Belief Revision and Progression of KBs in the Epistemic Situation Calculus

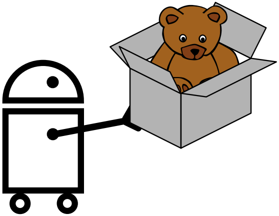
Christoph Schwering[1]   Gerhard Lakemeyer[1]   Maurice Pagnucco[2]

[1] RWTH Aachen University, Germany
[2] University of New South Wales, Australia

Actions and Beliefs

Robot is holding a box, does not know what is in it, but

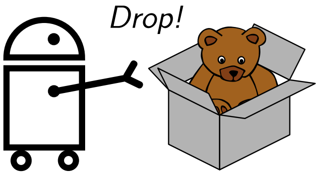
1. believes it is not fragile and not metallic
2. considers fragility more plausible than it being metallic
3. knows it is not broken yet



Actions and Beliefs

Robot is holding a box, does not know what is in it, but

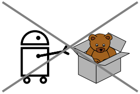
1. believes it is not fragile and not metallic
2. considers fragility more plausible than it being metallic
3. knows it is not broken yet

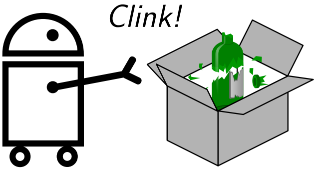


Actions and Beliefs

Robot is holding a box, does not know what is in it, but

1. believes it is not fragile and not metallic
2. considers fragility more plausible than it being metallic
3. knows it is not broken yet

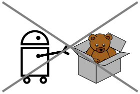


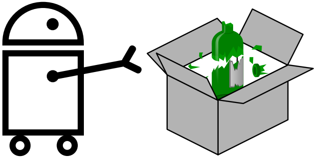


Actions and Beliefs

Robot is holding a box, does not know what is in it, but

1. believes it is not fragile and not metallic
2. considers fragility more plausible than it being metallic
3. knows it is not broken yet



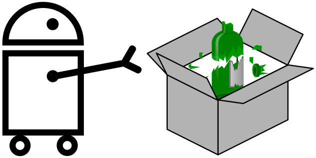


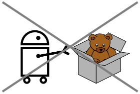
How to reason about this?

Actions and Beliefs

Robot is holding a box, does not know what is in it, but

1. believes it is not fragile and not metallic
2. considers fragility more plausible than it being metallic
3. knows it is not broken yet





If 1–3 is all it believed initially, what is all it believes now?

Disclaimer: talk includes improvements over paper (journal version in preparation)

Logic for Actions and Beliefs

First-order logic with modalities:

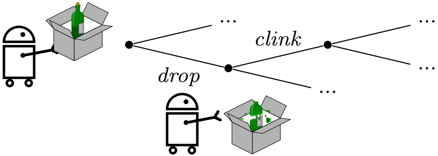
- α holds after action A $\quad [A]\alpha$
- α holds forever $\quad \Box\alpha$
- if ϕ held, ψ would hold $\quad \mathbf{B}(\phi \Rightarrow \psi) \quad \mathbf{B}\psi$
- all we believe is $\phi_i \Rightarrow \psi_i$ $\quad \mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$
- before forgetting $\mathcal{P}$, —"— $\quad \mathbf{O}_{\mathcal{P}}\{$—"—$\}$

Logic for Actions and Beliefs

First-order logic with modalities:

- α holds after action A — $[A]\alpha$
- α holds forever — $\Box\alpha$
- if ϕ held, ψ would hold — $\mathbf{B}(\phi \Rightarrow \psi)$ $\quad$ $\mathbf{B}\psi$
- all we believe is $\phi_i \Rightarrow \psi_i$ — $\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$
- before forgetting $\mathcal{P}$, —"— — $\mathbf{O}_{\mathcal{P}}\{$—"—$\}$

Semantics: worlds specify initial and future truth of fluents



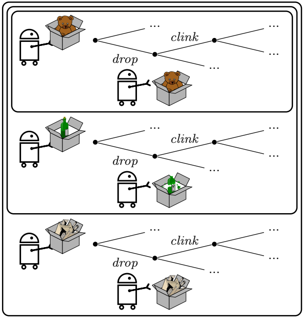
$$w \models \neg B \wedge [drop]B$$

Logic for Actions and Beliefs

First-order logic with modalities:

