Private Belief Revision

Patrick Girard

Philosophy Department, The University of Auckland

September 29, 2011

Joint work with

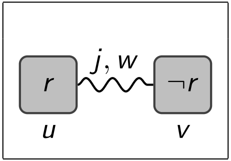
⋆ Liu Fenrong (Tsinghua) ⋆ Jeremy Seligman (Auckland) ⋆

Outline

bad.jpg

The lab. Knowledge

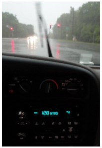
Walter and Jesse are cooking methamphetamine in an underground lab in New-Mexico. The lab has no windows and neither of them know whether it is raining. This can be represented by the following model:

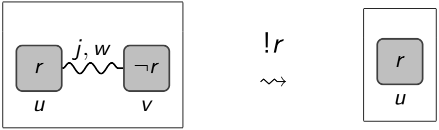


⋆ Throughout the talk, reflexive and transitive links are not drawn, for readability, but you should assume that models are always reflexive and transitive.

Scenario 1. Knowledge.

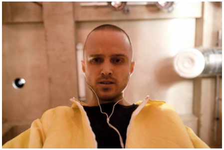
The weather reports on the radio announces that it is raining. After the announcement, Walter and Jesse know that it is raining.

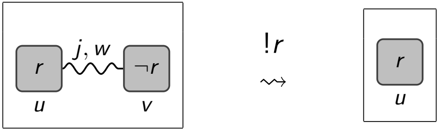




Scenario 2. Knowledge

Jesse is listening to his iPod and doesn't hear the radio announcement. After the announcement, Walter knows that it is raining, but Jesse doesn't.

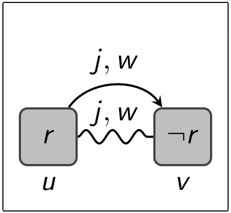




Problem: If we do simple world elimination, we get the wrong effect. Jesse now knows that it is raining.

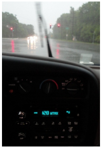
The lab. Knowledge and Belief.

Walter and Jesse are cooking methamphetamine in an underground lab in New-Mexico. The lab has no windows and neither of them know whether it is raining. But they live in New Mexico, where it rains very rarely, so they believe that it is not raining. This can be represented by the following model:



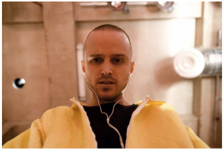
Scenario 3. Knowledge.

The weather reports on the radio announces that it is raining, but the reception is very bad in the lab, and neither Walter nor Jesse hear the report properly. After the announcement, Walter and Jesse still do not know that it is raining, but believe that it is.



Scenario 4. Knowledge and Belief.

Jesse is listening to his iPod and doesn't hear the radio announcement. The reception on the radio is very bad in the lab, so after the announcement, Walter only get to believe that it is raining, and Jesse is unaware of the announcement.



Problem: If we do a simple belief change for Walter, Jesse now knows that Walter believes that it is raining.

Outline

Doxastic Epistemic Logic

A language for reasoning about knowledge and belief.

$$[\sim_a]\varphi \quad = \quad \textit{agent } a \textit{ knows that } \varphi$$
$$[\leq_a]\varphi \quad = \quad \textit{agent } a \textit{ considers } \varphi \textit{ more plausible.}$$

- In state u, agent a knows that φ iff φ is a correct description of all states that a cannot distinguish from u.
    - Typically written $K_a\varphi$.
- In state u, agent a considers φ more plausible iff φ is a correct description of all states that a considers at least as plausible as u.
    - Sometimes known as "safe belief".

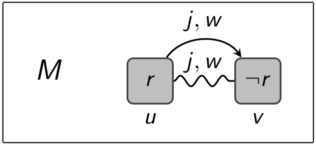
Doxastic Epistemic Models

A **doxastic epistemic model** $M = \langle W, A, \sim, \leq, V \rangle$ consists of

- a set A of *agents*, a set W of *epistemic alternatives* and a *valuation* function V from Prop to subsets of W.
- an equivalence relation $\sim_a$ on W of *indistinguishability*, for each agent a in A
- a preorder (reflexive and transitive relation) $\leq_a$ on W of *plausibility*, for each agent a in A

Semantics

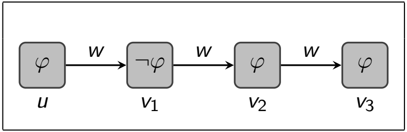
$M, w \models [\sim_a]\varphi$ iff $M, v \models \varphi$ for every $v \sim_a w$.
$M, w \models [\leq_a]\varphi$ iff $M, v \models \varphi$ for every $w \leq_a v$.



