Modal Logic
Dynamic Epistemic Logic

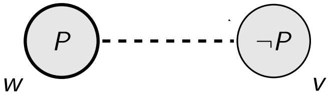
Eric Pacuit

University of Maryland, College Park
ai.stanford.edu/∼epacuit

April 24, 2012

Modeling Information Change

Models of Hard and Soft Information



Epistemic Model: $\mathcal{M} = \langle W, \{\sim_i\}_{i \in \mathcal{A}}, V \rangle$

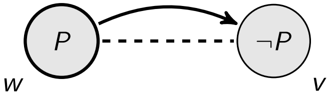
- $w \sim_i v$ means i cannot rule out v according to her information.

Language: $\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid K_i\varphi$

Truth:

- $\mathcal{M}, w \models p$ iff $w \in V(p)$ (p an atomic proposition)
- Boolean connectives as usual
- $\mathcal{M}, w \models K_i\varphi$ iff for all $v \in W$, if $w \sim_i v$ then $\mathcal{M}, v \models \varphi$

Models of Hard and Soft Information



Epistemic-Plausibility Model: $\mathcal{M} = \langle W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V\rangle$

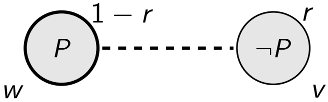
- $w \preceq_i v$ means v is at least as plausibility as w for agent i.

Language: $\varphi := p \mid \neg\varphi \mid \varphi\wedge\psi \mid K_i\varphi \mid B^{\varphi}\psi \mid [\preceq_i]\varphi$

Truth:

- $[\![\varphi]\!]_{\mathcal{M}} = \{w \mid \mathcal{M}, w \models \varphi\}$
- $\mathcal{M}, w \models B_i^{\varphi}\psi$ iff for all $v \in Min_{\preceq_i}([\![\varphi]\!]_{\mathcal{M}} \cap [w]_i)$, $\mathcal{M}, v \models \psi$
- $\mathcal{M}, w \models [\preceq_i]\varphi$ iff for all $v \in W$, if $v \preceq_i w$ then $\mathcal{M}, v \models \varphi$

Models of Hard and Soft Information



Epistemic-Plausibility Model: $\mathcal{M} = \langle W, \{\sim_i\}_{i\in\mathcal{A}}, \{\pi_i\}_{i\in\mathcal{A}}, V\rangle$

- $\pi_i : W \to [0,1]$ is a probability measure

Language: $\varphi := p \mid \neg\varphi \mid \varphi\wedge\psi \mid K_i\varphi \mid B^p\psi$

Truth:

- $[\![\varphi]\!]_{\mathcal{M}} = \{w \mid \mathcal{M}, w \models \varphi\}$
- $\mathcal{M}, w \models B^p\varphi$ iff $\pi_i([\![\varphi]\!]_{\mathcal{M}} \mid [w]_i) = \frac{\pi_i([\![\varphi]\!]_{\mathcal{M}}\cap[w]_i)}{\pi_i([w]_i)} \geq p$, $\mathcal{M}, v \models \psi$
- $\mathcal{M}, w \models K_i\varphi$ iff for all $v \in W$, if $w \sim_i v$ then $\mathcal{M}, v \models \varphi$

Models of Hard and Soft Information

- *Describing* what the agents know and believe rather than *defining* the agents' knowledge (and beliefs) in terms or more primitive notions

Models of Hard and Soft Information

- *Describing* what the agents know and believe rather than *defining* the agents' knowledge (and beliefs) in terms or more primitive notions

- Many group notions (common knowledge, distributed knowledge, common belief, common p-belief)

Models of Hard and Soft Information

- *Describing* what the agents know and believe rather than *defining* the agents' knowledge (and beliefs) in terms or more primitive notions
- Many group notions (common knowledge, distributed knowledge, common belief, common p-belief)
- Other types of informational attitudes (robust beliefs, strong beliefs, certainty, awareness, etc.)

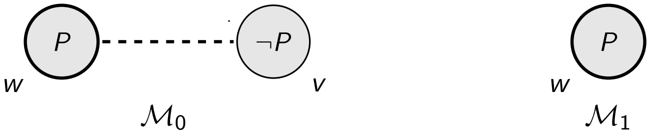
Models of Hard and Soft Information

- *Describing* what the agents know and believe rather than *defining* the agents' knowledge (and beliefs) in terms or more primitive notions
- Many group notions (common knowledge, distributed knowledge, common belief, common p-belief)
- Other types of informational attitudes (robust beliefs, strong beliefs, certainty, awareness, etc.)
- Represents the agents' information at a fixed moment in time

Models of Hard and Soft Information

- *Describing* what the agents know and believe rather than *defining* the agents' knowledge (and beliefs) in terms or more primitive notions
- Many group notions (common knowledge, distributed knowledge, common belief, common *p*-belief)
- Other types of informational attitudes (robust beliefs, strong beliefs, certainty, awareness, etc.)
- Represents the agents' information at a fixed moment in time

Finding out that *p* is true



Modeling Information Change: Two Methodologies

1. "Change-based modeling": describe the effect a *learning experience* has on a model

2. "Explicit-temporal modeling": explicitly describe different moments *in the model*

Modeling Information Change: Two Methodologies

1. "Change-based modeling": describe the effect a *learning experience* has on a model

2. "Explicit-temporal modeling": explicitly describe different moments *in the model*

Example

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

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

Is this procedure correct?

Example

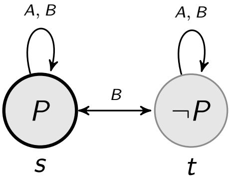
Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

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

Is this procedure correct? Yes, if

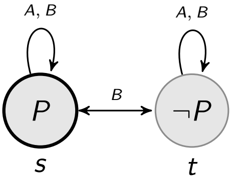
1. Ann knows about the talk.
2. Bob knows about the talk.
3. Ann knows that Bob knows about the talk.
4. Bob *does not* know that Ann knows that he knows about the talk.
5. *And nothing else.*

Example



P means "The talk is at 2PM".

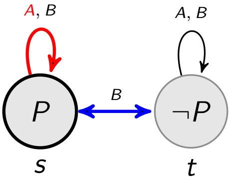
Example



P means "The talk is at 2PM".

