Learning and Epistemic Modal Logic
Part 1

Nina Gierasimczuk

Department of Applied Mathematics and Computer Science
Technical University of Denmark

FoPSS: Logic and Learning
Oxford, July 5th, 2018

Outline

Outline

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

A♠

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

A♠ Q♠

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

A♠ Q♠ 3♠

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

A♠ Q♠ 3♠ A♠

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

A♠ Q♠ 3♠ A♠ Q♠

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

A♠ Q♠ 3♠ A♠ Q♠ 4♡

Inductive Inference: Card Game

What is the rule behind this sequence of cards?

A♠ Q♠ 3♠ A♠ Q♠ 4♡ . . .

How many different abstract scenarios?

Assume we have at our disposal unlimited amount of playing cards.

How many different abstract scenarios?

Assume we have at our disposal unlimited amount of playing cards.

1. How many different (kinds of) playing cards do we have?

How many different abstract scenarios?

Assume we have at our disposal unlimited amount of playing cards.

1. How many different (kinds of) playing cards do we have?
2. How many different beginnings of length 1?

How many different abstract scenarios?

Assume we have at our disposal unlimited amount of playing cards.

1. How many different (kinds of) playing cards do we have?
2. How many different beginnings of length 1?
3. How many different beginnings of length 2?

How many different abstract scenarios?

Assume we have at our disposal unlimited amount of playing cards.

1. How many different (kinds of) playing cards do we have?
2. How many different beginnings of length 1?
3. How many different beginnings of length 2?
4. How many different infinite sequences?

1.	A♠	A♠	A♠	A♠	A♠	A♠	...
2.	A♣	A♣	A♣	A♣	A♣	A♣	...
3.	A♡	A♡	A♡	A♡	A♡	A♡	...
4.	A♢	Q♠	3♠	8♡	2♡	5♠	...
5.	A♠	Q♠	7♠	J♠	5♠	5♠	...
...							
m.	A♣	A♡	A♣	A♡	A♣	A♢	...
...							

1.	A♠	A♠	A♠	A♠	A♠	A♠	...
2.	A♣	A♣	A♣	A♣	A♣	A♣	...
3.	A♡	A♡	A♡	A♡	A♡	A♡	...
4.	A♢	Q♠	3♠	8♡	2♡	5♠	...
5.	A♠	Q♠	7♠	J♠	5♠	5♠	...
...							
m.	A♣	A♡	A♣	A♡	A♣	A♢	...
...							

1.	A♠	A♠	A♠	A♠	A♠	A♠	...
2.	A♣	A♣	A♣	A♣	A♣	A♣	...
3.	A♡	A♡	A♡	A♡	A♡	A♡	...
4.	A♢	Q♠	3♠	8♡	2♡	5♠	...
5.	A♠	Q♠	7♠	J♠	5♠	5♠	...
...							
m.	A♣	A♡	A♣	A♡	A♣	A♢	...
...							

The Infinities of Inductive Inference

1.	A♠	A♠	A♠	A♠	A♠	A♠	...
2.	A♣	A♣	A♣	A♣	A♣	A♣	...
3.	A♡	A♡	A♡	A♡	A♡	A♡	...
4.	A♢	Q♠	3♠	8♡	2♡	5♠	...
5.	A♠	Q♠	7♠	J♠	5♠	5♠	...
...							
m.	A♣	A♡	A♣	A♡	A♣	A♢	...
...							

The Infinities of Inductive Inference

1.	A♠	A♠	A♠	A♠	A♠	A♠	...
2.	A♣	A♣	A♣	A♣	A♣	A♣	...
3.	A♡	A♡	A♡	A♡	A♡	A♡	...
4.	A♢	Q♠	3♠	8♡	2♡	5♠	...
5.	A♠	Q♠	7♠	J♠	5♠	5♠	...
...							
m.	A♣	A♡	A♣	A♡	A♣	A♢	...
...							

The Infinities of Inductive Inference

1.	A♠	A♠	A♠	A♠	A♠	A♠	...
2.	A♣	A♣	A♣	A♣	A♣	A♣	...
3.	A♡	A♡	A♡	A♡	A♡	A♡	...
4.	A♢	Q♠	3♠	8♡	2♡	5♠	...
5.	A♠	Q♠	7♠	J♠	5♠	5♠	...
...							
m.	A♣	A♡	A♣	A♡	A♣	A♢	...
...							

1.	A♠	A♠	A♠	A♠	A♠	A♠	...	
2.	A♣	A♣	A♣	A♣	A♣	A♣	...	
3.	A♡	A♡	A♡	A♡	A♡	A♡	...	
4.	A♢	Q♠	3♠	8♡	2♡	5♠	...	
5.	A♠	Q♠	7♠	J♠	5♠	5♠	...	
...								
m.	A♣	A♡	A♣	A♡	A♣	A♢	...	m-th
...								

1.	A♠	A♠	A♠	A♠	A♠	A♠	. . .	
2.	A♣	A♣	A♣	A♣	A♣	A♣	. . .	
3.	A♡	A♡	A♡	A♡	A♡	A♡	. . .	
4.	A♢	Q♠	3♠	8♡	2♡	5♠	. . .	
5.	A♠	Q♠	7♠	J♠	5♠	5♠	. . .	
. . .								
m.	A♣	A♡	A♣	A♡	A♣	A♢	. . .	m-th
. . .								
								. . .

How many possible rules are there?

1. In principle...

How many possible rules are there?

1. In principle...
2. Rule written down on a piece of paper.
3. Rule expressed by a natural language sentence.
4. Rule described by a theory that fills a 300 pages book.

How many possible rules are there?

1. In principle...
2. Rule written down on a piece of paper.
3. Rule expressed by a natural language sentence.
4. Rule described by a theory that fills a 300 pages book.
5. Rule encoded by a Turing Machine program.

Descriptions are finite, and there are countably many of them.

How many sequences comply to one rule?

1. The sequence has solely A♠-cards.

How many sequences comply to one rule?

1. The sequence has solely A♠-cards.
2. The sequence has solely ♠-cards.

How many sequences comply to one rule?

1. The sequence has solely A♠-cards.
2. The sequence has solely ♠-cards.
3. The sequence has ♡-cards on even places.

How many sequences comply to one rule?

1. The sequence has solely A♠-cards.
2. The sequence has solely ♠-cards.
3. The sequence has ♡-cards on even places.
4. The sequence is definable in first-order logic.
5. etc...

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A♠

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A♠ Q♠

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A♠ Q♠ 3♠

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A♠ Q♠ 3♠ A♠

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A♠ Q♠ 3♠ A♠ Q♠

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A♠ Q♠ 3♠ A♠ Q♠ 4♡

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A$\spadesuit$ Q$\spadesuit$ 3$\spadesuit$ A$\spadesuit$ Q$\spadesuit$ 4$\heartsuit$ Q$\spadesuit$

Different Hypothesis Spaces

What can we know and how can we know it?

1. $\{(\text{all cards are } \spadesuit), (\text{all cards are } \diamondsuit)\}$
2. $\{(\spadesuit \text{ at the 4-th position}), \neg(\spadesuit \text{ at the 4-th position})\}$
3. $\underline{\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\}}$
4. $\{(\text{exactly } n \text{ cards are } \heartsuit) \mid n \in \mathbb{N}\} \cup \{(\infty \text{ cards are } \heartsuit)\}$

A♠ Q♠ 3♠ A♠ Q♠ 4♡ Q♠ Q♡...

The Game of Learning in the Limit: Learner and Nature

- A class of possible worlds (known by both players).
- Nature chooses one of them (learner does not know which).
- Nature generates data about the world.
- From inductively given data learner draws her conjectures.
- With each input learner can answer with a different hypothesis.
- Learner succeeds if

The Game of Learning in the Limit: Learner and Nature

- A class of possible worlds (known by both players).
- Nature chooses one of them (learner does not know which).
- Nature generates data about the world.
- From inductively given data learner draws her conjectures.
- With each input learner can answer with a different hypothesis.
- Learner succeeds if **she stabilizes to a correct hypothesis**.

The Game of Learning in the Limit: Learner and Nature

- A class of possible worlds (known by both players).
- Nature chooses one of them (learner does not know which).
- Nature generates data about the world.
- From inductively given data learner draws her conjectures.
- With each input learner can answer with a different hypothesis.
- Learner succeeds if **she stabilizes to a correct hypothesis**.

Her success depends on her skills and on the problem.

Outline

Function Learning

1. Possible realities:

   Functions

2. Hypotheses:

   Names of functions

3. Information accessible to the learner:

   Sequences of pairs (argument, value)

4. Learner:

   Function that takes a sequence and outputs a hypothesis

5. Success criterion:

   After finite number of outputs stabilize on a correct answer

Model-theoretic Learning

1. Possible realities:

   Models of a given signature

2. Hypotheses:

   First order sentences

3. Information accessible to the learner:

   Sequences of atomic formulas and negations thereof

4. Learner:

   Function that takes a sequence and outputs a hypothesis

5. Success criterion:

   After finite number of outputs stabilize on a correct answer

Set Learning

1. Possible realities:

   Sets of integers
2. Hypotheses:

   Names of sets
3. Information accessible to the learner:

   Sequences of numbers
4. Learner:

   Function that takes a sequence and outputs a hypothesis
5. Success criterion:

   After finite number of outputs stabilize on a correct answer

Outline

Another Game

Consider the following class:

$\{1, 2, 3, \quad 5, \ldots\}$

$\{1, \quad 3, 4, 5, \ldots\}$

$\{1, 2, \quad 4, 5, \ldots\}$

$\{\quad 2, 3, 4, 5, \ldots\}$

...

Another Game

Consider the following class:
$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{ \quad 2, 3, 4, 5, \ldots\}$
$\ldots$

1

Another Game

Consider the following class:
$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{\quad 2, 3, 4, 5, \ldots\}$
$\ldots$

1, 3

Another Game

Consider the following class:

$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{ \quad 2, 3, 4, 5, \ldots\}$
$\ldots$

1, 3, 4

Another Game

Consider the following class:
$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{\quad 2, 3, 4, 5, \ldots\}$
...