$M, u \not\models [\sim_j]r$
$M, u \not\models [\leq_w]\neg r$
$M, v \models [\leq_w]\neg r$

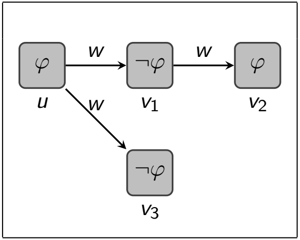
Defining Belief

It is typical to define beliefs as descriptions of most plausible states. For instance, in state u, we can say that φ is a description of the most plausible states for agent a if $M, u \models \langle \leq_a \rangle [\leq_a] \varphi$, as in:



$$M, u \models \langle \leq_w \rangle [\leq_w] \varphi$$

But this is not enough. Consider the model:



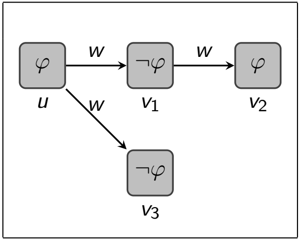
$$M, u \models \langle \leq_w \rangle [\leq_w] \varphi$$

Defining Belief

So we need to make beliefs more global. A typical solution is to use the global modality $U\varphi$ with the semantics given by:

$$M, u \models U\varphi \quad \text{iff} \quad M, v \models \varphi \text{ for every } v \in W$$

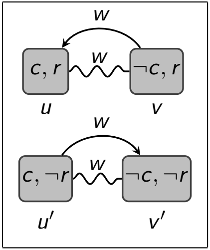
In state u, we can then say that φ is a description of the most plausible states for agent a if $M, u \models U\langle\leq_a\rangle[\leq_a]\varphi$. With this definition, we rule out the previous counter-example:



$$M, u \not\models U\langle\leq_w\rangle[\leq_w]\varphi$$

Defining

But this is too much in the presence of knowledge. For instance, if Walter knows that it is raining, he believes that it is cloudy, whereas if he knows that it's not raining, he believes that it is not cloudy.



$$M, u \not\models U\langle\leq_w\rangle[\leq_w]c$$

So instead of using the global modality $U\varphi$, we use $[\sim_a]\varphi$ and define belief as:

$$B\varphi \quad := \quad [\sim_a]\langle\leq_a\rangle[\leq_a]\varphi$$

Outline

Program Language

A language for reasoning about programs (PDL).

$$\alpha \quad := \quad \top \mid ?\varphi \mid \sim_i \mid \leq_i \mid \alpha;\alpha \mid (\alpha \cup \alpha)$$

Expression of type α are called *program terms*. Program terms α are names for relational programs $\underline{\alpha}$ defined by:

$$
\begin{array}{lcl}
\underline{\top} & = & W^2 \\
\underline{?\varphi} & = & \{\langle u, u\rangle \mid M, u \models \varphi\} \\
\underline{\sim_i} & = & \{\langle u, v\rangle \mid u \sim_i v \in M\} \\
\underline{\leq_i} & = & \{\langle u, v\rangle \mid u \leq_i v \in M\} \\
\underline{\alpha_1;\alpha_2} & = & \{\langle u, v\rangle \mid \exists w : \langle u, w\rangle \in \underline{\alpha_1}, \langle w, v\rangle \in \underline{\alpha_2}\} \\
\underline{(\alpha_1 \cup \alpha_2)} & = & \underline{\alpha_1} \cup \underline{\alpha_2}
\end{array}
$$

Programs

Let Prog $= \{[\epsilon_1, \ldots, \epsilon_n, \delta_1, \ldots, \delta_n]\}$, with ϵ_i an δ_i epistemic and doxastic programs respectively, be a set of tuples of programs. Each program $\pi \in$ Prog specifies a transformation on a model M:

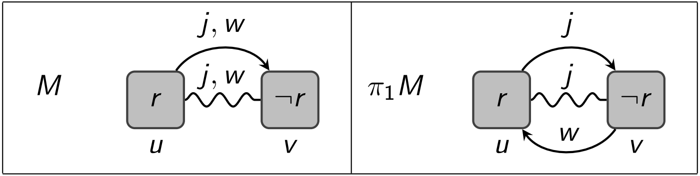
$$[\epsilon_1, \ldots, \epsilon_n, \delta_1, \ldots, \delta_n]M = \langle W, \underline{\epsilon_1}, \ldots, \underline{\epsilon_n}, \underline{\delta_1}, \ldots, \underline{\delta_n}, V \rangle$$