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

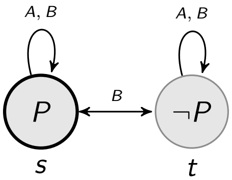
Example



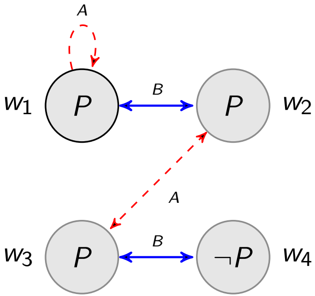
P means “The talk is at 2PM”.

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

Example



Prior Model



Posterior Model

Digression on Belief Change

Digression on Belief Change

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Digression on Belief Change

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Thus, the agent believes:

- q The bird caught in the trap is white.

Digression on Belief Change

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Thus, the agent believes:

- q The bird caught in the trap is white.

Now suppose the rational agent—for example, You—learn that the bird caught in the trap is black ($\neg q$).

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Thus, the agent believes:

- q The bird caught in the trap is white.

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Thus, the agent believes:

- q The bird caught in the trap is white.

Question: How should the agent incorporate $\neg q$ into his belief state to obtain a consistent belief state?

Digression on Belief Change, I

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Thus, the agent believes:

- q The bird caught in the trap is white.

Question: How should the agent incorporate $\neg q$ into his belief state to obtain a consistent belief state?

Problem: Logical considerations alone are insufficient to answer this question! Why??

Digression on Belief Change, II

Consider the following beliefs of a rational agent:

- p_1 All Europeans swans are white.
- p_2 The bird caught in the trap is a swan.
- p_3 The bird caught in the trap comes from Sweden.
- p_4 Sweden is part of Europe.

Thus, the agent believes:

- q The bird caught in the trap is white.

Question: How should the agent incorporate $\neg q$ into his belief state to obtain a consistent belief state?
Problem: Logical considerations alone are insufficient to answer this question!
There are several logically distinct ways to incorporate $\neg q$!

Digression on Belief Change, II

What extralogical factors serve to determine what beliefs to give up and what beliefs to retain?

Digression on Belief Change, III

Belief revision is a matter of choice, and the choices are to be made in such a way that:

1. The resulting theory squares with the experience;
2. It is simple; and
3. The choices disturb the original theory as little as possible.

Digression on Belief Change, III

Belief revision is a matter of choice, and the choices are to be made in such a way that:

1. The resulting theory squares with the experience;
2. It is simple; and
3. The choices disturb the original theory as little as possible.

Research has relied on the following related guiding ideas:

1. When accepting a new piece of information, an agent should aim at a minimal change of his old beliefs.
2. If there are different ways to effect a belief change, the agent should give up those beliefs which are least entrenched.

Digression: Belief Revision

A.P. Pedersen and H. Arló-Costa. *"Belief Revision."*. In Continuum Companion to Philosophical Logic. Continuum Press, 2011..

Hans Rott. *Change, Choice and Inference: A Study of Belief Revision and Nonmonotonic Reasoning*. Oxford University Press, 2001.

Digression: AGM Postulates

AGM 1: $K * \varphi$ is deductively closed

AGM 2: $\varphi \in K * \varphi$

AGM 3: $K * \varphi \subseteq Cn(K \cup \{\varphi\})$

AGM 4: If $\neg\varphi \notin K$ then $K * \varphi = Cn(K \cup \{\varphi\})$

AGM 5: $K * \varphi$ is inconsistent only if φ is inconsistent

AGM 6: If φ and ψ are logically equivalent then $K * \varphi = K * \psi$

AGM 7: $K * (\varphi \wedge \psi) \subseteq Cn(K * \varphi \cup \{\psi\})$

AGM 8 if $\neg\psi \notin K * \varphi$ then $Cn(K * \varphi \cup \{\psi\}) \subseteq K * (\varphi \wedge \psi)$

Digression: Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

Digression: Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

A subtle difference:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Digression: Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

A subtle difference:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Eg., Either the room is painted white or Queen's day (Koninginnedag) is on Sunday.

Digression: Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

A subtle difference:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Eg., Either the room is painted white or Queen's day (Koninginnedag) is on Sunday.

Complete vs. incomplete belief sets:

$K = Cn(\{p \vee q\})$ vs. $K = Cn(\{p \vee q, p, q\})$

Digression: Revision vs. Update

Suppose φ is some incoming information that should be incorporated into the agents beliefs (represented by a theory T).

A subtle difference:

- If φ describes facts about the current state of affairs
- If φ describes facts that have possible become true only after the original beliefs were formed.

Eg., Either the room is painted white or Queen's day (Koninginnedag) is on Sunday.

Complete vs. incomplete belief sets:

$K = Cn(\{p \vee q\})$ vs. $K = Cn(\{p \vee q, p, q\})$

Revising by $\neg p$ ($K * \neg p$) vs. Updating by $\neg p$ ($K \diamond \neg p$)

H. Katsuno and A. O. Mendelzon. *Propositional knowledge base revision and minimal change*. Artificial Intelligence, 52, pp. 263 - 294 (1991).

Aspects of Informative Events

1. The agents' *observational* powers.

   Agents may perceive the same event differently and this can be described in terms of what agents do or do not observe. Examples range from *public announcements* where everyone witnesses the same event to private communications between two or more agents with the other agents not even being aware that an event took place.

Aspects of Informative Events

1. The agents' *observational* powers.

2. The *type* of change triggered by the event.

   Agents may differ in precisely how they incorporate new information into their epistemic states. These differences are based, in part, on the agents' perception of the *source* of the information. For example, an agent may consider a particular source of information *infallible* (not allowing for the possibility that the source is mistaken) or merely *trustworthy* (accepting the information as reliable though allowing for the possibility of a mistake).

Aspects of Informative Events

1. The agents' *observational* powers.
2. The *type* of change triggered by the event.
3. The underlying *protocol* specifying which events (observations, messages, actions) are available (or permitted) at any given moment.

   This is intended to represent the rules or conventions that govern many of our social interactions. For example, in a conversation, it is typically not polite to "blurt everything out at the beginning", as we must speak in small chunks. Other natural conversational protocol rules include "do not repeat yourself", "let others speak in turn", and "be honest". Imposing such rules *restricts* the legitimate sequences of possible statements or events.

Aspects of Informative Events

1. The agents' *observational* powers.

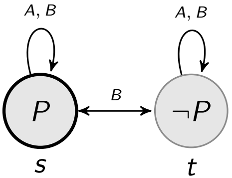
2. The *type* of change triggered by the event.