1, 3, 4, 2

Consider the following class:

$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{\quad 2, 3, 4, 5, \ldots\}$
$\ldots$

1, 3, 4, 2, 6

Another Game

Consider the following class:

$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{\quad 2, 3, 4, 5, \ldots\}$
$\ldots$

1, 3, 4, 2, 6, 7

Another Game

Consider the following class:

$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{\quad 2, 3, 4, 5, \ldots\}$
$\ldots$

1, 3, 4, 2, 6, 7, 8

Another Game

Consider the following class:
$\{1, 2, 3, \quad 5, \ldots\}$
$\{1, \quad 3, 4, 5, \ldots\}$
$\{1, 2, \quad 4, 5, \ldots\}$
$\{\quad 2, 3, 4, 5, \ldots\}$
$\ldots$

1, 3, 4, 2, 6, 7, 8, . . .

Another Game

1. Are you confident? What would make you change your guess?
2. What was your "guessing rule"?
3. How do you like winning if at least one of your guess is correct?
4. And if you succeed to make a right guess and never change your mind after that? How many wrong guesses could you make under this condition?

Another Game

1. Assume that I'll give you all and only truthful clues. What would be the guessing rule to win according to the last winning condition?
2. Add $\{1, 2, 3, 4, 5, \ldots\}$. Is your guessing rule still good?
3. While keeping $\{1, 2, 3, 4, 5, \ldots\}$ in, assume that I'll guarantee they are ordered increasingly. Can you win the game?
4. Now, remove $\{1, 2, 3, 4, 5, \ldots\}$. You get only one guess—would you object to this winning condition?

1960s: The Beginnings

- Hillary Putnam (1965). Trial and error predicates and the solution to...
- E. Mark Gold (1967). Language identification in the limit.
- Ray Solomonoff (1964). A formal theory of inductive inference.

Trial and Error Predicates

A predicate (set) P is decidable if there is a effective procedure φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \varphi(x) = 1; \\ \neg P(x) & \text{iff} & \varphi(x) = 0. \end{array}$$

What happens if we modify the condition by:

1. allowing φ to change her mind any finite number of times;
2. making it impossible to diagnose termination?

P is a trial and error predicate if there is a Turing Machine φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 1; \\ \neg P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 0. \end{array}$$

Trial and error predicates are decidable in the limit.

Trial and Error Predicates

A predicate (set) P is decidable if there is a recursive function φ such that

$$\begin{array}{ccc} P(x) & \text{iff} & \varphi(x) = 1; \\ \neg P(x) & \text{iff} & \varphi(x) = 0. \end{array}$$

What happens if we modify the condition by:

1. allowing φ to change her mind any finite number of times;
2. making it impossible to diagnose termination?

P is a trial and error predicate if there is a Turing Machine φ such that

$$\begin{array}{ccc} P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 1; \\ \neg P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 0. \end{array}$$

Trial and error predicates are decidable in the limit.

Trial and Error Predicates

A predicate (set) P is decidable if there is a Turing Machine φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \varphi(x) = 1; \\ \neg P(x) & \text{iff} & \varphi(x) = 0. \end{array}$$

What happens if we modify the condition by:

1. allowing φ to change her mind any finite number of times;
2. making it impossible to diagnose termination?

P is a trial and error predicate if there is a Turing Machine φ such that

$$\begin{array}{rcl} P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 1; \\ \neg P(x) & \text{iff} & \exists k \forall n \geq k \ \varphi(x, n) = 0. \end{array}$$

Trial and error predicates are decidable in the limit.

In this context one can think of φ as of a learning function,
Especially if more than two answers are possible.

The quantifier prefix in the definition of trial and error predicates
indicates their place in arithmetic hierarchy.

$$
\begin{array}{ccccccccccccc}
 & & \mathbf{\Sigma}^0_1 & & & & \mathbf{\Sigma}^0_2 & \cdots & & & \mathbf{\Sigma}^0_\alpha & & & \cdots \\
 & \nearrow & & \searrow & & \nearrow & & & & \nearrow & & \searrow & & \\
\mathbf{\Delta}^0_1 & & & & \mathbf{\Delta}^0_2 & & & \cdots & \mathbf{\Delta}^0_\alpha & & & & \mathbf{\Delta}^0_{\alpha+1} & \cdots \\
 & \searrow & & \nearrow & & \searrow & & & & \searrow & & \nearrow & & \\
 & & \mathbf{\Pi}^0_1 & & & & \mathbf{\Pi}^0_2 & \cdots & & & \mathbf{\Pi}^0_\alpha & & & \cdots
\end{array}
$$

We will focus on a more general case,
when learner has to pick from more than two options,
in fact, from countably many options.

Some Basic Definitions

Let $\mathbb{N}$ stand for positive integers and $S \subseteq \mathbb{N}$.
Let $\mathcal{S} = (S_i)_{i \in \mathbb{N}}$ be a family of sets.

Definition
By a stream t for S we mean an infinite sequence of elements from S enumerating all and only the elements from S (allowing repetitions).

Definition
We will use the following notation:

- t_n is the n-th element of t;
- $t[n]$ is the sequence $(t_0, t_2, \ldots, t_{n-1})$;
- $content(t)$ is the set of elements that occur in t;
- $\varphi : \mathbb{N}^* \to \mathbb{N}$ is a learning function.

Identifiability in the Limit

Definition

Learning function φ:

1. identifies $S_i \in \mathcal{S}$ in the limit on t iff for co-finitely many m, $\varphi(t[m]) = i$;
2. identifies $S_i \in \mathcal{S}$ in the limit iff it identifies S_i in the limit on every t for S_i;
3. identifies $\mathcal{S}$ in the limit iff it identifies in the limit every $S_i \in \mathcal{S}$.

$\mathcal{S}$ is identifiable in the limit iff some learning function identifies $\mathcal{S}$ in the limit.

Some Examples

Example
Let $\mathcal{S}_1 = \{S_i \mid i \in \mathbb{N} - \{0\}\}$, where $S_n = \{1, \ldots, n\}$.

$\mathcal{S}_1$ is identifiable in the limit by the following function $\varphi : \mathbb{N}^* \to \mathbb{N}$:

$$\varphi(t[n]) = \max(content(t[n])).$$

Example

Let $\mathcal{S}_2 = \{S_i \mid i \in \mathbb{N}\}$, where $S_0 = \mathbb{N}$ and for $n \geq 1$, $S_n = \{1, \ldots, n\}$.

$\mathcal{S}_2$ is not identifiable in the limit.

Argument

To show that this is the case, let us assume that there is a function φ that identifies $\mathcal{S}_2$. We will construct a text, t on which φ fails:
t starts by enumerating $\mathbb{N}$ in order: $0, 1, 2, \ldots$
if at a number k learner φ decides it is S_0, t starts repeating k indefinitely.
This means t is a text for S_k.
As soon as φ decides it is S_k we continue with $k+1, k+2, \ldots$, so t will become a text for S_0, etc.
This shows that there is a text for a set from $\mathcal{S}_2$ on which φ fails.

Some Examples

Example

Let $\mathcal{S}_4 = \{S_n \mid S_n = \mathbb{N} - \{n\}, n \in \mathbb{N}\}$.

$\mathcal{S}_4$ is identifiable in the limit by the learning function $\varphi : \mathbb{N}^* \to \mathbb{N}$:

$$\varphi(t[n]) = \min(\mathbb{N} - \mathit{content}(t[n])).$$

Gold's Theorems

Theorem (1)

The class of all finite languages is identifiable.

Theorem (2)

The class containing all finite and at least one infinite language is not identifiable.

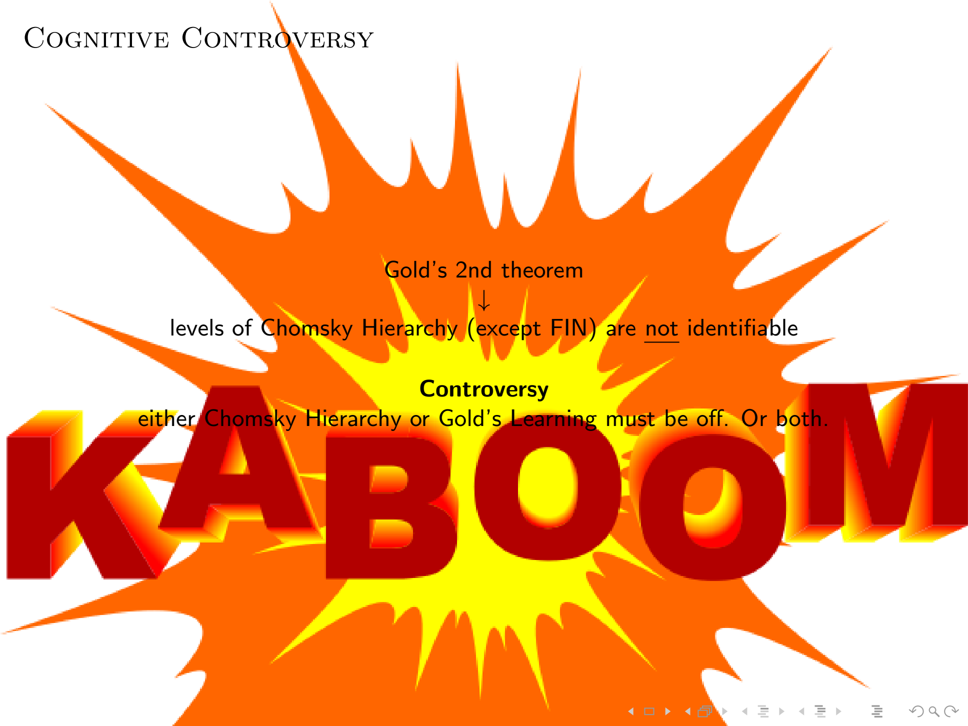
Gold's 2nd theorem

$\downarrow$

levels of Chomsky Hierarchy (except FIN) are <u>not</u> identifiable

Controversy

either Chomsky Hierarchy or Gold's Learning must be off. Or both.



