Logic and Artificial Intelligence

Lecture 10

Eric Pacuit

Currently Visiting the Center for Formal Epistemology, CMU

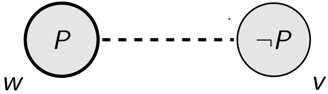
Center for Logic and Philosophy of Science
Tilburg University

ai.stanford.edu/∼epacuit
e.j.pacuit@uvt.nl

October 3, 2011

Modeling Information Change

Models of Hard and Soft Information



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

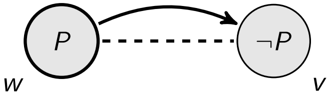
- $w \sim_i v$ means i cannot rule out v according to her information.

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

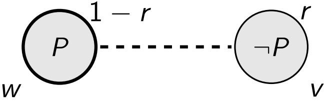
- $w \preceq_i v$ means v is at least as plausibility as w for agent i.

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.)

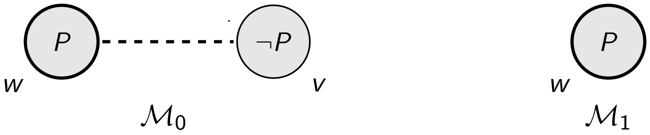
Models of Hard and Soft Information

- *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

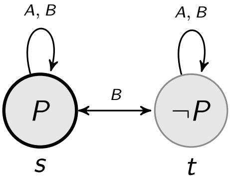
Ann would like Bob to attend her talk; however, she only wants Bob to attend if he is interested in the subject of her talk, not because he is just being polite.

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

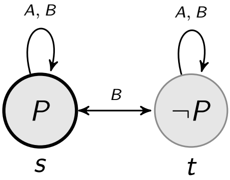
1. Ann knows about the talk.
2. Bob knows about the talk.
3. Ann knows that Bob knows about the talk.
4. Bob *does not* know that Ann knows that he knows about the talk.
5. *And nothing else.*

Example



P means "The talk is at 2PM".

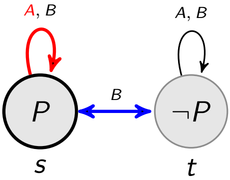
Example



P means "The talk is at 2PM".

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

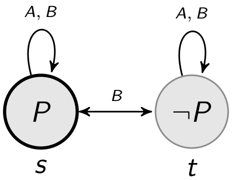
Example



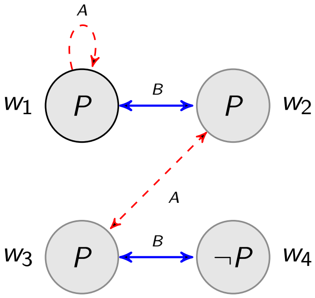
P means "The talk is at 2PM".

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

Example



Prior Model



Posterior Model

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.

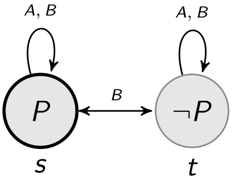
Aspects of Informative Events

1. The agents' *observational* powers.

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

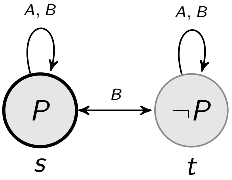
3. The underlying *protocol* specifying which events (observations, messages, actions) are available (or permitted) at any given moment.

Dynamic Events: Public Announcement



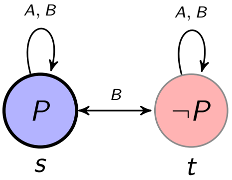
P means "The talk is at 2PM".

Dynamic Events: Public Announcement



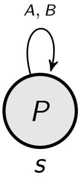
What happens if Ann publicly announces P?

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

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

Public Announcement Logic

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

Public Announcement Logic

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

Public Announcement Logic

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

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.

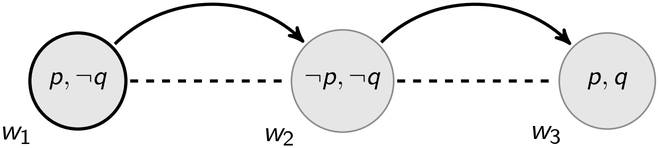
J. van Benthem, J. van Eijk, B. Kooi. *Logics of Communication and Change*. 2006.