- α holds after action A — $[A]\alpha$
- α holds forever — $\Box\alpha$
- if ϕ held, ψ would hold — $\mathbf{B}(\phi \Rightarrow \psi)$ — $\mathbf{B}\psi$
- all we believe is $\phi_i \Rightarrow \psi_i$ — $\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$
- before forgetting $\mathcal{P}$, —"— — $\mathbf{O}_{\mathcal{P}}\{$—"—$\}$

Semantics: possible worlds ranked by plausibilities



$$e \models \mathbf{B}\neg B \wedge [drop]\mathbf{B}\neg B$$

$$e \models [drop][clink]\mathbf{B}B \quad \text{(due to revision)}$$

All We Believe …

- Only-believing uniquely determines belief structure
- Related to Levesque's only-knowing, Pearl's Z-Ordering
- Subsumes only-knowing α by conditional $\neg\alpha \Rightarrow \bot$

Theorem: Unique-Model Property

$\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ has unique model if ϕ_i, ψ_i are obj.

Theorem generalizes for $\mathbf{O}_{\mathcal{P}}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$

All We Believe ...

- Only-believing uniquely determines belief structure
- Related to Levesque's only-knowing, Pearl's Z-Ordering
- Subsumes only-knowing α by conditional $\neg\alpha \Rightarrow \perp$

Theorem: Unique-Model Property

$\mathbf{O}\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ has unique model if ϕ_i, ψ_i are obj.

Usually all we believe is a Basic Action Theory (BAT) with

- initial beliefs Σ_{bel}
- knowledge about dynamics
  - physical effect (successor-state axioms due to Reiter):
    $\forall a. \Box [a] B \equiv a = \mathit{drop} \land F \lor B$
  - epistemic effect (action A leads to revision by $\mathit{IF}(A)$):
    $\forall a. \Box \mathit{IF}(a) \equiv (a = \mathit{clink} \supset B \lor M)$

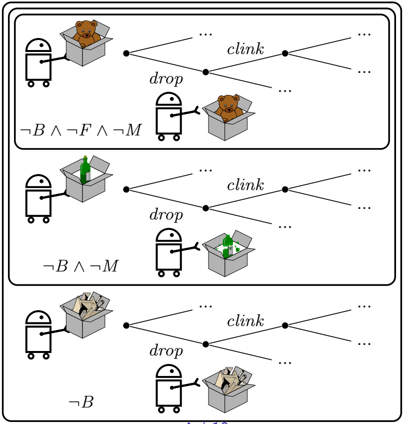
All We Believe ...

$$\mathbf{O}\{\top \Rightarrow \neg F \wedge \neg M,$$ believes it is not fragile and not metallic

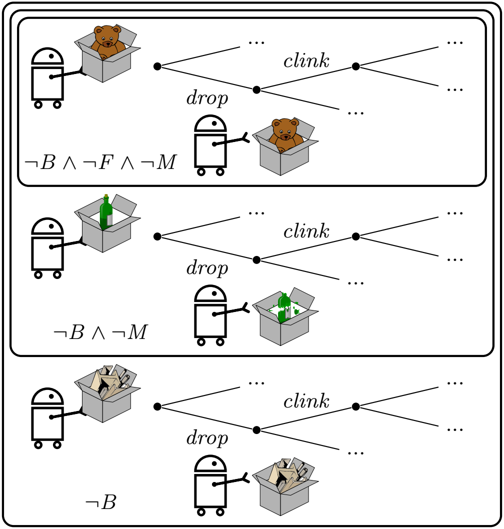
$$F \vee M \Rightarrow \neg M,$$ considers fragility more plausible than metallic