Gold's 2nd theorem

$\downarrow$

levels of Chomsky Hierarchy (except FIN) are <u>not</u> identifiable

Controversy

either Chomsky Hierarchy or Gold's Learning must be off. Or both.

Limitting knowledge

True, there are good reasons for preferring the computable way of deriving knowledge. We know the results of computations and only think we know the results of trial and error procedures. There are many reasons for preferring knowing to thinking (as Popper, 1966, observed). But that does not change the fact that sometimes thinking may be more appropriate.

Outline

- Besides of the current state of affairs,
- there is a number of other **possible states of affairs** or "worlds".

An agent knows a fact φ if φ is true at all the worlds she **considers possible**.

San Francisco Example

Agent a is walking on the streets in San Francisco on a sunny day. She has no information at all about the weather in London.

San Francisco Example

Agent *a* is walking on the streets in San Francisco on a sunny day. She has no information at all about the weather in London.

Thus, in all the worlds that she considers possible, it is sunny in San Francisco.

San Francisco Example

Agent *a* is walking on the streets in San Francisco on a sunny day. She has no information at all about the weather in London.

Thus, in all the worlds that she considers possible, it is sunny in San Francisco.

Since the agent has no information about the weather in London, there are worlds she considers possible in which it is sunny in London, and others in which it is raining in London.

San Francisco Example

Agent *a* is walking on the streets in San Francisco on a sunny day. She has no information at all about the weather in London.

Thus, in all the worlds that she considers possible, it is sunny in San Francisco.

Since the agent has no information about the weather in London, there are worlds she considers possible in which it is sunny in London, and others in which it is raining in London.

Thus, this agent knows that it is sunny in San Francisco, but she does not know whether it is sunny in London.

San Francisco Example

Agent *a* is walking on the streets in San Francisco on a sunny day. She has no information at all about the weather in London.

Thus, in all the worlds that she considers possible, it is sunny in San Francisco.

Since the agent has no information about the weather in London, there are worlds she considers possible in which it is sunny in London, and others in which it is raining in London.

Thus, this agent knows that it is sunny in San Francisco, but she does not know whether it is sunny in London.

Intuitively, **the fewer worlds = less uncertainty, and more knowledge**.

San Francisco Example

Agent a is walking on the streets in San Francisco on a sunny day. She has no information at all about the weather in London.

Thus, in all the worlds that she considers possible, it is sunny in San Francisco.

Since the agent has no information about the weather in London, there are worlds she considers possible in which it is sunny in London, and others in which it is raining in London.

Thus, this agent knows that it is sunny in San Francisco, but she does not know whether it is sunny in London.

Intuitively, **the fewer worlds = less uncertainty, and more knowledge**.

?

San Francisco Example

If the agent acquires additional information from a reliable source:

It is currently sunny in London.

She would no longer consider possible worlds in which it is raining in London.

If the agent acquires additional information from a reliable source:

It is currently sunny in London.

She would no longer consider possible worlds in which it is raining in London.

Epistemic Logic: Brief History

Epistemic logic has been introduced as a modal logic in **1962** by **Jaakko Hintikka**, a Finnish philosopher and logician.

In his logic both knowledge and belief are introduced as two separate concepts. His logic had two modal operators K and B (for knowledge and belief) to represent the two attitudes separately.

Definition (Language of Epistemic Logic)

Φ is a set of propositions, with $p \in \Phi$, and $\mathcal{A} = \{1, \ldots, n\}$ is a set of agents.

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi$$

where $\top$ abbreviates a tautology and $i \in \mathcal{A}$ is the name of some agent.

In case we are only dealing with one agent, we can also omit the index.

Syntax: Language of Epistemic Logic

$K\varphi$:

Syntax: Language of Epistemic Logic

$K\varphi$: I know that φ.

Syntax: Language of Epistemic Logic

$K\varphi$: I know that φ.

$\neg K\varphi$:

Syntax: Language of Epistemic Logic

$K\varphi$: I know that φ.

$\neg K\varphi$: I don't know that φ.

Syntax: Language of Epistemic Logic

$K\varphi$: I know that φ.

$\neg K\varphi$: I don't know that φ.

$K\neg\varphi$:

Syntax: Language of Epistemic Logic

$K\varphi$: I know that φ.

$\neg K\varphi$: I don't know that φ.

$K\neg\varphi$: I know that not φ.

Syntax: Language of Epistemic Logic

$K\varphi$: I know that φ.

$\neg K\varphi$: I don't know that φ.

$K\neg\varphi$: I know that not φ.

$\neg K\neg\varphi$:

Syntax: Language of Epistemic Logic

$$\begin{aligned} K\varphi &: \text{ I know that } \varphi. \\ \neg K\varphi &: \text{ I don't know that } \varphi. \\ K\neg\varphi &: \text{ I know that not } \varphi. \\ \neg K\neg\varphi &: \text{ I don't know that not } \varphi \end{aligned}$$

Syntax: Language of Epistemic Logic

$$\begin{aligned}
K\varphi &: \text{ I know that } \varphi. \\
\neg K\varphi &: \text{ I don't know that } \varphi. \\
K\neg\varphi &: \text{ I know that not } \varphi. \\
\neg K\neg\varphi &: \text{ I don't know that not } \varphi = \text{I consider it possible that } \varphi.
\end{aligned}$$

How can we express "I don't know whether φ"?

Syntax: Language of Epistemic Logic

$K\varphi$: I know that φ.

$\neg K\varphi$: I don't know that φ.

$K\neg\varphi$: I know that not φ.

$\neg K\neg\varphi$: I don't know that not φ = I consider it possible that φ.

How can we express "I don't know whether φ"?

$$\neg(K\varphi \vee K\neg\varphi)$$

or, equivalently:

$$\neg K\varphi \wedge \neg K\neg\varphi.$$

DEFINITION (POSSIBLE WORLD MODEL OR EPISTEMIC MODEL)

A **possible world model** M for n agents over Φ is $(S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$, where:

1. a non-empty set states (or worlds) S;
2. an interpretation π which associates with every state a truth assignment to the propositions, i.e.:

   for each state $s \in S$, $\pi(s) : Prop \rightarrow \{0, 1\}$;
3. for each agent i, $\mathcal{K}_i$ is a binary relation on S.

A **pointed possible world model** is a pair (M, s), where $M = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$ and $s \in S$.

Additional Explanations

1. $\pi(s)$ tells us whether a proposition is true or false in state s.
2. $\mathcal{K}_i$ captures the possibility relation according to agent i, i.e.,

   $(s, t) \in \mathcal{K}_i$ if agent i considers world t possible, given her information in world s.

   $\mathcal{K}_i$ is a possibility (or accessibility) relation; it says what worlds agent i considers possible (or can access) in any given world.

Equivalence Possibility Relation

$\mathcal{K}_i$ is an **equivalence** relation on S, i.e., it is a binary relation that is:

1. reflexive: for all $s \in S$, we have $(s, s) \in \mathcal{K}_i$,
2. symmetric: for all $s, t \in S$, we have $(s, t) \in \mathcal{K}_i$ iff $(t, s) \in \mathcal{K}_i$,
3. transitive: for all $s, t, u \in S$, we have that if $(s, t) \in \mathcal{K}_i$ and $(t, u) \in \mathcal{K}_i$, then $(s, u) \in \mathcal{K}_i$.

Equivalence Possibility Relation

$\mathcal{K}_i$ is an **equivalence** relation on S, i.e., it is a binary relation that is:

1. reflexive: for all $s \in S$, we have $(s, s) \in \mathcal{K}_i$,
2. symmetric: for all $s, t \in S$, we have $(s, t) \in \mathcal{K}_i$ iff $(t, s) \in \mathcal{K}_i$,
3. transitive: for all $s, t, u \in S$, we have that if $(s, t) \in \mathcal{K}_i$ and $(t, u) \in \mathcal{K}_i$, then $(s, u) \in \mathcal{K}_i$.

Unless specified otherwise, we will assume the relation to be equivalence.

When is a formula true in a situation?

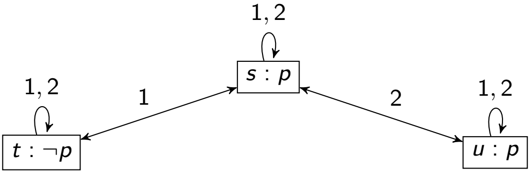
We write $(M, s) \models \varphi$ to say that φ is true at s in M, and $(M, s) \not\models \varphi$ to express that φ is false at s in M.

Definition

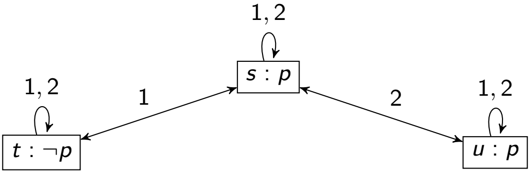
$(M, s) \models p$	iff	$\pi(s)(p) = 1$
$(M, s) \models \neg\varphi$	iff	$(M, s) \not\models \varphi$
$(M, s) \models \varphi \wedge \psi$	iff	$(M, s) \models \varphi$ and $(M, s) \models \psi$
$(M, s) \models K_i\varphi$	**iff**	**for all** v **with** $(s, v) \in \mathcal{K}_i$**,** $(M, v) \models \varphi$

So, $K_i\varphi$ is false at state s exactly when there is some v with $(s, v) \in \mathcal{K}_i$ such that φ is false at v.

Example 1: Is it sunny in San Francisco?

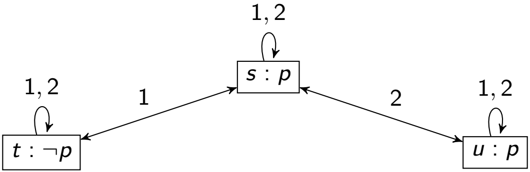


Example 1: Is it sunny in San Francisco?



