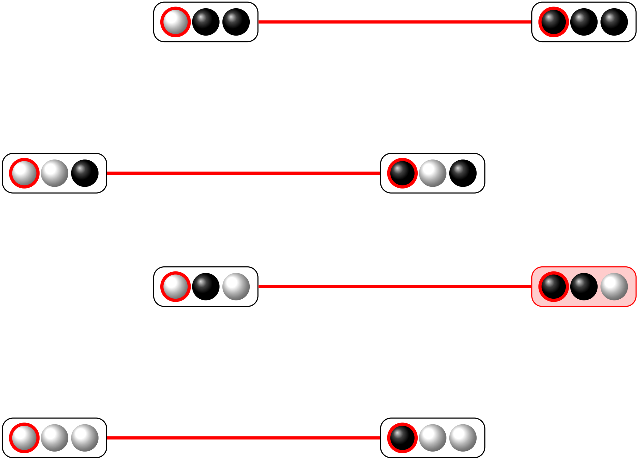
The 8 possible situations

# Muddy Children



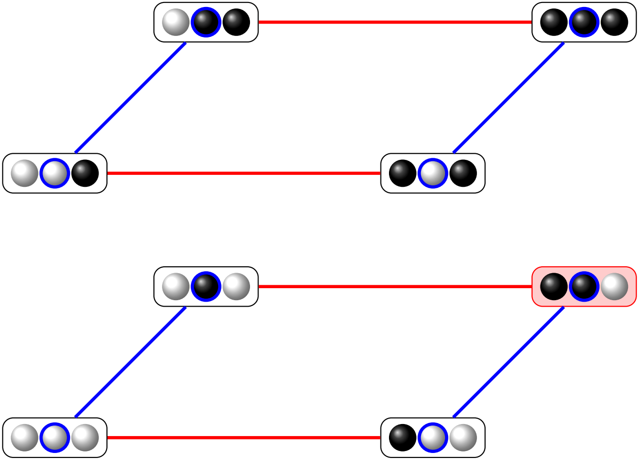
The actual situation

# Muddy Children



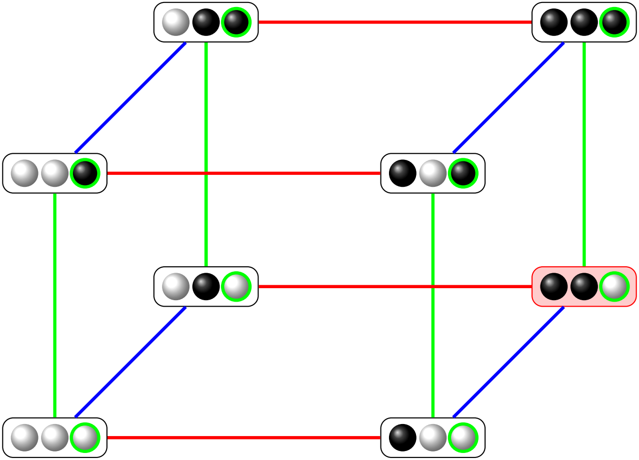
Ann's uncertainty

# Muddy Children



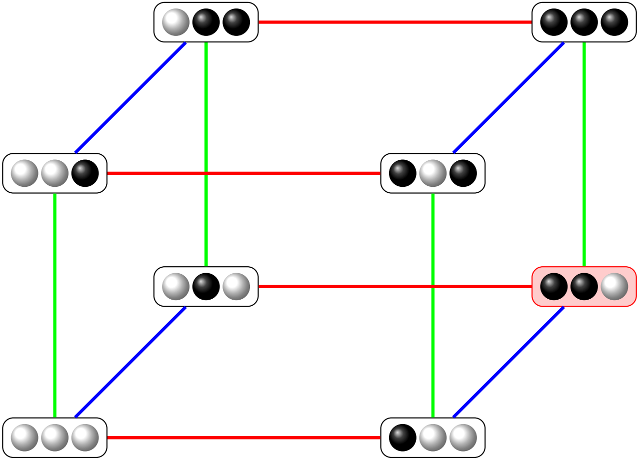
Bob's uncertainty

# Muddy Children

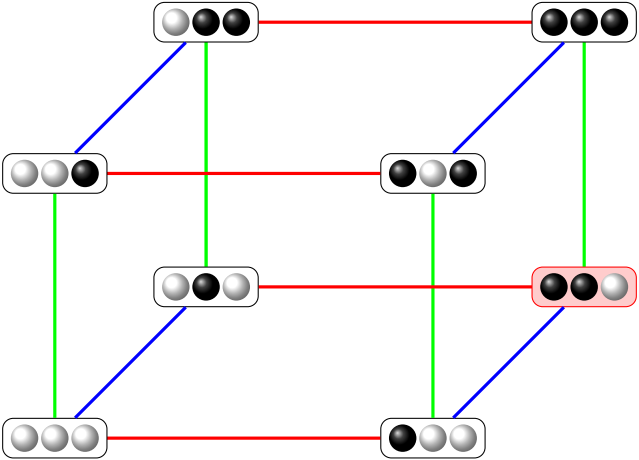


Charles' uncertainty