Example

Let

- $\epsilon_w \;=\; (?r; \sim_w; ?r) \cup (?\neg r; \sim_w; ?\neg r)$
  - Walter learns whether it is raining.
- $\delta_w \;=\; (?r; \leq_w; ?r) \cup (?\neg r; \leq_w; ?\neg r) \cup (?\neg r; \top; ?r)$
  - Walter starts believing that it is raining.
- $\pi_1 = [\epsilon_w, \sim_j, \delta_w, \leq_j]$

We compute the transformations specified by π_1 on the following model M:



Outline

Program Models

A **program model** $\Delta = \langle D, A, \sim, \leq, \text{Prog} \rangle$ consists of

- a set D of *program states*
- a set A of *agents*
- an equivalence relation $\sim_a$ on D of *indistinguishability*, for each agent a in A
- a preorder (reflexive and transitive relation) $\leq_a$ on d of *plausibility*, for each agent a in A
- a *valuation* function V from D to Prog.

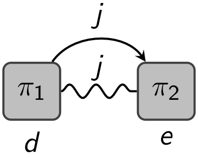
Restriction: Agents know which actions they are performing:

$$d \sim_i e \text{ implies that } \begin{array}{lcl} 1)\ P(d)_{\epsilon_i} & = & P(e)_{\epsilon_i} \\ 2)\ P(d)_{\delta_i} & = & P(e)_{\delta_i} \end{array}$$

Example

Let

- $\epsilon_w = (?r; \sim_w; ?r) \cup (?\neg r; \sim_w; ?\neg r)$
    - Walter learns whether it is raining.
- $\delta_w = (?r; \leq_w; ?r) \cup (?\neg r; \leq_w; ?\neg r) \cup (?\neg r; \top; ?r)$
    - Walter starts believing that it is raining.
- $\pi_1 = [\epsilon_w, \sim_j, \delta_w, \leq_j]$
    - Walter learns whether it is raining and starts believing that it is, and Jesse does nothing.
- $\pi_2 = [\sim_w, \sim_j, \leq_w, \leq_j]$
    - Both Walter and Jesse do nothing.



Program Actions

Given a doxastic epistemic model M and a program model Δ, define the *programmed model* $M \otimes \Delta = \langle W \times D, A, \sim', \leq', V' \rangle$ where

- V' is a propositional valuation from Prop to subsets of $W \times D$ such that $V'(p) = \{\langle u, d \rangle \mid u \in V(p)\}$
- the relations $\sim'$ and $\leq'$ are computed as follows.
  - $\langle u, d \rangle \sim'_i \langle v, e \rangle$ iff

    1) $d = e \quad \Rightarrow \quad \langle u, v \rangle \in \underline{P(d)_{\epsilon_i}}$

    2) $u = v \quad \Rightarrow \quad \langle u, v \rangle \in \underline{P(d)_{\epsilon_i}} \cap \underline{P(e)_{\epsilon_i}}$

    3) $u \neq v$ and $d \neq e \quad \Rightarrow \quad \langle u, v \rangle \in \sim_i$

    4) $\langle d, e \rangle \in \sim_i$
  - $\langle u, d \rangle \leq'_i \langle v, e \rangle$ iff

    1) $d = e \quad \Rightarrow \quad \langle u, v \rangle \in \underline{P(d)_{\delta_i}}$

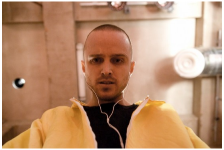
    2) $u = v \quad \Rightarrow \quad \langle u, v \rangle \in \underline{P(d)_{\delta_i}} \cap \underline{P(e)_{\delta_i}}$

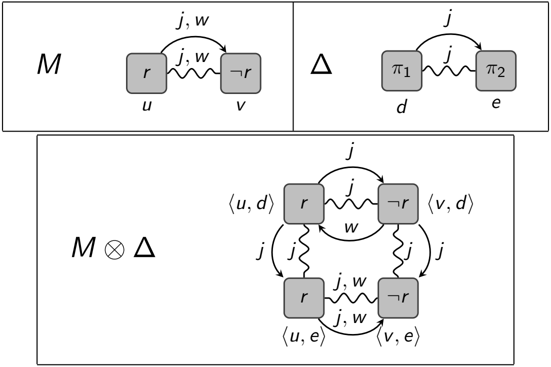
    3) $u \neq v$ and $d \neq e \quad \Rightarrow \quad \langle u, v \rangle \in \leq_i$

    4) $\langle d, e \rangle \in \leq_i$

Scenario 4. Knowledge and Belief.