Is it true that:

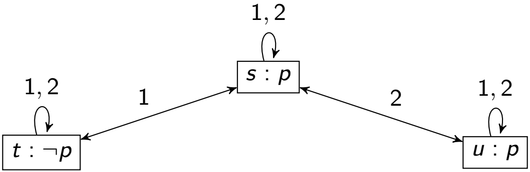
Example 1: Is it sunny in San Francisco?



Is it true that:

$$(M, s) \models p$$

Example 1: Is it sunny in San Francisco?

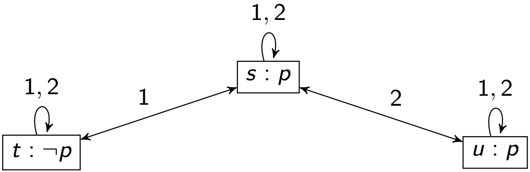


Is it true that:

$$(M, s) \models p$$

$$(M, s) \models \neg K_1 p$$

Example 1: Is it sunny in San Francisco?



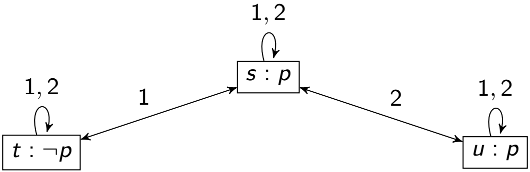
Is it true that:

$$(M, s) \models p$$

$$(M, s) \models \neg K_1 p$$

$$(M, u) \models K_1 p$$

Example 1: Is it sunny in San Francisco?



Is it true that:

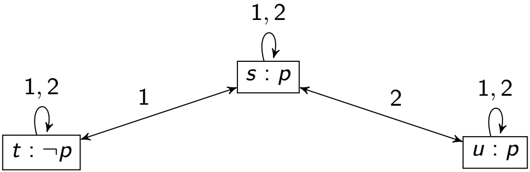
$(M, s) \models p$

$(M, s) \models K_2 p$

$(M, s) \models \neg K_1 p$

$(M, u) \models K_1 p$

Example 1: Is it sunny in San Francisco?



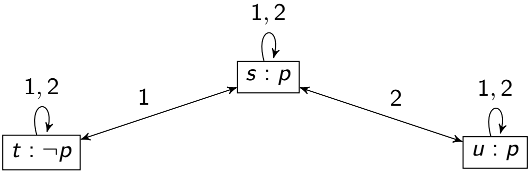
Is it true that:

$$(M, s) \models p \qquad (M, s) \models K_2 p$$

$$(M, s) \models \neg K_1 p \qquad (M, s) \models K_1(K_2 p \vee K_2 \neg p)$$

$$(M, u) \models K_1 p$$

Example 1: Is it sunny in San Francisco?



Is it true that:

$$(M, s) \models p$$

$$(M, s) \models K_2 p$$

$$(M, s) \models \neg K_1 p$$

$$(M, s) \models K_1(K_2 p \vee K_2 \neg p)$$

$$(M, u) \models K_1 p$$

$$(M, s) \models \neg K_2 \neg K_1 p$$

Imagine n children are playing outside together. Now it happens that during their play some of them, say k get mud on their foreheads. Each can see mud on others but not on his own forehead.

Along comes the father, who says, "At least one of you have mud on your forehead". The father then asks the following question, over and over: "Does any of you know whether you have mud on your own forehead?"
Assuming that all the children are perceptive, intelligent, truthful, and they answer simultaneously, what will happen?

Surprisingly, after the father asks the question for the k^{th} time all muddy children will say "yes". How come?

Muddy Children: The Underlying Assumptions

- **Common knowledge** that the father is truthful,
- that all the children hear the father,
- that all the children see each other,
- that none of them can see their own forehead,
- and that all the children are truthful and intelligent.