# Muddy Children

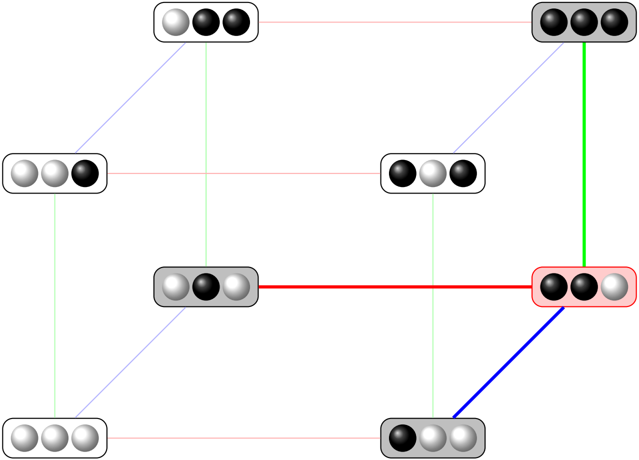


# Muddy Children



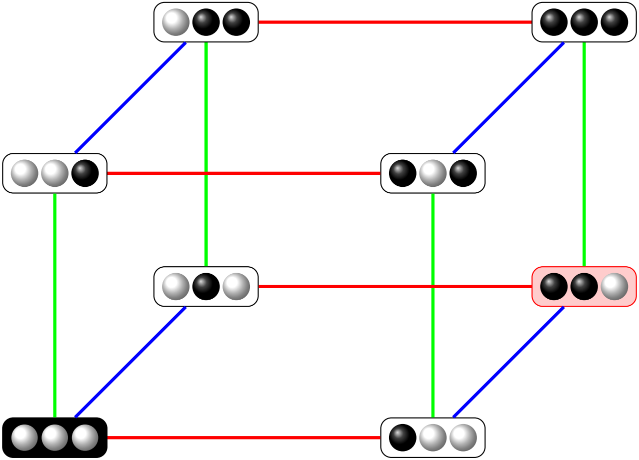
None of the children know if they are muddy

# Muddy Children



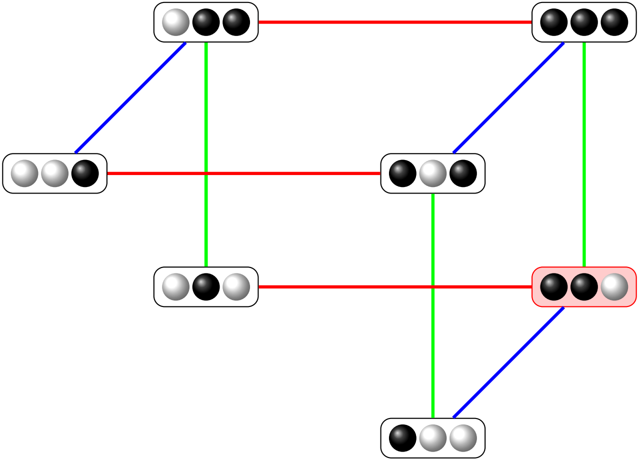
None of the children know if they are muddy

# Muddy Children



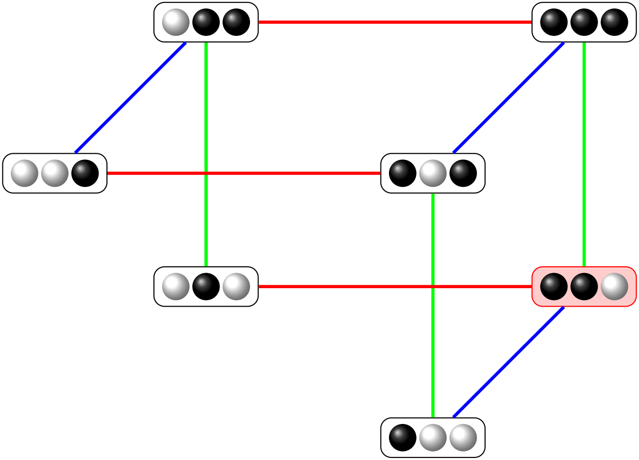
"At least one has mud on their forehead."

