

Iterated Belief Change in the Situation Calculus

(Steven Shapiro, Maurice Pagnucco, Yves Lespérance, Hector J. Levesque)

Kaveh Raji

Universität Hamburg Fachbereich Informatik

Gliederung

- Motivation
- Was bedeutet „Iterated Believe Change“
- Situationskalkül (Definitionen)
- Belief Change / Belief Operator
- Beispiel
- Eigenschaften
- Diskussion

[Motivation]

- Ein intelligentes System sollte über Aktionen und die Wirkungen dieser „nachdenken“ können.
- Aktionen verändern die Welt oder lassen Agenten wahrnehmen.
- Aktionen sollten adäquate Glaubensveränderungen initiieren.

[Iterated Belief Change]

- Agenten sollen in einer Domäne bei ausreichender Information eine möglichst passende interne Repräsentation der Domäne enthalten.
- Veränderungen der Welt führen zu Veränderungen der Repräsentation.
- Mehrere Veränderungen sollen handhabbar sein.

[Iterated Belief Change]

- Die Autoren haben einen Ansatz entwickelt, basierend auf dem Situationskalkül.
- Revision und Update werden abgebildet.
- Belief change resultiert in dem Ansatz lediglich aus Aktionen somit wird iterated belief change unterstützt.
- Selbstbeobachtung und Selbstprüfung wird unterstützt.

[Situationskalkül]

- Der Situationskalkül ist eine Prädikatenlogische Sprache zur Repräsentation dynamischer Domänen.
- Eine Situation stellt eine „Momentaufnahme“ der Domäne dar.

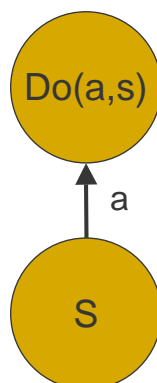
Situationskalkül

- Es existieren Anfangssituationen. Diese repräsentieren die Welt, wie der Agent sie anfänglich annimmt.
- Definition:

$$Init(s) \stackrel{def}{=} \neg \exists a, s' s = do(a, s')$$

Situationskalkül

- $do(a,s)$ beschreibt die Situation die entsteht, wenn in der Situation s die Aktion a ausgeführt wird.



[Situationskalkül]

- Aktionen werden in zwei Gruppen unterteilt.
- 📁 Aktionen die die Welt verändern.
- 📄 Sensorische Aktionen.

[Situationskalkül]

- Eine Erreichbarkeitsrelation für Situationen wird definiert.
- $B(s',s)$ falls in Situation s , die Situation s' als möglich angesehen wird.

[Situationskalkül]

- Eine andere Relation wird eingeführt: $SF(a,s)$.
- Diese Relation gilt falls in der Situation s der Sensor, der mit der Aktion a assoziiert wird, den Wert 1 liefert.

[Situationskalkül]

- Axiom:

$$B(s'', do(a, s)) \equiv$$

$$\exists s' [B(s', s) \wedge s'' = do(a, s') \wedge (SF(a, s') \equiv SF(a, s))]$$

Belief Change / Belief Operator

- Belief Update wird vollzogen wenn sich die Welt in der sich der Agent befindet verändert.
- Belief Revision behandelt den Fall, in dem sich die Welt nicht verändert. Hierbei ermöglichen neue Informationen dem Agenten die Schließung von Wissenslücken und die Eliminierung falscher Annahmen.

Belief Change / Belief Operator

- Scherl und Levesque haben einen Believe-Operator definiert:
- Definition:

$$Bel_{SL}(\phi, s) \stackrel{def}{=} \forall s' (B(s', s) \Rightarrow \phi[s'])$$

Belief Change / Belief Operator

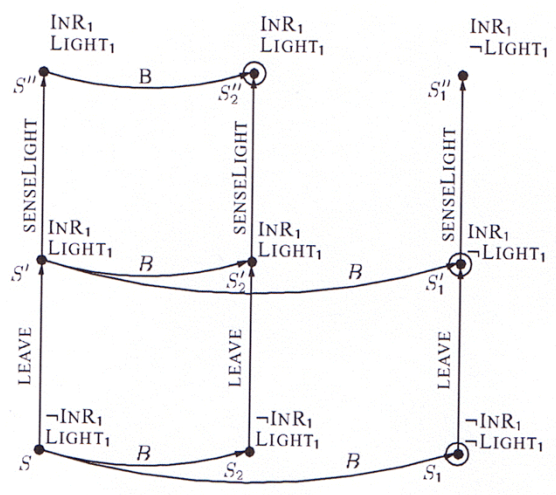
- Die Autoren definieren eine ähnliche Believe Relation.
- Jedoch werden die Situationen nach Plausibilität geordnet.
- Die Funktion $pl(s)$ weist jeder Situation s eine natürliche Zahl zu. Je geringer diese Zahl, desto plausibler die Situation.
- Axiom: $pl(do(a,s))=pl(s)$