A Muddy Children Model

$(w_1 : m_a, m_b, m_c)$

$(w_3 : m_a, m_b, \neg m_c)$

$(w_2 : m_a, \neg m_b, m_c)$ $(w_4 : \neg m_a, m_b, m_c)$

$(w_5 : m_a, \neg m_b, \neg m_c)$ $(w_7 : \neg m_a, m_b, \neg m_c)$

$(w_6 : \neg m_a, \neg m_b, m_c)$

$(w_8 : \neg m_a, \neg m_b, \neg m_c)$

Figure: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

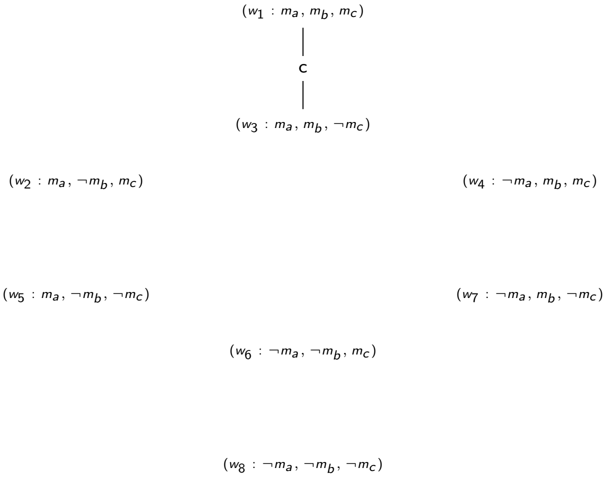


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

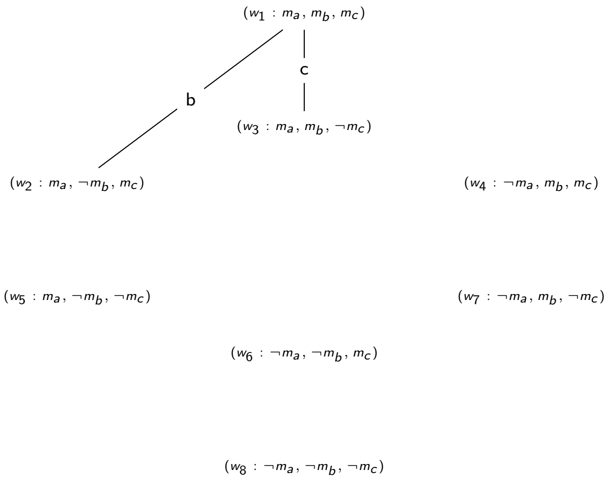


Figure: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

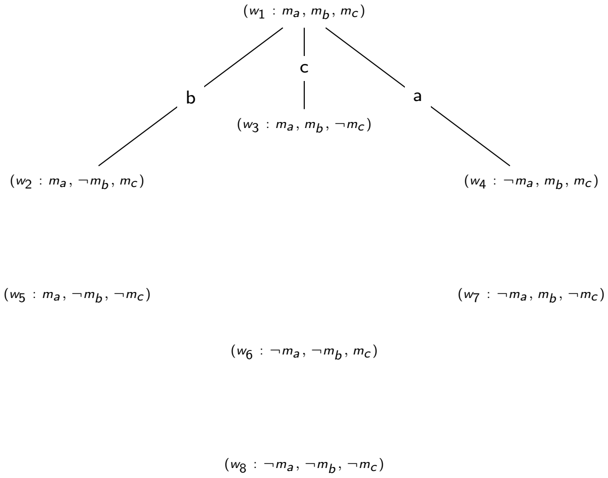


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

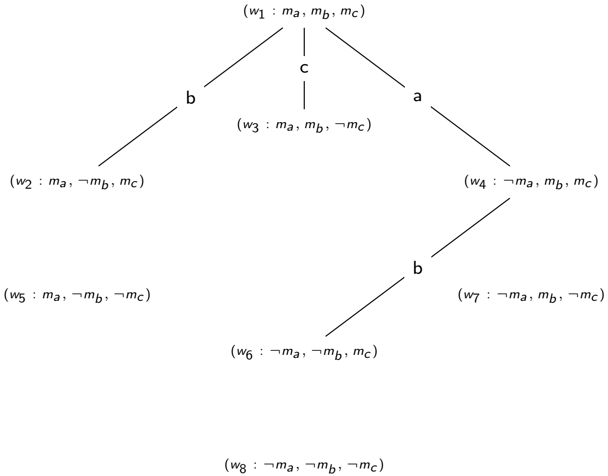


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

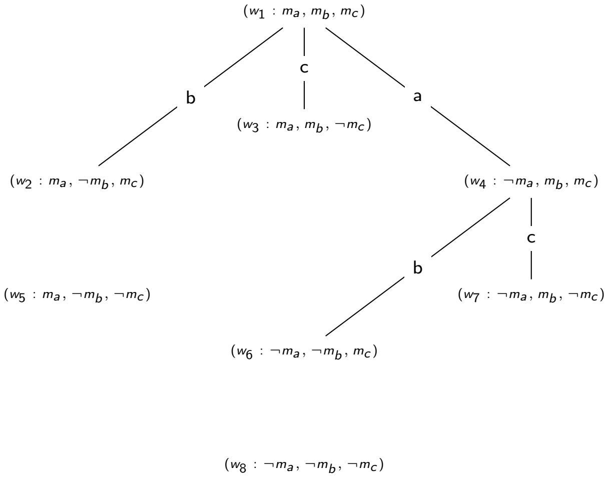


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

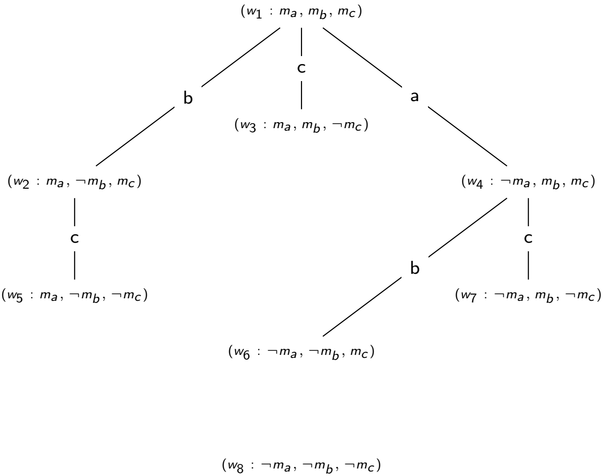


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

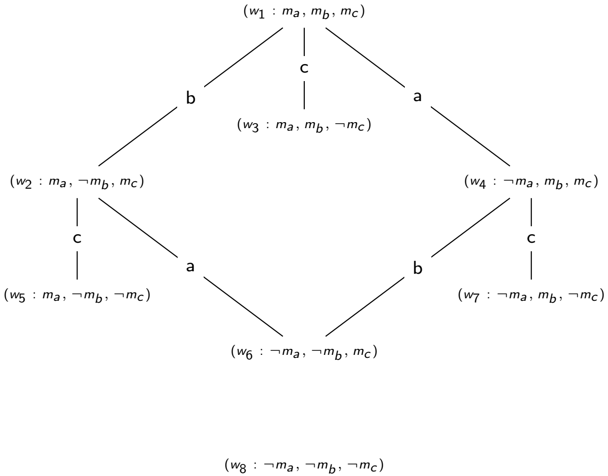


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

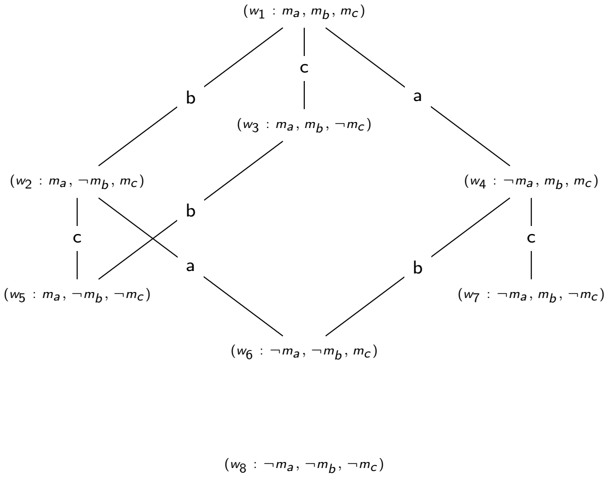


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

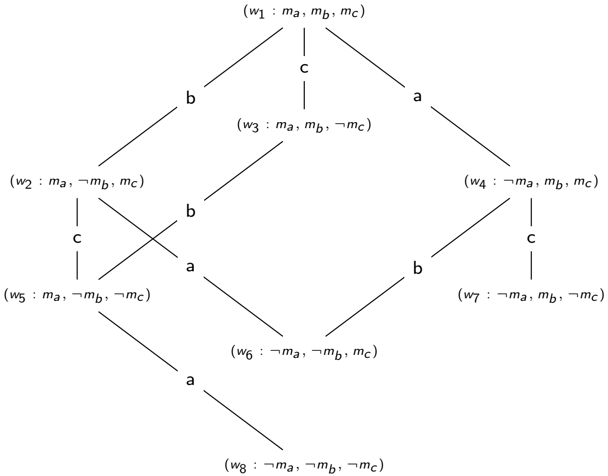


Figure: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

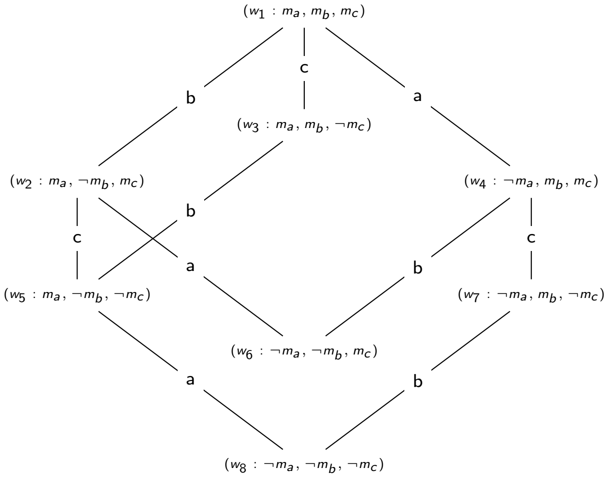


Figure: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

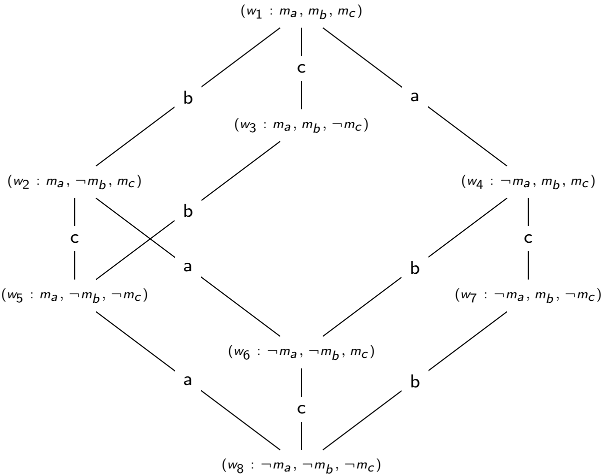


FIGURE: Initial epistemic model of the Muddy Children puzzle

A Muddy Children Model

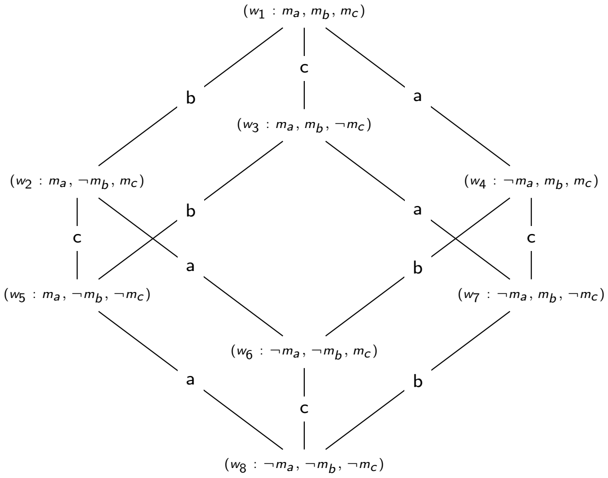
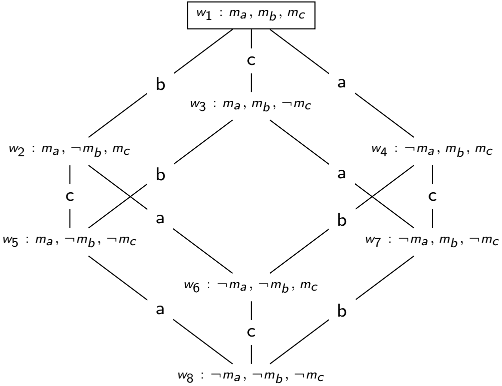


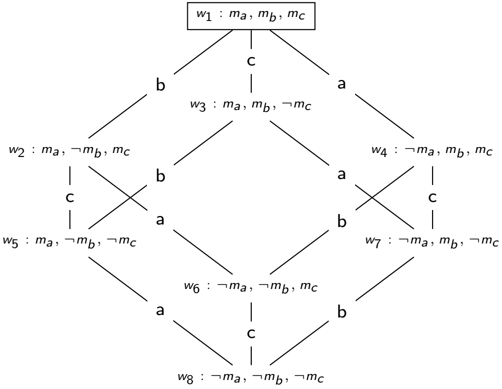
FIGURE: Initial epistemic model of the Muddy Children puzzle

Muddy Children Scenario Modelled in Epistemic Logic



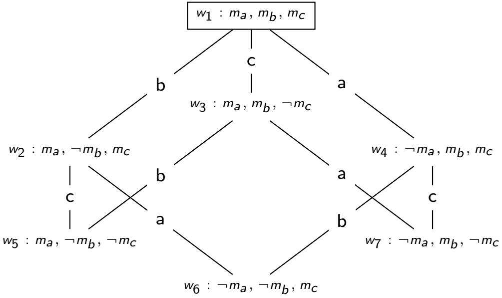
Before the announcement.

Muddy Children Scenario Modelled in Epistemic Logic



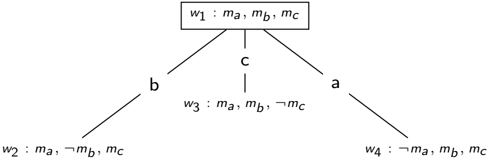
Father says:
"$m_a \vee m_b \vee m_c$"

Muddy Children Scenario Modelled in Epistemic Logic



– Does any of you know?
– No!

Muddy Children Scenario Modelled in Epistemic Logic



– Does any of you know?
– No!

Muddy Children Scenario Modelled in Epistemic Logic

$\boxed{w_1 : m_a, m_b, m_c}$

– Does any of you know?
– Yes!

Motivation

What are the properties of K?

How well does the K operator model knowledge?

Motivation

What are the properties of K?

How well does the K operator model knowledge?

We will try to answer this question
by looking at formulas about knowledge that are always true.

Definition

Given a model $M = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$, we say that:

- φ is **valid in** M, $M \models \varphi$, if $(M, s) \models \varphi$ for every state $s \in S$.

Formulas and Truth

Definition

Given a model $M = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$, we say that:

- φ is **valid in** M, $M \models \varphi$, if $(M, s) \models \varphi$ for every state $s \in S$.
- φ is **satisfiable in** M, if $(M, s) \models \varphi$ for some state $s \in S$.

Formulas and Truth

Definition

Given a model $M = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$, we say that:

- φ is **valid in** M, $M \models \varphi$, if $(M, s) \models \varphi$ for every state $s \in S$.
- φ is **satisfiable in** M, if $(M, s) \models \varphi$ for some state $s \in S$.
- φ is **valid**, $\models \varphi$, if φ is valid in all models.

DEFINITION

Given a model $M = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$, we say that:

- φ is **valid in** M, $M \models \varphi$, if $(M, s) \models \varphi$ for every state $s \in S$.
- φ is **satisfiable in** M, if $(M, s) \models \varphi$ for some state $s \in S$.
- φ is **valid**, $\models \varphi$, if φ is valid in all models.
- φ is **satisfiable**, if φ is satisfiable in some models.