3. The underlying *protocol* specifying which events (observations, messages, actions) are available (or permitted) at any given moment.

Aspects of Informative Events

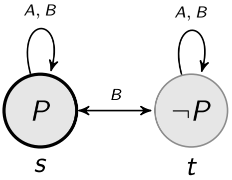
1. The agents' *observational* powers.
2. The *type* of change triggered by the event.
3. The underlying *protocol* specifying which events (observations, messages, actions) are available (or permitted) at any given moment.

Dynamic Events: Public Announcement



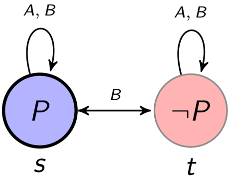
P means "The talk is at 2PM".

Dynamic Events: Public Announcement



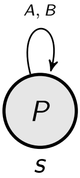
What happens if Ann publicly announces P?

Dynamic Events: Public Announcement



What happens if Ann publicly announces P?

Dynamic Events: Public Announcement



What happens if Ann publicly announces P? $s \models CP$

Public Announcement Logic

J. Plaza. *Logics of Public Communications*. 1989.

J. Gerbrandy. *Bisimulations on Planet Kripke*. 1999.

J. van Benthem. *One is a lonely number*. 2002.

Public Announcement Logic

The Public Announcement Language is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid C\varphi \mid [\psi]\varphi$$

where $p \in \mathrm{At}$ and $i \in \mathcal{A}$.

Public Announcement Logic

The Public Announcement Language is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid C\varphi \mid [\psi]\varphi$$

where $p \in \mathrm{At}$ and $i \in \mathcal{A}$.

- $[\psi]\varphi$ is intended to mean "After publicly announcing ψ, φ is true".

Public Announcement Logic

The Public Announcement Language is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid C\varphi \mid [\psi]\varphi$$

where $p \in \mathrm{At}$ and $i \in \mathcal{A}$.

- $[P]K_iP$: "After publicly announcing P, agent i knows P"

Public Announcement Logic

The Public Announcement Language is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid C\varphi \mid [\psi]\varphi$$

where $p \in \mathrm{At}$ and $i \in \mathcal{A}$.

- $[\neg K_i P]CP$: "After announcing that agent i does not know P, then P is common knowledge"

Public Announcement Logic

The Public Announcement Language is generated by the following grammar:

$$p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K_i\varphi \mid C\varphi \mid [\psi]\varphi$$

where $p \in \mathrm{At}$ and $i \in \mathcal{A}$.

- $[\neg K_i P]K_i P$: "after announcing i does not know P, then i knows P. "

Public Announcement Logic

Suppose $\mathcal{M} = \langle W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V\rangle$ is a multi-agent Kripke Model

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

where $\mathcal{M}|_\psi = \langle W', \{\sim_i'\}_{i\in\mathcal{A}}, \{\preceq_i'\}_{i\in\mathcal{A}}, V'\rangle$ with

- $W' = W \cap \{w \mid \mathcal{M}, w \models \psi\}$
- For each i, $\sim_i' = \sim_i \cap (W' \times W')$
- For each i, $\preceq_i' = \preceq_i \cap (W' \times W')$
- for all $p \in \text{At}$, $V'(p) = V(p) \cap W'$

Public Announcement Logic

$$[\psi]p \quad \leftrightarrow \quad (\psi \rightarrow p)$$

Public Announcement Logic

$$
\begin{aligned}
[\psi]p &\leftrightarrow (\psi \rightarrow p) \\
[\psi]\neg\varphi &\leftrightarrow (\psi \rightarrow \neg[\psi]\varphi)
\end{aligned}
$$

Public Announcement Logic

$$
\begin{aligned}
[\psi]p &\leftrightarrow (\psi \to p) \\
[\psi]\neg\varphi &\leftrightarrow (\psi \to \neg[\psi]\varphi) \\
[\psi](\varphi \wedge \chi) &\leftrightarrow ([\psi]\varphi \wedge [\psi]\chi)
\end{aligned}
$$

Public Announcement Logic

$$
\begin{array}{rcl}
[\psi]p & \leftrightarrow & (\psi \rightarrow p) \\
[\psi]\neg\varphi & \leftrightarrow & (\psi \rightarrow \neg[\psi]\varphi) \\
[\psi](\varphi \wedge \chi) & \leftrightarrow & ([\psi]\varphi \wedge [\psi]\chi) \\
[\psi]K_i\varphi & \leftrightarrow & (\psi \rightarrow K_i(\psi \rightarrow [\psi]\varphi))
\end{array}
$$

Public Announcement Logic

$$
\begin{array}{rcl}
[\psi]p & \leftrightarrow & (\psi \rightarrow p) \\
[\psi]\neg\varphi & \leftrightarrow & (\psi \rightarrow \neg[\psi]\varphi) \\
[\psi](\varphi \wedge \chi) & \leftrightarrow & ([\psi]\varphi \wedge [\psi]\chi) \\
[\psi]K_i\varphi & \leftrightarrow & (\psi \rightarrow K_i(\psi \rightarrow [\psi]\varphi))
\end{array}
$$

Public Announcement Logic

$$
\begin{array}{rcl}
[\psi]p & \leftrightarrow & (\psi \rightarrow p) \\
[\psi]\neg\varphi & \leftrightarrow & (\psi \rightarrow \neg[\psi]\varphi) \\
[\psi](\varphi \wedge \chi) & \leftrightarrow & ([\psi]\varphi \wedge [\psi]\chi) \\
[\psi]K_i\varphi & \leftrightarrow & (\psi \rightarrow K_i(\psi \rightarrow [\psi]\varphi))
\end{array}
$$

Theorem Every formula of Public Announcement Logic is equivalent to a formula of Epistemic Logic.

Public Announcement Logic

$$
\begin{array}{rcl}
[\psi]p & \leftrightarrow & (\psi \rightarrow p) \\
[\psi]\neg\varphi & \leftrightarrow & (\psi \rightarrow \neg[\psi]\varphi) \\
[\psi](\varphi \wedge \chi) & \leftrightarrow & ([\psi]\varphi \wedge [\psi]\chi) \\
[\psi]K_i\varphi & \leftrightarrow & (\psi \rightarrow K_i(\psi \rightarrow [\psi]\varphi))
\end{array}
$$

The situation is more complicated with common knowledge.