- $[q]Kq$

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

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

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



- $[\varphi]\varphi$

Public Announcement vs. Conditional Belief

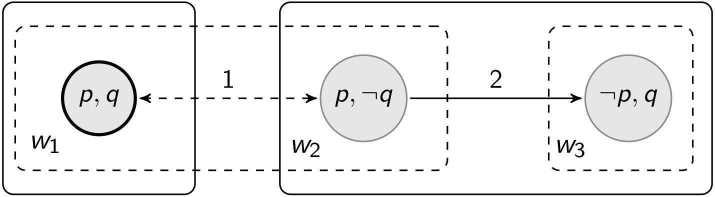
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different?

Public Announcement vs. Conditional Belief

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

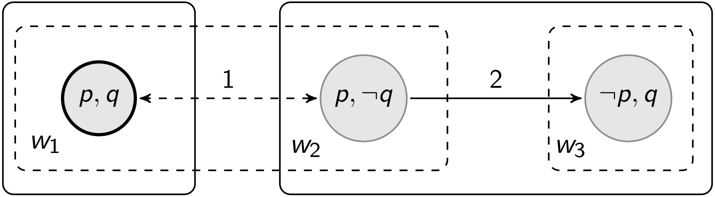
Public Announcement vs. Conditional Belief

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



Public Announcement vs. Conditional Belief

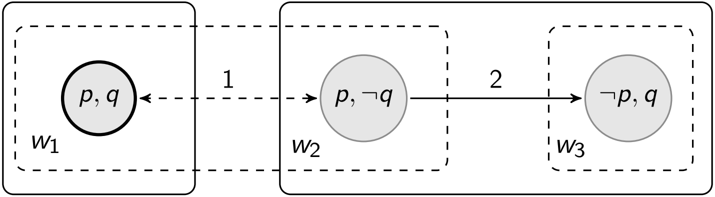
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



- $w_1 \models B_1 B_2 q$

Public Announcement vs. Conditional Belief

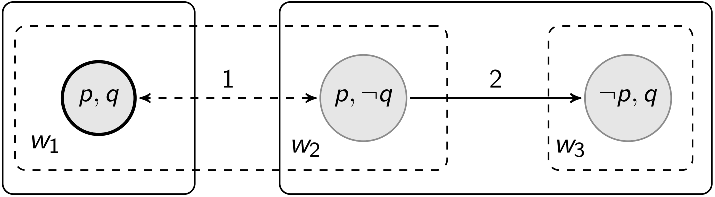
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



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

Public Announcement vs. Conditional Belief

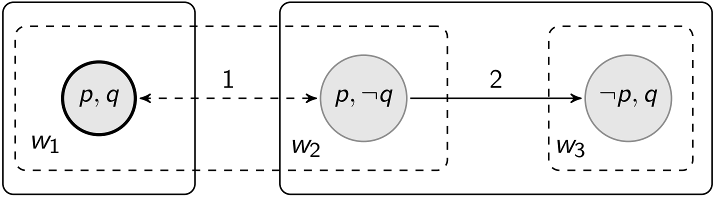
Are $[\varphi]B\psi$ and $B^{\varphi}\psi$ different? Yes!



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

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.

Recursion Axioms: Belief and Conditional Belief

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

Recursion Axioms: Belief and Conditional Belief

$$[\varphi][\preceq_i]\psi \leftrightarrow (\varphi \rightarrow [\preceq_i](\varphi \rightarrow [\varphi]\psi))$$

$$[\varphi]B\psi \not\leftrightarrow (\varphi \rightarrow B(\varphi \rightarrow [\varphi]\psi))$$

Recursion Axioms: Belief and Conditional Belief

$$[\varphi][\preceq_i]\psi \leftrightarrow (\varphi \rightarrow [\preceq_i](\varphi \rightarrow [\varphi]\psi))$$

$$[\varphi]B\psi \not\leftrightarrow (\varphi \rightarrow B(\varphi \rightarrow [\varphi]\psi))$$

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

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