# Muddy Children



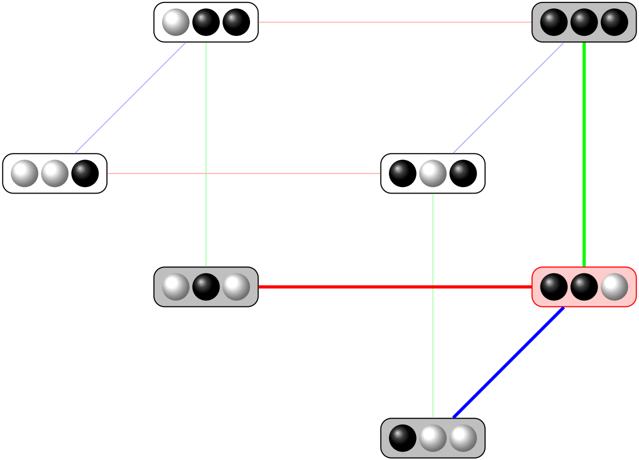
“At least one has mud on their forehead.”

# Muddy Children



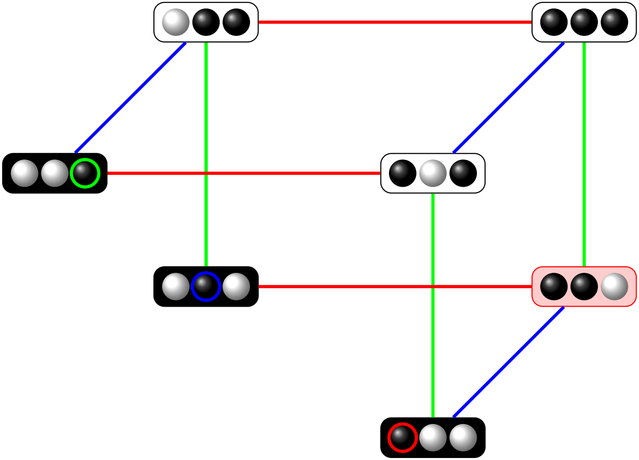
"Who has mud on their forehead?"

# Muddy Children



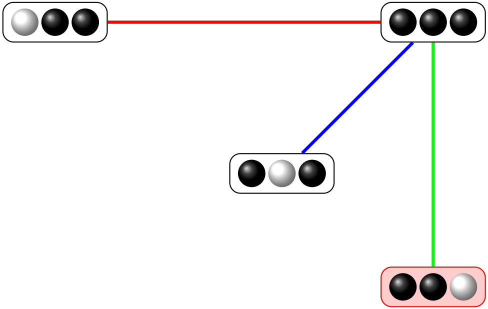
"Who has mud on their forehead?"

# Muddy Children



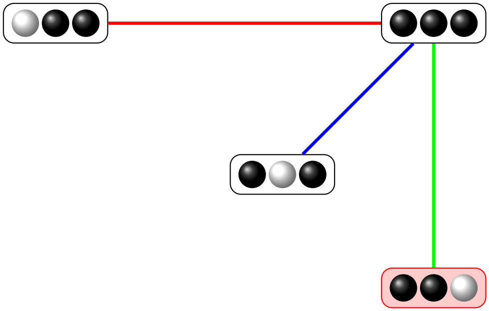
No one steps forward.

# Muddy Children



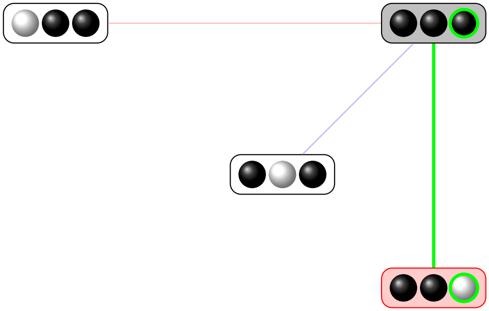
No one steps forward.

# Muddy Children



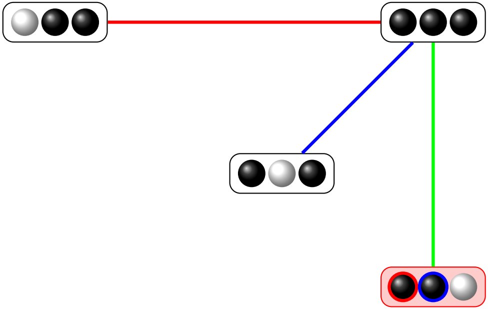
“Who has mud on their forehead?”

# Muddy Children



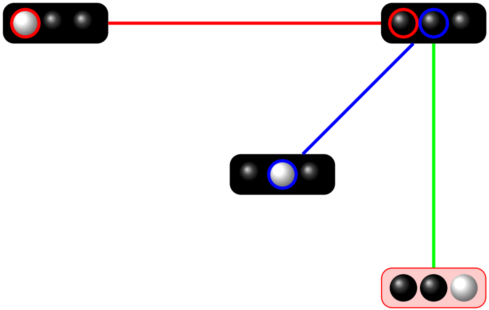
Charles does not know he is clean.

# Muddy Children



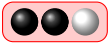
Ann and Bob step forward.

# Muddy Children



Ann and Bob step forward.

# Muddy Children



Now, Charles knows he is clean.