J. van Benthem, J. van Eijk, B. Kooi. *Logics of Communication and Change*. 2006.

- $[q]Kq$

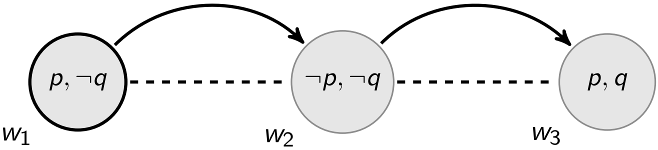
- $[q]Kq$
- $Kp \to [q]Kp$

- $[q]Kq$
- $Kp \to [q]Kp$
- $B\varphi \to [\psi]B\varphi$

- $[q]Kq$

- $Kp \rightarrow [q]Kp$

- $B\varphi \rightarrow [\psi]B\varphi$



- $[\varphi]\varphi$

Public Announcement vs. Conditional Belief

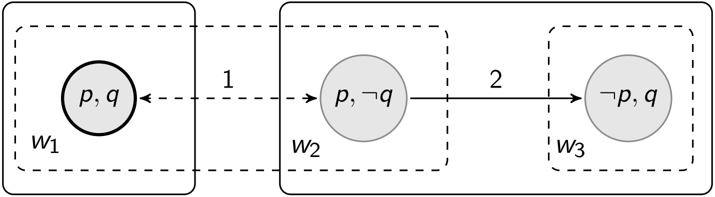
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different?

Public Announcement vs. Conditional Belief

Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!

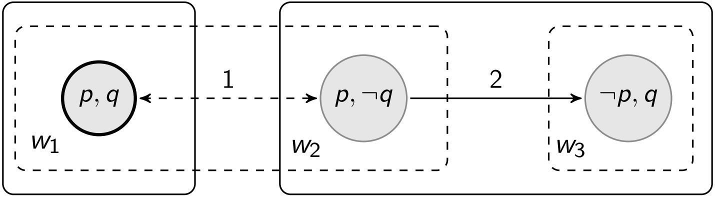
Public Announcement vs. Conditional Belief

Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



Public Announcement vs. Conditional Belief

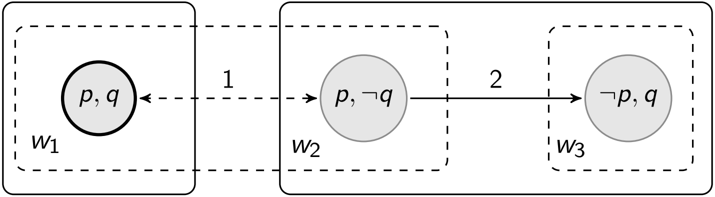
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



- $w_1 \models B_1 B_2 q$

Public Announcement vs. Conditional Belief

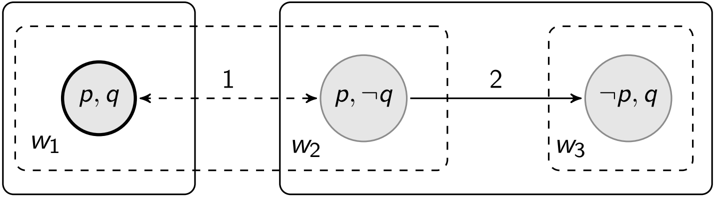
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



- $w_1 \models B_1 B_2 q$
- $w_1 \models B_1^p B_2 q$

Public Announcement vs. Conditional Belief

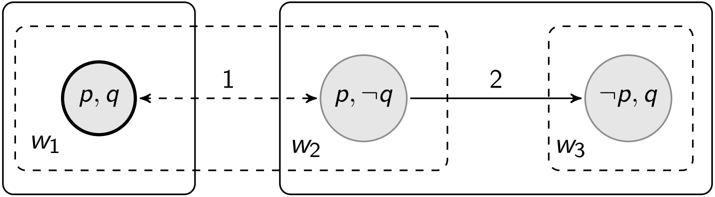
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



- $w_1 \models B_1 B_2 q$
- $w_1 \models B_1^p B_2 q$
- $w_1 \models [p]\neg B_1 B_2 q$

Public Announcement vs. Conditional Belief

Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



- $w_1 \models B_1 B_2 q$
- $w_1 \models B_1^p B_2 q$
- $w_1 \models [p]\neg B_1 B_2 q$
- More generally, $B_i^p(p \wedge \neg K_i p)$ is satisfiable but $[p]B_i(p \wedge \neg K_i p)$ is not.

The Logic of Public Observation

- $[!\psi]K\varphi \leftrightarrow (\psi \rightarrow K(\psi \rightarrow [!\psi]\varphi))$

The Logic of Public Observation

- $[!\psi]K\varphi \leftrightarrow (\psi \to K(\psi \to [!\psi]\varphi))$

- $[\varphi][\preceq_i]\psi \leftrightarrow (\varphi \to [\preceq_i](\varphi \to [\varphi]\psi))$

The Logic of Public Observation

- $[!\psi]K\varphi \leftrightarrow (\psi \to K(\psi \to [!\psi]\varphi))$
- $[\varphi][\preceq_i]\psi \leftrightarrow (\varphi \to [\preceq_i](\varphi \to [\varphi]\psi))$
- **Belief**: $[!\psi]B\varphi \not\leftrightarrow (\psi \to B(\psi \to [!\psi]\varphi))$

The Logic of Public Observation

- $[!\psi]K\varphi \leftrightarrow (\psi \rightarrow K(\psi \rightarrow [!\psi]\varphi))$
- $[\varphi][\preceq_i]\psi \leftrightarrow (\varphi \rightarrow [\preceq_i](\varphi \rightarrow [\varphi]\psi))$
- **Belief**: $[!\psi]B\varphi \not\leftrightarrow (\psi \rightarrow B(\psi \rightarrow [!\psi]\varphi))$

  $[\varphi]B\psi \leftrightarrow (\varphi \rightarrow B^{\varphi}[\varphi]\psi)$

The Logic of Public Observation

- $[!\psi]K\varphi \leftrightarrow (\psi \to K(\psi \to [!\psi]\varphi))$

- $[\varphi][\preceq_i]\psi \leftrightarrow (\varphi \to [\preceq_i](\varphi \to [\varphi]\psi))$

- **Belief**: $[!\psi]B\varphi \not\leftrightarrow (\psi \to B(\psi \to [!\psi]\varphi))$

  $[\varphi]B\psi \leftrightarrow (\varphi \to B^{\varphi}[\varphi]\psi)$
  $[\varphi]B^{\alpha}\psi \leftrightarrow (\varphi \to B^{\varphi \wedge [\varphi]\alpha}[\varphi]\psi)$

The Logic of Public Observation, continued

- **Common Knowledge**: $[!p]Cp$, what is the reduction axiom for 'C'?