$$B \Rightarrow \bot,$$ knows it is not broken yet

dynamic axioms$\}$

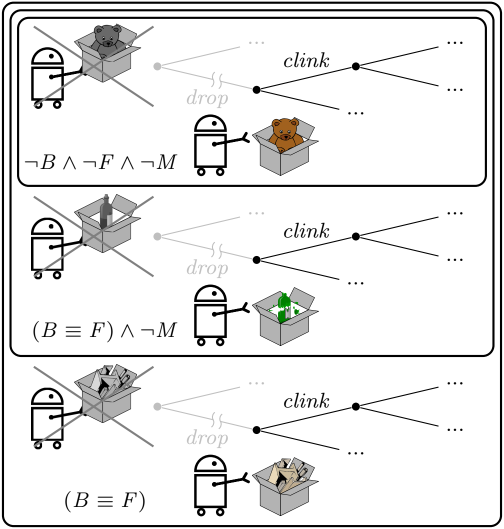


Progression of an Epistemic State



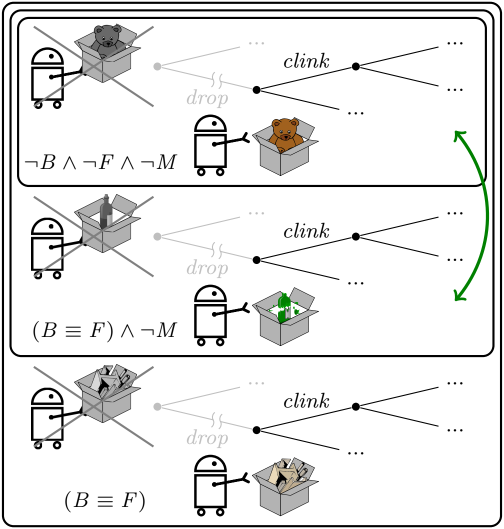
Initially $e \models \mathbf{B}\neg B$

Progression of an Epistemic State



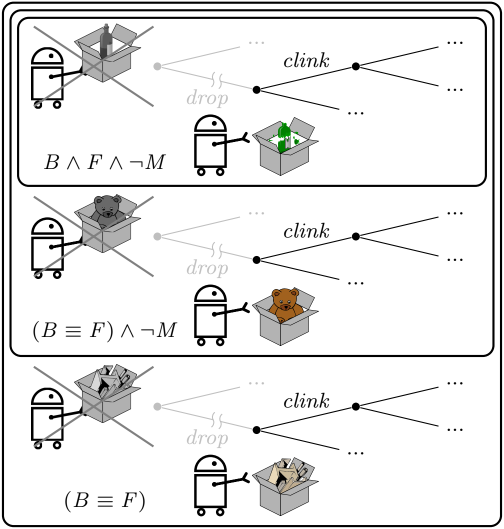
After progression still $e \gg drop \models \mathbf{B}\neg B$

Progression of an Epistemic State



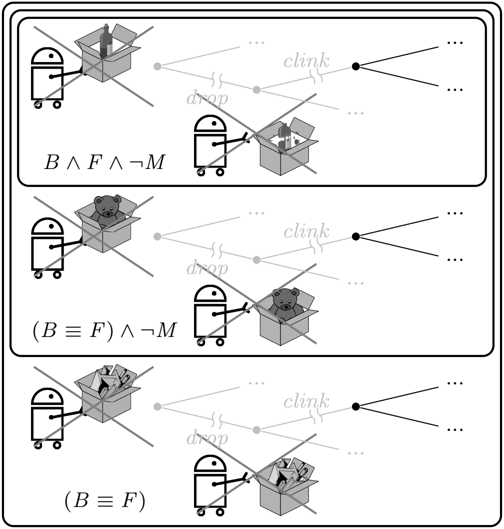
Natural revision by $B \vee M$ promotes B-worlds to the top

Progression of an Epistemic State



After revision $(e \gg drop) * (B \vee M) \models \mathbf{B}B$

Progression of an Epistemic State



After progression $e \gg drop \gg clink \models \mathbf{B}B$

Progression of a BAT by a Physical Action

- Similar to Lin and Reiter's progression
- Let A have no epistemic effect
- Let $\mathcal{F}$ be fluents of BAT with axioms $\Box[a]F(\vec{x}) \equiv \gamma_F$
- Let $\mathcal{P}$ be new predicates

Beliefs after doing A

$$\Sigma_{\text{bel}} \gg A = \Sigma_{\text{bel}}{}^{\mathcal{F}}_{\mathcal{P}} \cup \{\neg(\forall \vec{x}.F(\vec{x}) \equiv \gamma_F{}^{a}_{A}{}^{\mathcal{F}}_{\mathcal{P}}) \Rightarrow \bot \mid F \in \mathcal{F}\}$$

- Substitute $\mathcal{P}$ for $\mathcal{F}$ to capture pre-A beliefs
- Assert $\forall \vec{x}.F(\vec{x}) \equiv \gamma_F{}^{a}_{A}{}^{\mathcal{F}}_{\mathcal{P}}$ to set post-A beliefs

Progression of a BAT by an Epistemic Action

- Let A have no physical effect
- Progression $\Sigma_{\text{bel}} \gg A = \Sigma_{\text{bel}} * IF(A)$
- Let $\Delta = \{\phi \Rightarrow \psi \in \Sigma_{\text{bel}} \mid \mathbf{O}\Sigma_{\text{bel}} \models \mathbf{B}(\alpha \Rightarrow \phi \supset \psi)\}$
- Let P be a new predicate