Belief Change / Belief Operator

- Definition:

$$Bel(\phi, s) \stackrel{def}{=} \forall s' [B(s', s) \wedge (\forall s'' B(s'', s) \Rightarrow pl(s') \leq pl(s''))] \Rightarrow \phi[s']$$

Beispiel



- Es gibt zwei Räume R1 und R2.
- Light1 gibt an ob das Licht in R1 an ist.
- Es gibt die Aktionen SenseLight und Leave.
- InR1 gibt an ob der Agent in R1 ist oder nicht.

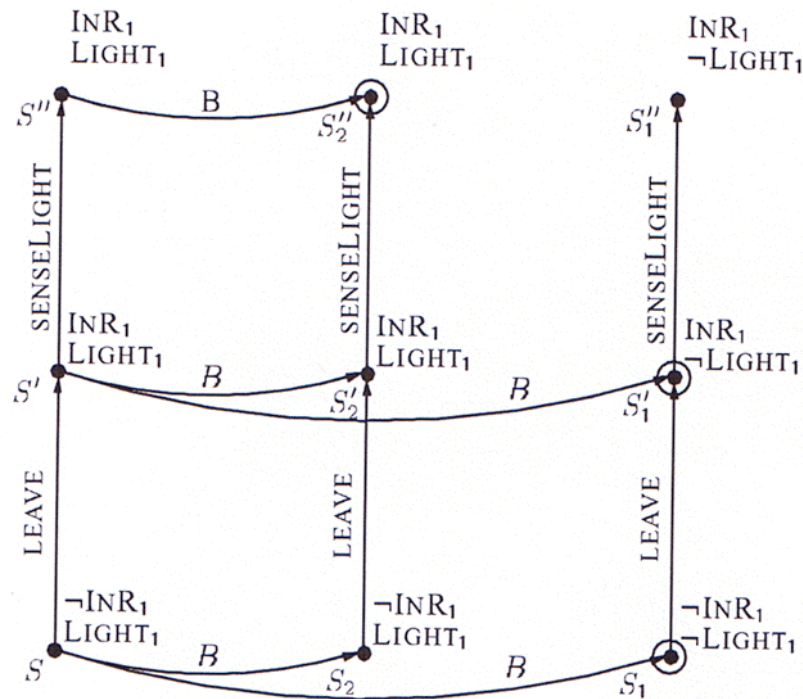
Situationskalkül

■ Formeln:

$$B(s'', do(a, s)) \equiv \exists s' [B(s', s) \wedge s'' = do(a, s') \wedge (SF(a, s') \equiv SF(a, s))]$$

$$Bel(\phi, s) \stackrel{def}{=} \forall s' [B(s', s) \wedge (\forall s'' B(s'', s) \Rightarrow pl(s') \leq pl(s''))] \Rightarrow \phi[s']$$

[Beispiel]



[Eigenschaften]

- Definition:

Eine Aktion A für die Revision der Formel ϕ erfüllt folgendes:

$$[\forall s SF(A, s) \Leftrightarrow \phi[s]] \wedge [\forall s \forall \vec{x} F(\vec{x}, s) \Leftrightarrow F(\vec{x}, do(A, s))]$$

Eigenschaften

- Es folgt:

$$[\forall s \phi[s] \Rightarrow Bel(\phi, do(A, s))] \wedge \\ \forall s \neg \phi[s] \Rightarrow Bel(\neg \phi, do(A, s))]$$

Eigenschaften

- Definition:

Eine Aktion A für das Updaten der Formel ϕ erfüllt folgendes:

$$\forall s \phi[do(A, s)] \wedge \forall s SF(A, s)$$

[Eigenschaften]

- Es folgt:

$$\forall s \text{ Bel}(\phi, do(A, s))$$