The Logic of Public Observation, continued

- **Common Knowledge**: $[!p]Cp$, what is the reduction axiom for 'C'?

  $\mathcal{M}, w \models C^{\varphi}\psi$ iff ψ is true in all worlds reachable by a finite path starting at w *going through states satisfying* φ.

The Logic of Public Observation, continued

- **Common Knowledge**: $[!p]Cp$, what is the reduction axiom for 'C'?

  $\mathcal{M}, w \models C^{\varphi}\psi$ iff ψ is true in all worlds reachable by a finite path starting at w *going through states satisfying* φ.

  $[!\psi]C\varphi \leftrightarrow (\psi \rightarrow C^{\psi}[!\psi]\varphi)$
  $[!\psi]C^{\alpha}\varphi \leftrightarrow (\psi \rightarrow C^{\psi \wedge [!\psi]\alpha}[!\psi]\varphi)$

The Logic of Public Observation, continued

- **Common Knowledge**: $[!p]Cp$, what is the reduction axiom for 'C'?

  $\mathcal{M}, w \models C^{\varphi}\psi$ iff ψ is true in all worlds reachable by a finite path starting at w *going through states satisfying* φ.

  $[!\psi]C\varphi \leftrightarrow (\psi \rightarrow C^{\psi}[!\psi]\varphi)$
  $[!\psi]C^{\alpha}\varphi \leftrightarrow (\psi \rightarrow C^{\psi\wedge[!\psi]\alpha}[!\psi]\varphi)$

- **Make time explicit**: $[!\varphi]CY\varphi$: "After finding out that φ, it is common knowledge that φ *was* true"

Agents may differ in precisely how they incorporate new information into their epistemic states. These differences are based, in part, on the agents' perception of the *source* of the information. For example, an agent may consider a particular source of information *infallible* (not allowing for the possibility that the source is mistaken) or merely *trustworthy* (accepting the information as reliable, though allowing for the possibility of a mistake).

Hard and Soft Updates

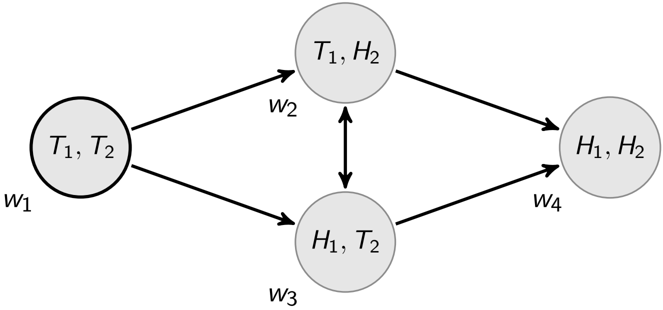
$$\mathcal{M} = \langle W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i\}_{i\in\mathcal{A}}, V\rangle$$

$\Downarrow$

Find out that φ

$\Downarrow$

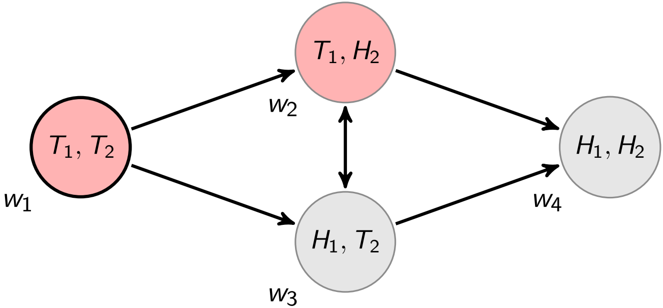
$$\mathcal{M} = \langle W', \{\sim'_i\}_{i\in\mathcal{A}}, \{\preceq'_i\}_{i\in\mathcal{A}}, V|_{W'}\rangle$$



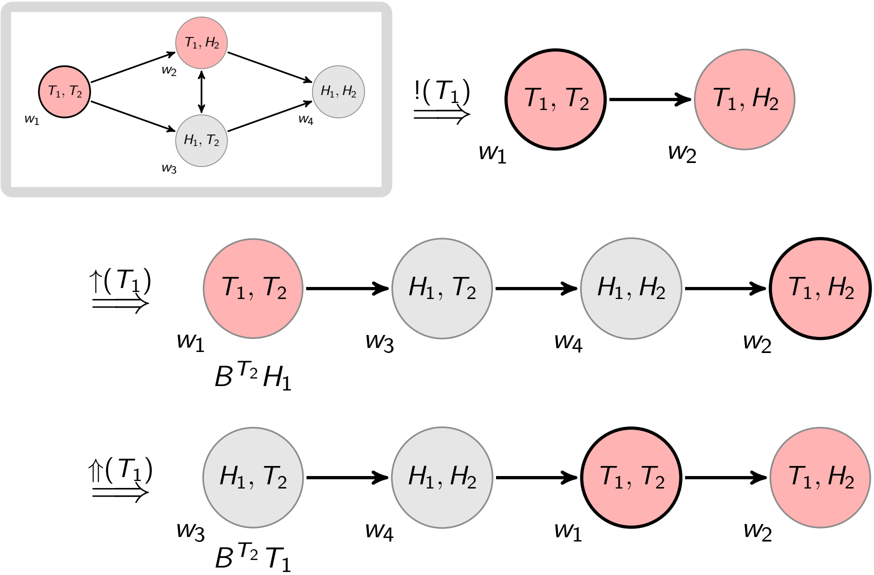
$Min_{\preceq}([w_1]) = \{w_4\}$, so $w_1 \models B(H_1 \wedge H_2)$

$Min_{\preceq}([w_1] \cap [\![T_1]\!]_{\mathcal{M}}) = \{w_2\}$, so $w_1 \models B^{T_1} H_2$

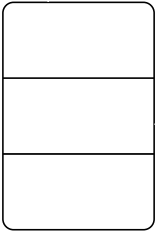
$Min_{\preceq}([w_1] \cap [\![T_1]\!]_{\mathcal{M}}) = \{w_3\}$, so $w_1 \models B^{T_2} H_1$



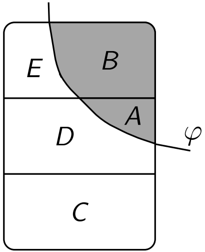
Suppose the agent *finds out that* T_1 *is/may be true*.