Jesse is listening to his iPod and doesn't hear the radio announcement. The reception on the radio is very bad in the lab, so after the announcement, Walter only get to believe that it is raining, and Jesse is unaware of the announcement.





Program Language

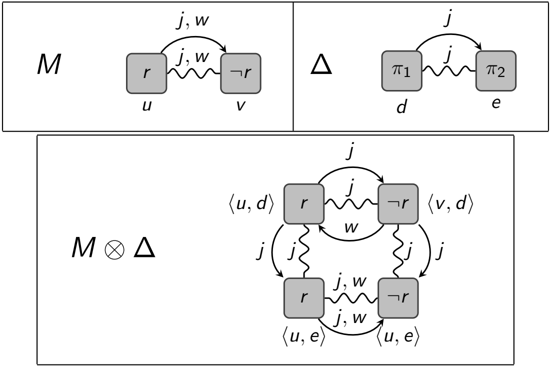
$$\alpha \; := \; \top \mid ?\varphi \mid \sim_i \mid \leq_i \mid \alpha;\alpha \mid (\alpha \cup \alpha)$$
$$\varphi \; := \; p \mid \neg\varphi \mid (\varphi \vee \varphi) \mid [\alpha]\varphi \mid [\Delta, d]\varphi$$

Semantics

$M, w \models [\alpha]\varphi$ iff $M, v \models \varphi$ for every $w\underline{\alpha}v$.

$M, w \models [\Delta, d]\varphi$ iff $M \otimes \Delta, v \models \varphi$.

Scenario 4. Knowledge and Belief.



$$M, u \models [\Delta, d]([\sim_w]r \;\&\; \neg[\sim_j][\sim_w]r)$$

i.e., $$M, u \models [\Delta, d](K_w r \;\&\; \neg K_j K_w r)$$

$$M, u \models [\Delta, d](B_w r \;\&\; B_j \neg B_w r)$$

Axiomatisation

Theorem

The set of validities of dynamic doxastic epistemic logic is axiomatised by the normal axioms and rules of PDL, standard axioms for $[\sim_i], [\leq_i]$ *and* $[\top]$*, together with the following dynamic axioms:*

$$
\begin{array}{lcl}
\langle \Delta, d \rangle p & \equiv & p \\
\langle \Delta, d \rangle \neg \varphi & \equiv & \neg \langle \Delta, d \rangle \varphi \\
\langle \Delta, d \rangle (\varphi \vee \psi) & \equiv & (\langle \Delta, d \rangle \varphi \vee \langle \Delta, d \rangle \psi) \\
\langle \Delta, d \rangle \langle \sim_i \rangle \varphi & \equiv & \langle P(d)_{\epsilon_i} \rangle \langle \Delta, d \rangle \varphi \\
\langle \Delta, d \rangle \langle \leq_i \rangle \varphi & \equiv & \langle P(d)_{\delta_i} \rangle \langle \Delta, d \rangle \varphi \\
\langle \Delta, d \rangle \langle \top \rangle \varphi & \equiv & \langle \top \rangle \langle \Delta, d \rangle \varphi
\end{array}
$$

Some References

A. Baltag, L.S. Moss, and S. Solecki.
The logic of public announcements, common knowledge and private suspicions.
In *TARK 98*, 1998.

Alexandru Baltag and Sonja Smets.
The logic of conditional doxastic actions: A theory of dynamic multi-agent belief revision.
In Sergei Artemov and Rohit Parikh, editors, *Proceedings of the Workshop on Rationality and Knowledge, ESSLLI 2006*, 2006.

J. van Benthem.
Dynamic logic for belief revision.
Journal of Applied Non-classical Logic, 17(2):129–155, 2007.

Johan van Benthem and Fenrong Liu.
The dynamics of preference upgrade.
Journal of Applied Non-Classical Logics, 17(2):157–182, 2007.

Johan van Benthem, Jan van Eijck, and Barteld Kooi.
Logics of communication and change.
Information and computation, 204(11):1620–1662, 2006.

Jan van Eijck.
Yet more modal logics of preference change and belief revision.
In Krzysztof R. Apt and Robert van Rooij, editors, *New Perspectives on Games and Interaction*, Texts in Logic and Games 4, page 81104. Amsterdam University Press, 2008.