Formulas and Truth

Definition

Given a model $M = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$, we say that:

- φ is **valid in** M, $M \models \varphi$, if $(M, s) \models \varphi$ for every state $s \in S$.
- φ is **satisfiable in** M, if $(M, s) \models \varphi$ for some state $s \in S$.
- φ is **valid**, $\models \varphi$, if φ is valid in all models.
- φ is **satisfiable**, if φ is satisfiable in some models.

A formula φ is valid (in M) if $\neg\varphi$ is not satisfiable (in M).

Distribution Axiom

Each agent knows all the logical consequences of her knowledge.

Distribution Axiom

Each agent knows all the logical consequences of her knowledge.

$\models (K_i\varphi \wedge K_i(\varphi \to \psi)) \to K_i\psi$

Distribution Axiom

Each agent knows all the logical consequences of her knowledge.

$\models (K_i\varphi \wedge K_i(\varphi \rightarrow \psi)) \rightarrow K_i\psi$

This axiom allows us to **distribute** the K_i operator over implication.

Knowledge Generalization Rule

Each agent knows all the formulas that are valid in a given model.

Knowledge Generalization Rule

Each agent knows all the formulas that are valid in a given model.

For all models M, if $M \models \varphi$ then $M \models K_i\varphi$.

Knowledge Generalization Rule

Each agent knows all the formulas that are valid in a given model.

For all models M, if $M \models \varphi$ then $M \models K_i\varphi$.

Note, that the above is not the same as: $\varphi \rightarrow K_i\varphi$

Knowledge Axiom (Truth Axiom)

Agents can only know facts.

Knowledge Axiom (Truth Axiom)

Agents can only know facts.

$\models K_i\varphi \to \varphi$

Knowledge Axiom (Truth Axiom)

Agents can only know facts.

$\models K_i\varphi \to \varphi$

This property distinguishes knowledge from **belief**.
It holds because the agents always consider the actual world possible.

Axioms of Positive and Negative Introspection

Agents know what they know and what they do not know.

Axioms of Positive and Negative Introspection

Agents know what they know and what they do not know.

$\models K_i\varphi \to K_iK_i\varphi$

$\models \neg K_i\varphi \to K_i\neg K_i\varphi$

The S5 System

1. $\models (K_i\varphi \wedge K_i(\varphi \rightarrow \psi)) \rightarrow K_i\psi$
2. if $M \models \varphi$ then $M \models K_i\varphi$
3. $\models K_i\varphi \rightarrow \varphi$
4. $\models K_i\varphi \rightarrow K_iK_i\varphi$
5. $\models \neg K_i\varphi \rightarrow K_i\neg K_i\varphi$

The S5 System

Theorem

For all formulas φ and ψ, all models M, where each possibility relation $\mathcal{K}_i$ is an equivalence, and all agents $i = 1, \ldots, n$, we have:

1. $M \models (K_i\varphi \wedge K_i(\varphi \rightarrow \psi)) \rightarrow K_i\psi$,
2. *if* $M \models \varphi$ *then* $M \models K_i\varphi$,
3. $M \models K_i\varphi \rightarrow \varphi$,
4. $M \models K_i\varphi \rightarrow K_iK_i\varphi$,
5. $M \models \neg K_i\varphi \rightarrow K_i\neg K_i\varphi$.

The S5 System and Equivalence Accessibility Relation

1. $\models (K_i\varphi \wedge K_i(\varphi \to \psi)) \to K_i\psi$
2. if $M \models \varphi$ then $M \models K_i\varphi$
3. $\models K_i\varphi \to \varphi$ **Reflexivity**
4. $\models K_i\varphi \to K_iK_i\varphi$ **Transitivity**
5. $\models \neg K_i\varphi \to K_i\neg K_i\varphi$ **Symmetricity and Transitivity**

Axiomatic Characterisations of Knowledge

properties of knowledge in terms of **some** valid formulas

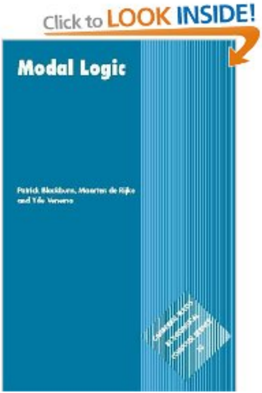
are there additional properties that are not consequences of those?

Is it possible to give a complete characterisation of the properties of knowledge?

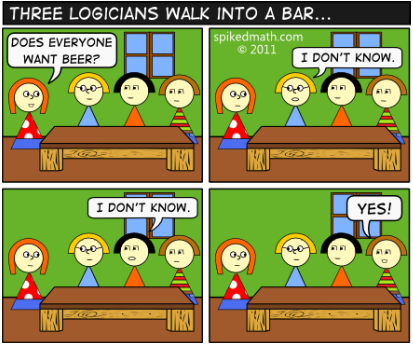
Outlook: Completeness Results

K	the class of all frames
K4	the class of transitive frames
T	the class of reflexive frames
B	the class of symmetric frames
KD	the class of right-unbounded frames
S4	the class of reflexive, transitive frames
S5	the class of frames whose relation is an equivalence relation
K4.3	the class of transitive frames with no branching to the right
S4.3	the class of reflexive, transitive frames with no branching to the right
KL	the class of finite transitive trees (*weak* completeness only)

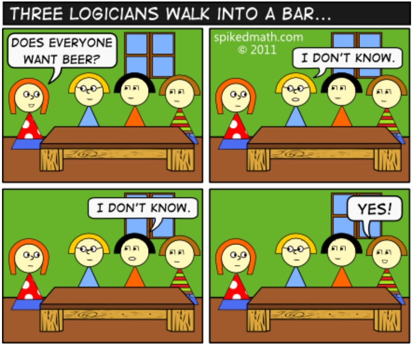
To Know More...


Click to LOOK INSIDE!
Modal Logic
Patrick Blackburn, Maarten de Rijke
and Yde Venema

Three logicians walk into a bar...

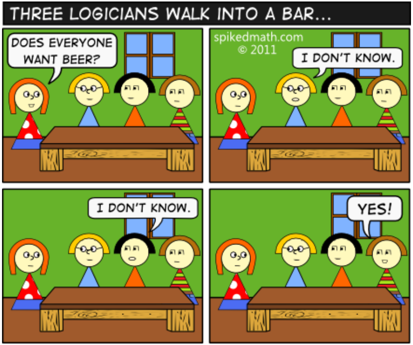


Three logicians walk into a bar...



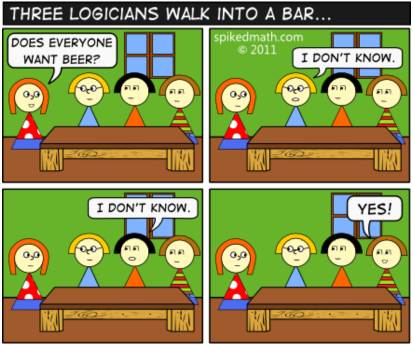
The three logicians are perfectly rational agents.

Three logicians walk into a bar...



The three logicians are perfectly rational agents. The question: "Does everyone want beer?" triggers the following instruction for agent i:

Three logicians walk into a bar...

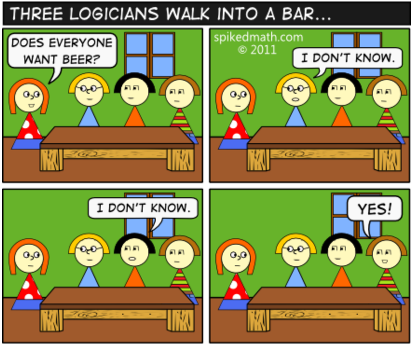


The three logicians are perfectly rational agents. The question: "Does everyone want beer?" triggers the following instruction for agent i:

- If $K_i(b_1 \wedge b_2 \wedge b_3)$,

  then i says "Yes".

Three logicians walk into a bar...



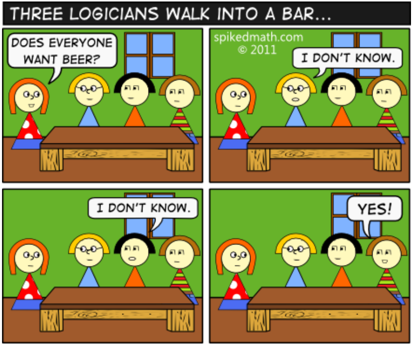
The three logicians are perfectly rational agents. The question: "Does everyone want beer?" triggers the following instruction for agent i:

- If $K_i(b_1 \wedge b_2 \wedge b_3)$,

  then i says "Yes".
- If $K_i\neg(b_1 \wedge b_2 \wedge b_3)$,

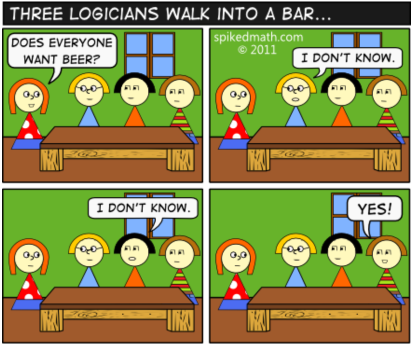
  then i says "No".

Three logicians walk into a bar...



The three logicians are perfectly rational agents. The question: "Does everyone want beer?" triggers the following instruction for agent i:

- If $K_i(b_1 \wedge b_2 \wedge b_3)$,
  then i says "Yes".
- If $K_i\neg(b_1 \wedge b_2 \wedge b_3)$,
  then i says "No".
- Otherwise,
  i says "I don't know".



The three logicians are perfectly rational agents. The question: "Does everyone want beer?" triggers the following instruction for agent i:

- If $K_i(b_1 \wedge b_2 \wedge b_3)$,

  then i says "Yes".
- If $K_i\neg(b_1 \wedge b_2 \wedge b_3)$,

  then i says "No".
- Otherwise,

  i says "I don't know".

Three Logicians: Updates

- i says "Yes": public announcement of $K_i(b_1 \wedge b_2 \wedge b_3)$.
- i says "No": public announcement of $K_i\neg(b_1 \wedge b_2 \wedge b_3)$.
- i says "I don't know": public announcement of $\neg K_i(b_1 \wedge b_2 \wedge b_3) \wedge \neg K_i\neg(b_1 \wedge b_2 \wedge b_3)$.