T_1, H_2
w_2
T_1, T_2
w_1
H_1, H_2
w_4
H_1, T_2
w_3
$\overset{!(T_1)}{\Longrightarrow}$
T_1, T_2 w_1
T_1, H_2 w_2
$\overset{\uparrow(T_1)}{\Longrightarrow}$
T_1, T_2 w_1 $B^{T_2} H_1$
H_1, T_2 w_3
H_1, H_2 w_4
T_1, H_2 w_2
$\overset{\Uparrow(T_1)}{\Longrightarrow}$
H_1, T_2 w_3 $B^{T_2} T_1$
H_1, H_2 w_4
T_1, T_2 w_1
T_1, H_2 w_2

Informative Actions

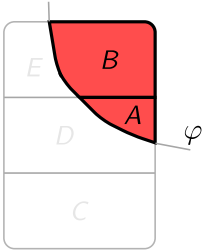


Informative Actions



Incorporate the new information φ

Informative Actions

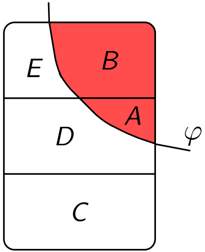


Public Announcement: Information from an infallible source

$(!\varphi)$: $A \prec_i B$ $\quad\mathcal{M}^{!\varphi} = \langle W^{!\varphi}, \{\sim_i^{!\varphi}\}_{i\in\mathcal{A}}, V^{!\varphi}\rangle$

$W^{!\varphi} = [\![\varphi]\!]_{\mathcal{M}}$
$\sim_i^{!\varphi} = \sim_i \cap (W^{!\varphi} \times W^{!\varphi})$
$\preceq_i^{!\varphi} = \preceq_i \cap (W^{!\varphi} \times W^{!\varphi})$

Informative Actions

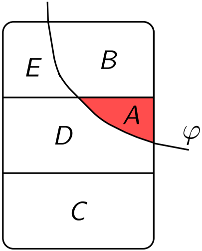


Radical Upgrade: $(\Uparrow\varphi)$: $A \prec_i B \prec_i C \prec_i D \prec_i E$,
$\mathcal{M}^{\Uparrow\varphi} = \langle W, \{\sim_i\}_{i\in\mathcal{A}}, \{\preceq_i^{\Uparrow\varphi}\}_{i\in\mathcal{A}}, V\rangle$

Let $[\![\varphi]\!]_i^w = \{x \mid \mathcal{M}, x \models \varphi\} \cap [w]_i$

- for all $x \in [\![\varphi]\!]_i^w$ and $y \in [\![\neg\varphi]\!]_i^w$, set $x \prec_i^{\Uparrow\varphi} y$,
- for all $x, y \in [\![\varphi]\!]_i^w$, set $x \preceq_i^{\Uparrow\varphi} y$ iff $x \preceq_i y$, and
- for all $x, y \in [\![\neg\varphi]\!]_i^w$, set $x \preceq_i^{\Uparrow\varphi} y$ iff $x \preceq_i y$.

Informative Actions



Conservative Upgrade: $(\uparrow\varphi)$: $A \prec_i C \prec_i D \prec_i B \cup E$

Conservative upgrade is radical upgrade with the formula

$$best_i(\varphi, w) := Min_{\preceq_i}([w]_i \cap \{x \mid \mathcal{M}, x \models \varphi\})$$

1. If $v \in best_i(\varphi, w)$ then $v \prec_i^{\uparrow\varphi} x$ for all $x \in [w]_i$, and
2. for all $x, y \in [w]_i - best_i(\varphi, w)$, $x \preceq_i^{\uparrow\varphi} y$ iff $x \preceq_i y$.

Reduction Axioms

$$[\Uparrow\varphi]B^{\psi}\chi \leftrightarrow (L(\varphi \wedge [\Uparrow\varphi]\psi) \wedge B^{\varphi\wedge[\Uparrow\varphi]\psi}[\Uparrow\varphi]\chi)\vee$$
$$(\neg L(\varphi \wedge [\Uparrow\varphi]\psi) \wedge B^{[\Uparrow\varphi]\psi}[\Uparrow\varphi]\chi)$$

Reduction Axioms

$$[\Uparrow\varphi]B^{\psi}\chi \leftrightarrow (L(\varphi \wedge [\Uparrow\varphi]\psi) \wedge B^{\varphi\wedge[\Uparrow\varphi]\psi}[\Uparrow\varphi]\chi)\vee$$
$$(\neg L(\varphi \wedge [\Uparrow\varphi]\psi) \wedge B^{[\Uparrow\varphi]\psi}[\Uparrow\varphi]\chi)$$

$$[\uparrow\varphi]B^{\psi}\chi \leftrightarrow (B^{\varphi}\neg[\uparrow\varphi]\psi \wedge B^{[\uparrow\varphi]\psi}[\uparrow\varphi]\chi)\vee(\neg B^{\varphi}\neg[\uparrow\varphi]\psi \wedge B^{\varphi\wedge[\uparrow\varphi]\psi}[\uparrow\varphi]\chi)$$

What happens as beliefs change over time (iterated belief revision)?

$$\underset{\text{initial model}}{\mathcal{M}_0} \xrightarrow{!\varphi_1} \mathcal{M}_1 \xrightarrow{!\varphi_2} \mathcal{M}_2 \xrightarrow{!\varphi_3} \cdots \xrightarrow{!\varphi_n} \underset{\text{fixed-point}}{\mathcal{M}_f}$$

$$\mathcal{M}_0 \xRightarrow{\Uparrow\varphi_1} \mathcal{M}_1 \xRightarrow{\Uparrow\varphi_2} \mathcal{M}_2 \xRightarrow{\Uparrow\varphi_3} \cdots \xRightarrow{\Uparrow\varphi_n} \mathcal{M}_f$$

initial model

fixed-point

$$\mathcal{M}_0 \xRightarrow{!\varphi_1} \mathcal{M}_1 \xRightarrow{\Uparrow\varphi_2} \mathcal{M}_2 \xRightarrow{\uparrow\varphi_3} \cdots \xRightarrow{\Uparrow\varphi_n} \mathcal{M}_f$$

initial model

fixed-point

$$\underset{\text{initial model}}{\mathcal{M}_0} \xRightarrow{\tau(\varphi_1)} \mathcal{M}_1 \xRightarrow{\tau(\varphi_2)} \mathcal{M}_2 \xRightarrow{\tau(\varphi_3)} \cdots \xRightarrow{\tau(\varphi_n)} \underset{\text{fixed-point}}{\mathcal{M}_f}$$

Where do the φ_k come from?