Beliefs after promoting the most-plausible α-worlds

$$\begin{aligned}
\Sigma_{\text{bel}} * \alpha = {} & \{\top \Rightarrow P\} \cup \\
& \{\neg(P \supset \alpha) \Rightarrow \bot\} \cup \\
& \{\neg(\phi \wedge P \supset \psi) \Rightarrow \bot \mid \phi \Rightarrow \psi \in \Delta\} \cup \\
& \{\phi \wedge \neg P \Rightarrow \psi \mid \phi \Rightarrow \psi \in \Sigma_{\text{bel}}\}
\end{aligned}$$

Progression of a BAT by an Epistemic Action

- Let A have no physical effect
- Progression $\Sigma_{\text{bel}} \gg A = \Sigma_{\text{bel}} * IF(A)$
- Let $\Delta = \{\phi \Rightarrow \psi \in \Sigma_{\text{bel}} \mid \mathbf{O}\Sigma_{\text{bel}} \models \mathbf{B}(\alpha \Rightarrow \phi \supset \psi)\}$
- Let P be a new predicate

Beliefs after promoting the most-plausible α-worlds

$$\begin{aligned}
\Sigma_{\text{bel}} * \alpha = \; & \{\top \Rightarrow P\} & \cup \\
& \{\neg(P \supset \alpha) \Rightarrow \bot\} & \cup \\
& \{\neg(\phi \wedge P \supset \psi) \Rightarrow \bot \mid \phi \Rightarrow \psi \in \Delta\} & \cup \\
& \{\phi \wedge \neg P \Rightarrow \psi \mid \phi \Rightarrow \psi \in \Sigma_{\text{bel}}\}
\end{aligned}$$

- P-worlds are the most plausible worlds
- P-worlds represent promoted α-worlds
- $\neg P$-worlds represent original belief structure

Progression of a BAT

Briefly: BAT progression matches semantic progression

Theorem: Progression #1

$$\models \mathbf{O}\Sigma \supset [A]\mathbf{O}_{\mathcal{P}\cup\{P\}}(\Sigma \gg A)$$

Roughly: If all we believe is Σ, then all we believe after A is $\Sigma \gg A$.

Theorem: Progression #2

$$\models \mathbf{O}\Sigma \supset [A]\alpha \quad \text{iff} \quad \models \mathbf{O}_{\mathcal{P}\cup\{P\}}(\Sigma \gg A) \supset \alpha$$

Roughly: $\Sigma \gg A$ entails the same beliefs as Σ has after A.

Belief Revision Postulates

- Alchourron–Gärdenfors–Makinson (AGM) hold
- Darwiche–Pearl (DP) hold with a little restriction on DP2
  Original DP2 is violated because we cannot recover from an inconsistent state
- Nayak–Pagnucco–Peppas (NPP) violated because the order matters in natural revision

Conclusion and Ongoing/Future Work

- Situation calculus plus natural revision
- Belief progression using only-believing

- Other revision schemes, e.g., lexicographic
- Projection by regression
- Elimination of (nested) beliefs

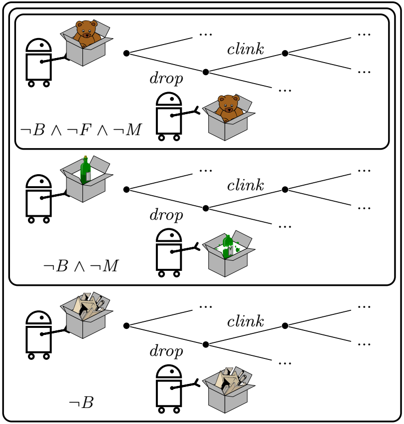
similar to our AAAI-15 paper

- When is progression first-order-definable?
- Feasible subclass   based on Lakemeyer & Levesque, KR-14
- Implementation

Appendix

All We Believe ...

- Believe it is not fragile and not metallic $\top \Rightarrow \neg F \wedge \neg M$
- Fragility is more plausible than metallic $F \vee M \Rightarrow \neg M$
- Know that it is not broken $B \Rightarrow \perp$
- Know dynamic axioms (omitted)



1. $(\top \supset \neg F \wedge \neg M) \wedge (F \vee M \supset \neg M) \wedge (B \supset \perp)$
2. $(F \vee M \supset \neg M) \wedge (B \supset \perp)$
3. $(B \supset \perp)$