Three logicians: Possible-world model

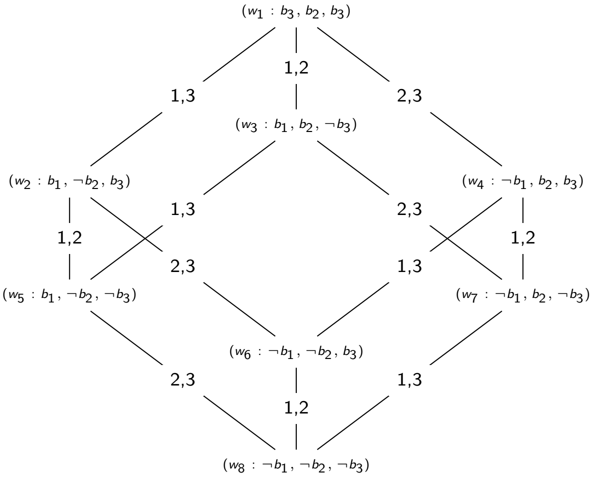


Figure: Initial epistemic model of the Three Logicians puzzle

Three logicians: Possible-world model

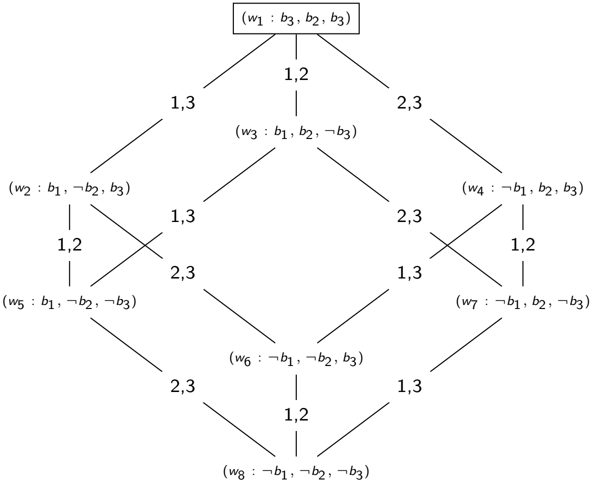


Figure: Public Announcement: $\neg K_1(b_1 \wedge b_2 \wedge b_3) \wedge \neg K_1 \neg(b_1 \wedge b_2 \wedge b_3)$

Three logicians: Possible-world model

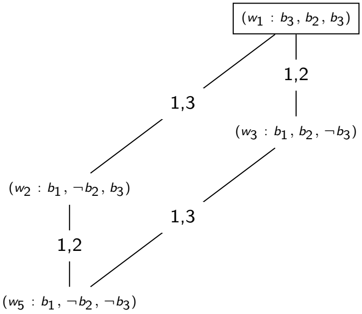


Figure: Public Announcement: $\neg K_1(b_1 \wedge b_2 \wedge b_3) \wedge \neg K_1 \neg(b_1 \wedge b_2 \wedge b_3)$

Three logicians: Possible-world model

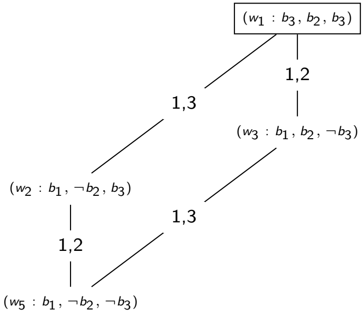


Figure: Public Announcement: $\neg K_2(b_1 \wedge b_2 \wedge b_3) \wedge \neg K_2 \neg (b_1 \wedge b_2 \wedge b_3)$

Three logicians: Possible-world model

$$\begin{tikzcd} \boxed{(w_1 : b_3, b_2, b_3)} \arrow[d, dash, "1{,}2" description] \\ (w_3 : b_1, b_2, \neg b_3) \end{tikzcd}$$

Figure: Public Announcement: $\neg K_2(b_1 \wedge b_2 \wedge b_3) \wedge \neg K_2 \neg (b_1 \wedge b_2 \wedge b_3)$

Three logicians: Possible-world model

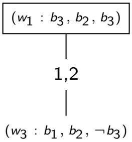


Figure: Public Announcement: $K_3(b_1 \wedge b_2 \wedge b_3)$, and the waitress knows!

Three logicians: Possible-world model

$\boxed{(w_1 : b_3, b_2, b_3)}$

Figure: Public Announcement: $K_3(b_1 \wedge b_2 \wedge b_3)$, and the waitress knows!

- PAL (Public Announcement Logic) was first proposed by Jan Plaza in 1989

Dynamic Epistemic Logics formalize informational changes:
the dynamics of knowledge/belief.

Dynamic Modalities

To express informational changes, dynamic epistemic logics use a new kind of propositional operators, called **dynamic modalities**:

$$[\alpha]\varphi,$$

where α is the name of some action involving communication.

To express informational changes, dynamic epistemic logics use a new kind of propositional operators, called **dynamic modalities**:

$$[\alpha]\varphi,$$

where α is the name of some action involving communication.

Such actions are called **epistemic actions** (as opposed to ontic actions) since they affect only the knowledge/beliefs of the agents.

Dynamic Modalities

To express informational changes, dynamic epistemic logics use a new kind of propositional operators, called **dynamic modalities**:

$$[\alpha]\varphi,$$

where α is the name of some action involving communication.

Such actions are called **epistemic actions** (as opposed to ontic actions) since they affect only the knowledge/beliefs of the agents.

The intended meaning of $[\alpha]\varphi$ is:
if action α is performed, then φ will become true.

Dynamic Modalities

To express informational changes, dynamic epistemic logics use a new kind of propositional operators, called **dynamic modalities**:

$$[\alpha]\varphi,$$

where α is the name of some action involving communication.

Such actions are called **epistemic actions** (as opposed to ontic actions) since they affect only the knowledge/beliefs of the agents.

The intended meaning of $[\alpha]\varphi$ is:
if action α is performed, then φ will become true.

Note: $[\alpha]\varphi$ is by definition true in the worlds in which α can't be performed.

DEFINITION (SYNTAX)

Φ is a set of propositions, with $p \in \Phi$, and $\mathcal{A} = \{1, \ldots, n\}$ is a set of agents, and $G \subseteq \mathcal{A}$.

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid [!\varphi]\varphi$$

where $\top$ abbreviates a tautology and $i \in \mathcal{A}$ is the name of some agent.

Language of Public Announcement Logic

Definition (Syntax)

Φ is a set of propositions, with $p \in \Phi$, and $\mathcal{A} = \{1, \ldots, n\}$ is a set of agents, and $G \subseteq \mathcal{A}$.

$$\varphi := \top \mid p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid [!\varphi]\varphi$$

where $\top$ abbreviates a tautology and $i \in \mathcal{A}$ is the name of some agent.

As before, these formulas are interpreted in possible world models.

Public Announcement as joint update

How can we model the effect of a public announcement?

Public Announcement as joint update

How can we model the effect of a public announcement?

As we saw in many examples, this can be done by **deleting worlds**.

LEARNING = ELIMINATING POSSIBILITIES

Public Announcement as joint update

How can we model the effect of a public announcement?

As we saw in many examples, this can be done by **deleting worlds**.

LEARNING = ELIMINATING POSSIBILITIES

From now on, we denote by $!\varphi$ the operation of deleting the non-φ worlds, and call it **public announcement with** φ, or **joint update with** φ.

DEFINITION

Let $M = (S, \pi, \mathcal{K}_1, \ldots, \mathcal{K}_n)$ and let $\|\varphi\|_M$ stand for the set of worlds in M that makes $\|\varphi\|$ true.

$(M, s) \models p$	iff	$\pi(s)(p) = 1$
$(M, s) \models \neg\varphi$	iff	$(M, s) \not\models \varphi$
$(M, s) \models \varphi \wedge \psi$	iff	$(M, s) \models \varphi$ and $(M, s) \models \psi$
$(M, s) \models \varphi \vee \psi$	iff	$(M, s) \models \varphi$ or $(M, s) \models \psi$
$(M, s) \models K_i\varphi$	iff	for all v with $(s, v) \in \mathcal{K}_i$, $(M, v) \models \varphi$
$(M, s) \models [!\varphi]\psi$	**iff**	**if** $(M, s) \models \varphi$ **then** $(M\|\varphi, s) \models \psi$

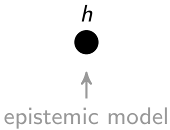
where $M|\varphi = (S', \pi', \mathcal{K}'_1, \ldots, \mathcal{K}'_n)$ is defined as follows:

- $S' := \|\varphi\|_M$
- $\pi' := \pi \cap \|\varphi\|_M$
- for each $i \in \{1, \ldots, n\}$, $\mathcal{K}'_i := \mathcal{K}_i \cap (\|\varphi\|_M \times \|\varphi\|_M)$

DEL BY EXAMPLE: A HIDDEN COIN TOSS

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



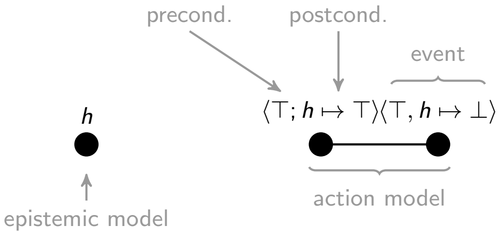
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL by example: A hidden coin toss

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



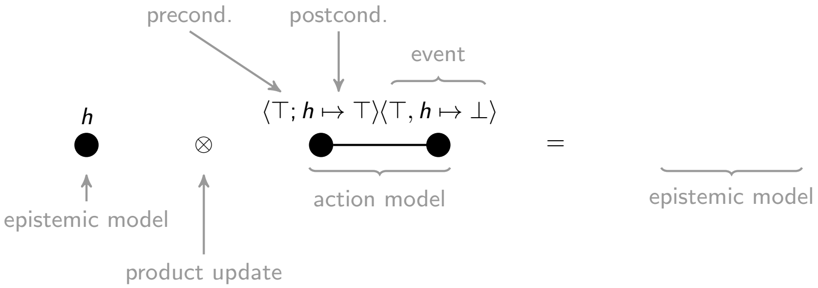
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL by example: A hidden coin toss