Iterated Updates

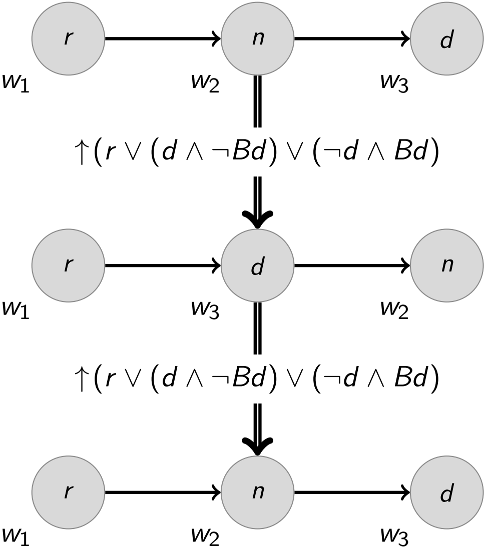
$!\varphi_1, !\varphi_2, !\varphi_3, \ldots, !\varphi_n$
always reaches a fixed-point

$\Uparrow p \Uparrow \neg p \Uparrow p \cdots$
Contradictory beliefs leads to oscillations

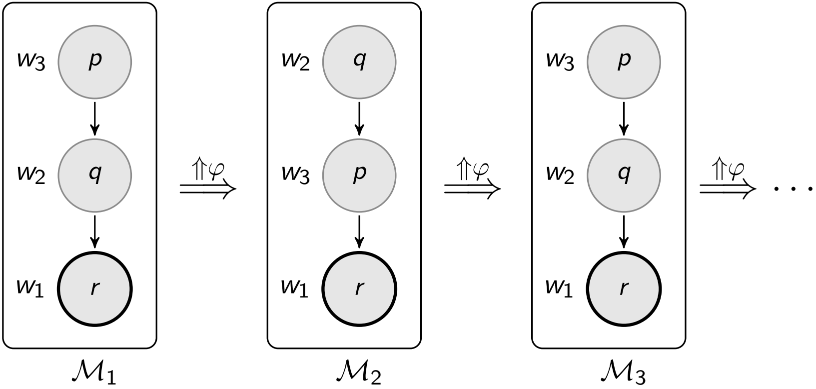
$\uparrow \varphi, \uparrow \varphi, \ldots$
Simple beliefs may never stabilize

$\Uparrow \varphi, \Uparrow \varphi, \ldots$
Simple beliefs stabilize, but conditional beliefs do not

A. Baltag and S. Smets. *Group Belief Dynamics under Iterated Revision: Fixed Points and Cycles of Joint Upgrades*. TARK, 2009.


r
n
d
w_1
w_2
w_3
$\uparrow(r \vee (d \wedge \neg Bd) \vee (\neg d \wedge Bd))$
r
d
n
w_1
w_3
w_2
$\uparrow(r \vee (d \wedge \neg Bd) \vee (\neg d \wedge Bd))$
r
n
d
w_1
w_2
w_3

Let φ be $(r \vee (B^{\neg r} q \wedge p) \vee (B^{\neg r} p \wedge q))$



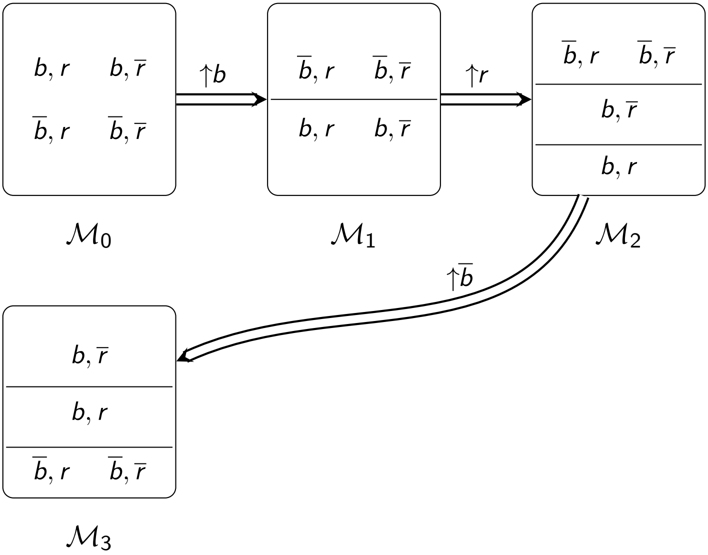
Suppose that you are in the forest and happen to a see strange-looking animal.

Suppose that you are in the forest and happen to a see strange-looking animal. You consult your animal guidebook and find a picture that seems to match the animal you see.

Suppose that you are in the forest and happen to a see strange-looking animal. You consult your animal guidebook and find a picture that seems to match the animal you see. The guidebook says that the animal is a type of bird, so that is what you conclude: The animal before you is a bird. After looking more closely, you also notice that the animal is also red.

Suppose that you are in the forest and happen to a see strange-looking animal. You consult your animal guidebook and find a picture that seems to match the animal you see. The guidebook says that the animal is a type of bird, so that is what you conclude: The animal before you is a bird. After looking more closely, you also notice that the animal is also red. So, you also update your beliefs with that fact.

