Belief Revision and Progression of KBs in the Epistemic Situation Calculus

Christoph Schwering\textsuperscript{1}  Gerhard Lakemeyer\textsuperscript{1}  Maurice Pagnucco\textsuperscript{2}

\textsuperscript{1} RWTH Aachen University, Germany
\textsuperscript{2} University of New South Wales, Australia
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
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

Clink!
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:

- $\alpha$ holds after action $A$  
  $[A]\alpha$
- $\alpha$ holds forever  
  $\square \alpha$
- if $\phi$ held, $\psi$ 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 $\mathcal{P}$, —”—  
  $\mathbf{O}_\mathcal{P}\{—”—\}$
Logic for Actions and Beliefs

First-order logic with modalities:

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

Semantics: worlds specify initial and future truth of fluents

\[
\begin{align*}
w \models \neg B \land [drop]B
\end{align*}
\]
Logic for Actions and Beliefs

First-order logic with modalities:

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

Semantics: possible worlds ranked by plausibilities

$$e \models B \neg B \land [drop]B \neg B$$
$$e \models [drop][clink]BB \quad \text{(due to revision)}$$
All We Believe…

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

### Theorem: Unique-Model Property

$O\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ has unique model if $\phi_i$, $\psi_i$ are obj.

Theorem generalizes for $O_P\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$
All We Believe ...

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

**Theorem: Unique-Model Property**

$O\{\phi_1 \Rightarrow \psi_1, \ldots, \phi_m \Rightarrow \psi_m\}$ has unique model if $\phi_i$, $\psi_i$ are obj.

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

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

\[ O \{ \top \Rightarrow \neg F \land \neg M, \quad \text{believes it is not fragile and not metallic} \]

\[ F \lor M \Rightarrow \neg M, \quad \text{considers fragility more plausible than metallic} \]

\[ B \Rightarrow \bot, \quad \text{knows it is not broken yet} \]

\[ \text{dynamic axioms} \]
Progression of an Epistemic State

Initially $e \models B \neg B$
Progression of an Epistemic State

After progression still $e \gg drop \models B\neg B$
Progression of an Epistemic State

Natural revision by $B \lor M$ promotes $B$-worlds to the top
Progression of an Epistemic State

\[
B \land F \land \neg M \quad (B \equiv F) \land \neg M \quad (B \equiv F')
\]

After revision \((e \gg drop) \ast (B \lor M) \models BB\)
Progression of an Epistemic State

\[ B \land F \land \neg M \]

\[ (B \equiv F) \land \neg M \]

\[ (B \equiv F') \]

After progression \( e \gg drop \gg clink \models B B \)
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 \( \square[a]F(\vec{x}) \equiv \gamma_F \)
- Let \( \mathcal{P} \) be new predicates

Beliefs after doing \( A \)

\[
\Sigma_{\text{bel}} \gg A = \Sigma_{\text{bel}}^{\mathcal{F}} \cup \{ \neg(\forall \vec{x}.F(\vec{x}) \equiv \gamma_{F_A}^{\mathcal{F}}) \Rightarrow \bot \mid F \in \mathcal{F} \}
\]

- Substitute \( \mathcal{P} \) for \( \mathcal{F} \) to capture pre-\( A \) beliefs
- Assert \( \forall \vec{x}.F(\vec{x}) \equiv \gamma_{F_A}^{\mathcal{F}} \) to set post-\( A \) beliefs
Progression of a BAT by an Epistemic Action

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

Beliefs after promoting the most-plausible \( \alpha \)-worlds

\[
\Sigma_{bel} \ast \alpha = \{ \top \Rightarrow P \} \cup \\
\{ \neg (P \supset \alpha) \Rightarrow \bot \} \cup \\
\{ \neg (\phi \land P \supset \psi) \Rightarrow \bot \mid \phi \Rightarrow \psi \in \Delta \} \cup \\
\{ \phi \land \neg P \Rightarrow \psi \mid \phi \Rightarrow \psi \in \Sigma_{bel} \}
\]
Progression of a BAT by an Epistemic Action

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

Beliefs after promoting the most-plausible $\alpha$-worlds

$\Sigma_{\text{bel}} \ast \alpha = \{ \top \Rightarrow P \} \cup \{ \neg(P \supset \alpha) \Rightarrow \bot \} \cup \{ \neg(\phi \land P \supset \psi) \Rightarrow \bot \mid \phi \Rightarrow \psi \in \Delta \} \cup \{ \phi \land \neg P \Rightarrow \psi \mid \phi \Rightarrow \psi \in \Sigma_{\text{bel}} \}$

- $P$-worlds are the most plausible worlds
- $P$-worlds represent promoted $\alpha$-worlds
- $\neg P$-worlds represent original belief structure
Progression of a BAT

Briefly: BAT progression matches semantic progression

Theorem: Progression #1

\[ \models O\Sigma \supset [A] O_{P\cup\{P\}}(\Sigma \gg A) \]

Roughly: If all we believe is \( \Sigma \), then all we believe after \( A \) is \( \Sigma \gg A \).

Theorem: Progression #2

\[ \models O\Sigma \supset [A] \alpha \iff \models O_{P\cup\{P\}}(\Sigma \gg A) \supset \alpha \]

Roughly: \( \Sigma \gg A \) entails the same beliefs as \( \Sigma \) 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

- 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
- Fragility is more plausible than metallic
- Know that it is not broken
- Know dynamic axioms

\[
\begin{align*}
\top & \Rightarrow \neg F \land \neg M \\
F \lor M & \Rightarrow \neg M \\
B & \Rightarrow \bot
\end{align*}
\]

1. \((\top \supset \neg F \land \neg M) \land (F \lor M \supset \neg M) \land (B \supset \bot)\)

2. \((F \lor M \supset \neg M) \land (B \supset \bot)\)

3. \((B \supset \bot)\)