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



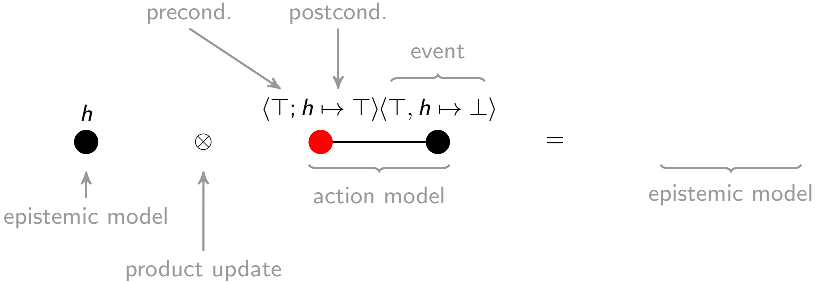
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL by example: A hidden coin toss

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



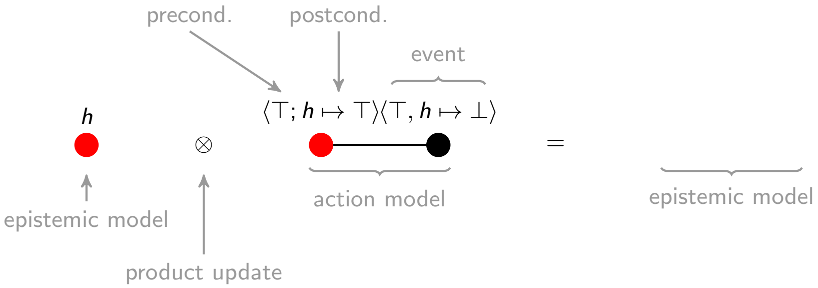
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL BY EXAMPLE: A HIDDEN COIN TOSS

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



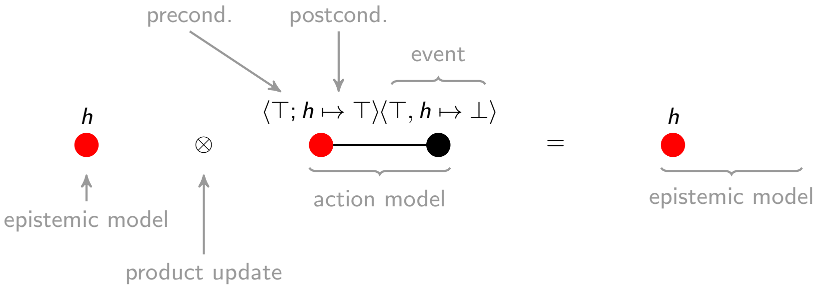
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL BY EXAMPLE: A HIDDEN COIN TOSS

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



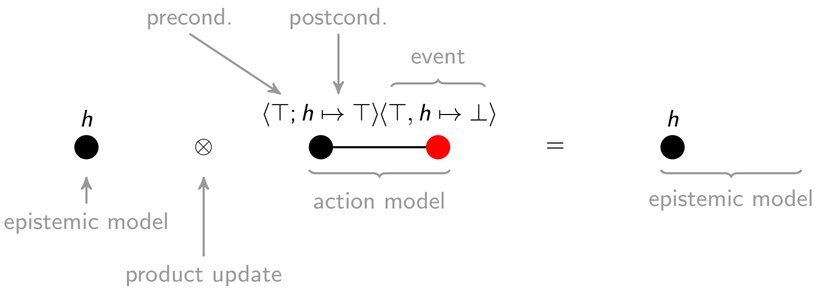
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL by example: A hidden coin toss

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



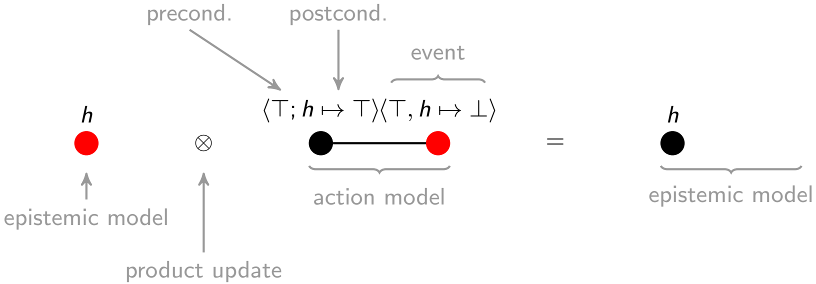
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL BY EXAMPLE: A HIDDEN COIN TOSS

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



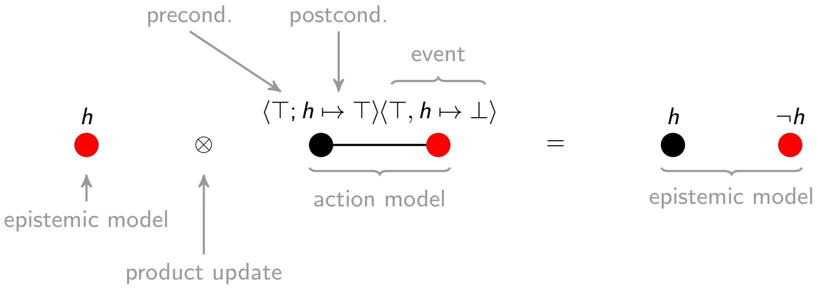
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL by example: A hidden coin toss

We use the action models of DEL with postconditions (ontic actions).

h:= "the coin faces heads up"



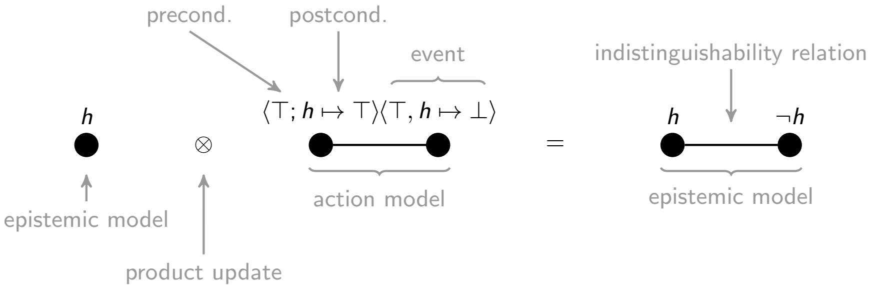
Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

DEL BY EXAMPLE: A HIDDEN COIN TOSS

We use the action models of DEL with postconditions (ontic actions).

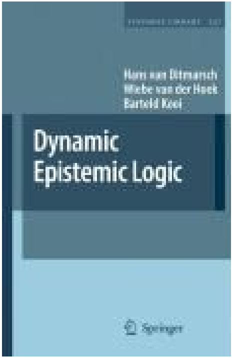
h:= "the coin faces heads up"



Baltag, Moss, and Solecki. The logic of public announcements, common knowledge, and private suspicions. TARK 1998.

van Ditmarsch and Kooi. Semantic results for ontic and epistemic change. LOFT 2008.

To Know More...


Hans van Ditmarsch
Wiebe van der Hoek
Barteld Kooi
Dynamic
Epistemic Logic
Springer

END OF PART 1

Group Knowledge: Everybody Knows

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Group Knowledge: Everybody Knows

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $E_G\varphi$.

Group Knowledge: Everybody Knows

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $E_G\varphi$.

Semantics $(M, s) \models E_G\varphi$ iff $(M, s) \models K_i\varphi$ for every $i \in G$.

Example
Everybody in this class knows that logic is useful.

Group Knowledge: Everybody Knows

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $E_G\varphi$.

Semantics $(M, s) \models E_G\varphi$ iff $(M, s) \models K_i\varphi$ for every $i \in G$.

Example
Everybody in this class knows that logic is useful.

Definition
We define inductively:

$$E_G^0\varphi := \varphi$$

$$E_G^{k+1}\varphi := E_G E_G^k\varphi$$

Group Knowledge: Common Knowledge

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Group Knowledge: Common Knowledge

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $C_G\varphi$.

Group Knowledge: Common Knowledge

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $C_G\varphi$.

Semantics $(M, s) \models C_G\varphi$ iff $(M, s) \models E_G^k\varphi$ for every $k = 1, 2, \ldots$

Group Knowledge: Common Knowledge

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $C_G\varphi$.

Semantics $(M, s) \models C_G\varphi$ iff $(M, s) \models E_G^k\varphi$ for every $k = 1, 2, \ldots$

Example

It is common knowledge in Denmark that to drive on the right side of the road.

Group Knowledge: Distributed Knowledge

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Group Knowledge: Distributed Knowledge

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $D_G\varphi$.

Group Knowledge: Distributed Knowledge

Let $G \subseteq \mathcal{A}$ be a non-empty subset of agents.

Syntax If φ is a formula, then so is $D_G\varphi$.

Semantics $(M, s) \models D_G\varphi$ iff $(M, t) \models \varphi$ for all t such that $(s, t) \in \bigcap_{i \in G} \mathcal{K}_i$.

Semantics of Epistemic Logic with Group Operators

Definition

$(M,s) \models p$	iff	$\pi(s)(p) = 1$
$(M,s) \models \neg\varphi$	iff	$(M,s) \not\models \varphi$
$(M,s) \models \varphi \wedge \psi$	iff	$(M,s) \models \varphi$ and $(M,s) \models \psi$
$(M,s) \models K_i\varphi$	iff	for all v with $(s,v) \in \mathcal{K}_i$, $(M,v) \models \varphi$
$(M,s) \models E_G\varphi$	iff	$(M,s) \models K_i\varphi$ for every $i \in G$
$(M,s) \models C_G\varphi$	iff	$(M,s) \models E_G^k\varphi$ for every $k = 1, 2, \ldots$
$(M,s) \models D_G\varphi$	iff	$(M,t) \models \varphi$ for all t such that $(s,t) \in \bigcap_{i \in G} \mathcal{K}_i$