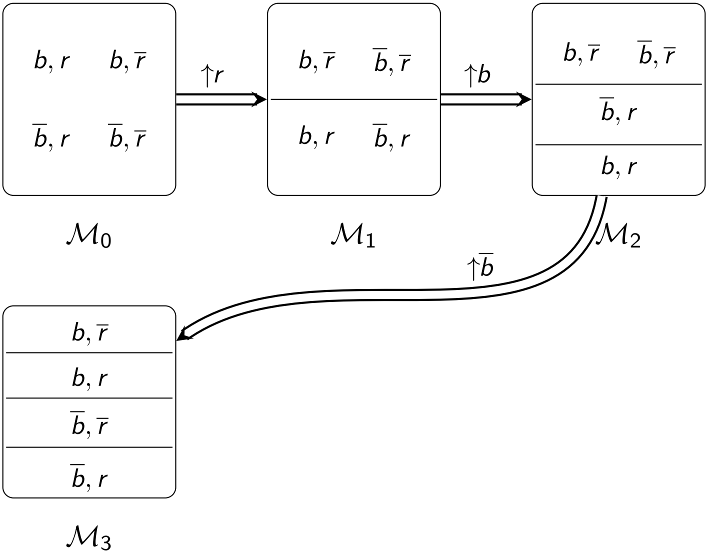
Suppose that you are in the forest and happen to a see strange-looking animal. You consult your animal guidebook and find a picture that seems to match the animal you see. The guidebook says that the animal is a type of bird, so that is what you conclude: The animal before you is a bird. After looking more closely, you also notice that the animal is also red. So, you also update your beliefs with that fact. Now, suppose that an expert (whom you trust) happens to walk by and tells you that the animal is, in fact, not a bird.


b, r b, r̄
b̄, r b̄, r̄
$\mathcal{M}_0$
↑b
b̄, r b̄, r̄
b, r b, r̄
$\mathcal{M}_1$
↑r
b̄, r b̄, r̄
b, r̄
b, r
$\mathcal{M}_2$
↑b̄
b, r̄
b, r
b̄, r b̄, r̄
$\mathcal{M}_3$

Note that in the last model, $\mathcal{M}_3$, the agent does not believe that the bird is red.

Note that in the last model, $\mathcal{M}_3$, the agent does not believe that the bird is red. The problem is that there does not seem to be any justification for why the agent drops her belief that the bird is red. This seems to result from the accidental fact that the agent started by updating with the information that the animal is a bird.

Note that in the last model, $\mathcal{M}_3$, the agent does not believe that the bird is red. The problem is that there does not seem to be any justification for why the agent drops her belief that the bird is red. This seems to result from the accidental fact that the agent started by updating with the information that the animal is a bird. In particular, note that the following sequence of updates is not problematic:


b, r $b, \overline{r}$
$\overline{b}, r$ $\overline{b}, \overline{r}$
$\mathcal{M}_0$
$\uparrow r$
$b, \overline{r}$ $\overline{b}, \overline{r}$
b, r $\overline{b}, r$
$\mathcal{M}_1$
$\uparrow b$
$b, \overline{r}$ $\overline{b}, \overline{r}$
$\overline{b}, r$
b, r
$\mathcal{M}_2$
$\uparrow \overline{b}$
$b, \overline{r}$
b, r
$\overline{b}, \overline{r}$
$\overline{b}, r$
$\mathcal{M}_3$

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Three switches wired such that a light is on iff all three switches are up or all three are down.

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Three switches wired such that a light is on iff all three switches are up or all three are down.
- Three independent (reliable) observers report on the switches: Alice says switch 1 is U, Bob says switch 2 is D and Carla says switch 3 is U.

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Three switches wired such that a light is on iff all three switches are up or all three are down.
- Three independent (reliable) observers report on the switches: Alice says switch 1 is U, Bob says switch 2 is D and Carla says switch 3 is U.
- I receive the information that the light is on. What should I believe?

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Three switches wired such that a light is on iff all three switches are up or all three are down.
- Three independent (reliable) observers report on the switches: Alice says switch 1 is U, Bob says switch 2 is D and Carla says switch 3 is U.
- I receive the information that the light is on. What should I believe?
- Cautious: UUU, DDD; Bold: UUU

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Suppose there are two switches: L_1 is the main switch and L_2 is a secondary switch controlled by the first two lights. (So $L_1 \to L_2$, but not the converse)

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Suppose there are two switches: L_1 is the main switch and L_2 is a secondary switch controlled by the first two lights. (So $L_1 \to L_2$, but not the converse)
- Suppose I receive $L_1 \wedge L_2$, this does not change the story.

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Suppose there are two switches: L_1 is the main switch and L_2 is a secondary switch controlled by the first two lights. (So $L_1 \to L_2$, but not the converse)
- Suppose I receive $L_1 \wedge L_2$, this does not change the story.
- Suppose I learn that L_2. This is irrelevant to Carla's report, but it means either Ann or Bob is wrong.

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Suppose there are two switches: L_1 is the main switch and L_2 is a secondary switch controlled by the first two lights. (So $L_1 \rightarrow L_2$, but not the converse)
- Suppose I receive $L_1 \wedge L_2$, this does not change the story.
- Suppose I learn that L_2. This is irrelevant to Carla's report, but it means either Ann or Bob is wrong.

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Suppose there are two switches: L_1 is the main switch and L_2 is a secondary switch controlled by the first two lights. (So $L_1 \rightarrow L_2$, but not the converse)
- Suppose I receive $L_1 \wedge L_2$, this does not change the story.
- Suppose I learn that L_2. This is irrelevant to Carla's report, but it means either Ann or Bob is wrong.
- Now, after learning L_1, the only rational thing to believe is that all three switches are up.

UUU	DDD
UUD	DDU
UDU	DUD
UDD	DUU

- Suppose there are two switches: L_1 is the main switch and L_2 is a secondary switch controlled by the first two lights. (So $L_1 \to L_2$, but not the converse)
- Suppose I receive $L_1 \wedge L_2$, this does not change the story.
- Suppose I learn that L_2. This is irrelevant to Carla's report, but it means either Ann or Bob is wrong.
- Now, after learning L_1, the only rational thing to believe is that all three switches are up.

Many of the recent developments in this area have been driven by analyzing *concrete* examples.

This raises an important methodological issue: Implicit assumptions about what the actors know and believe about the situation being modeled often guide the analyst's intuitions. In many cases, it is crucial to make these underlying assumptions explicit.

The general point is that *how* the agent(s) come to know or believe that some proposition p is true is as important (or, perhaps, more important) than the fact that the agent(s) knows or believes that p is the case

Discussion

A key aspect of any formal model of a (social) interactive situation or situation of rational inquiry is the way it accounts for the

> *...information about how I learn some of the things I learn, about the sources of my information, or about what I believe about what I believe and don't believe. If the story we tell in an example makes certain information about any of these things relevant, then it needs to be included in a proper model of the story, if it is to play the right role in the evaluation of the abstract principles of the model.* *(Stalnaker, pg. 203)*

R. Stalnaker. *Iterated Belief Revision*. Erkentnis 70, pgs. 189 209, 2009.