Epistemic Change in the Action Language Framework

Aaron Hunter
Simon Fraser University
amhunter@cs.sfu.ca

Problem Description

We extend the standard action language framework to reason about epistemic change, using $\mathcal{A}$ as the primary example. We emphasize the following goals.

- Flexible, intuitive extension:
  - represent direct and indirect epistemic effects
  - represent multi-agent action domains
  - minimize technical machinery introduced
- Generic extension:
  - suitable for a variety of action languages
  - suitable for both knowledge and belief
- Applications:
  - hypothetical reasoning
  - protocol verification

Related Work

Epistemic Extensions of $\mathcal{A}$

There have been two previous extensions suitable for reasoning about the knowledge of individual fluents.

- $\mathcal{A}_L$ - Lobo, Mendez and Taylor (2001)
  - A **causes to know** F **if** P
  - A **may cause** F
- $\mathcal{A}_S$ - Son and Baral (2001)
  - A **determines** F

Neither approach can be naturally extended to represent nested knowledge operators in a multi-agent domain.

Reasoning About Multi-Agent Action Effects

There are several related formalisms for reasoning about action effects in multi-agent domains. We give two recent examples.

- De Vos and Vermeir (2002)
  - choice logic programs for representing games
  - no explicit treatment of knowledge
- van Ditmarsch (2002)
  - logical approach to multi-agent epistemic action effects
  - difficult to translate into action language framework

Motivating Application: Cryptographic Protocol Verification

Cryptographic protocols can be viewed as sequences of actions with epistemic goals.

Example:
1. $A \rightarrow B$: random n encrypted with key K
2. $B \rightarrow A$: the number n unencrypted

Question:
Does A **know** that B is alive on the network?

A variety of formal tools have been employed for the verification of such protocols.

- Burrow, Abadi, Needham (1990)
  - protocol verification as reasoning in epistemic logic
  - explicit epistemic goals
  - implicit agents and actions
- Hernández-Orallo and Pinto(2000)
  - Situation Calculus
- Aiello and Massacci (2000)
  - Logic Programming

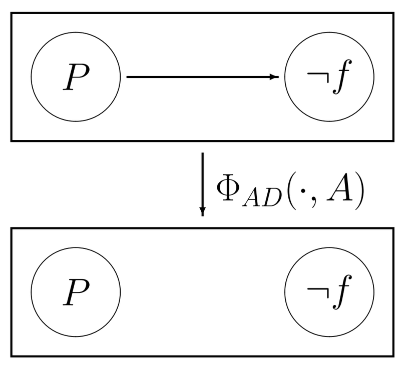
Solution

Modal Refinements

- Extend $\mathcal{A}$ with a modal operator in the underlying propositional logic
- If the underlying modal logic is standard epistemic logic, then let $\mathcal{A}[K]$ denote the modal extension
- With each action description AD, associate a transition function on Φ_{AD} on Kripke structures.
  1. For non-modal effects, update each state as in $\mathcal{A}$.
  2. For modal effects $\Box\phi$ with precondition P, remove edges of accessibility relation from P-states to $\neg\phi$-states.

Schematic Example

Let AD = $\{A$ **causes** $\Box f$ **if** $P\}$. Circled regions represent the set of worlds where the circled formula holds.



Note, in the case of knowledge, transitions of this form do not preserve symmetry. We need some additional restrictions.

Preservation Results

Definition
Let Π be a class of structures. A modal action description AD preserves Π if the transition function maps elements of Π to elements of Π.

Example
Let Π be the class of reflexive structures. Let AD be a set of propositions such that $P \models \phi$ for every rule in AD of the form

$$A \textbf{ causes } \Box\phi \textbf{ if } P.$$

Then AD preserves Π.

Definition
An action description is *admissible* in the epistemic extension of $\mathcal{A}$ just in case it preserves equivalence.

Providing sufficient conditions for natural preservation properties can be straightforward, but providing necessary conditions is not easy.

Comparison with Existing Work

In $\mathcal{A}_L$, the semantics is given by transition functions between sets of fluents. We can give a natural translation τ from $\mathcal{A}_L$ to $\mathcal{A}[K]$ such that we have the following result.

Theorem
Let AD be an action description, S a set of states A an action. The set of states S' is the $\mathcal{A}_L$ result of A if and only if S' is an equivalence class in the Kripke structure obtained from the transition function $\Phi_{\tau(AD)}$.

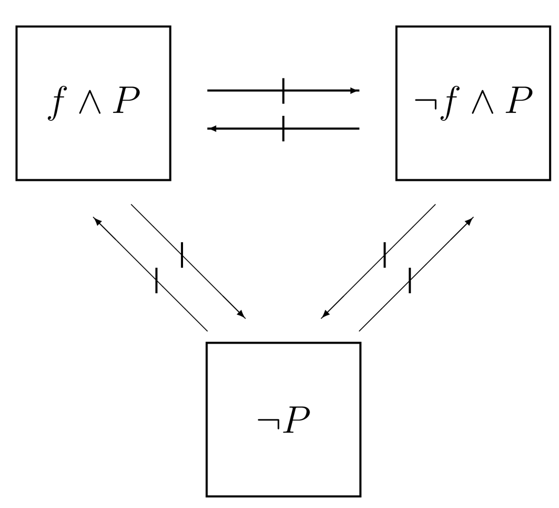
Example
Let

$$AD = \{A \textbf{ causes to know } f \textbf{ if } P\}$$

be an action description in the extension of Lobo et. al. The following is a provably equivalent $\mathcal{A}[K]$ description.

$$\{A \textbf{ causes } \Box(f \wedge P) \textbf{ if } f \wedge P,$$
$$A \textbf{ causes } \Box(\neg f \wedge P) \textbf{ if } \neg f \wedge P,$$
$$A \textbf{ causes } \Box\neg P \textbf{ if } \neg P\}$$

The intuition behind the proof is that both descriptions remove edges as follows.



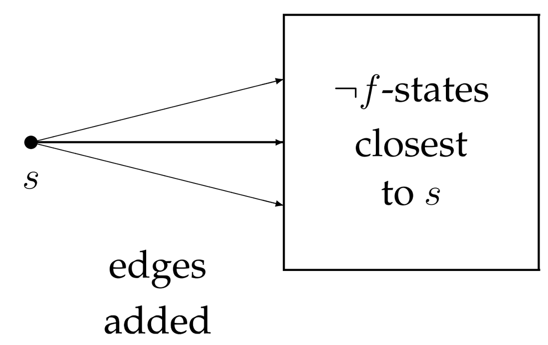
Future Thesis Work

Generalize With Distance Function

- Current approach only allows refinements to accessibility relations.
- Want to address knowledge loss through *expansions* of accessibility relations.
- We introduce a distance function on states to address this problem.
- Extend definition of Φ_{AD} with one more case.
  3. For negative modal effects $\neg\Box\phi$ with precondition P, add edges from P-states to $\neg\phi$ states that are minimally distant.

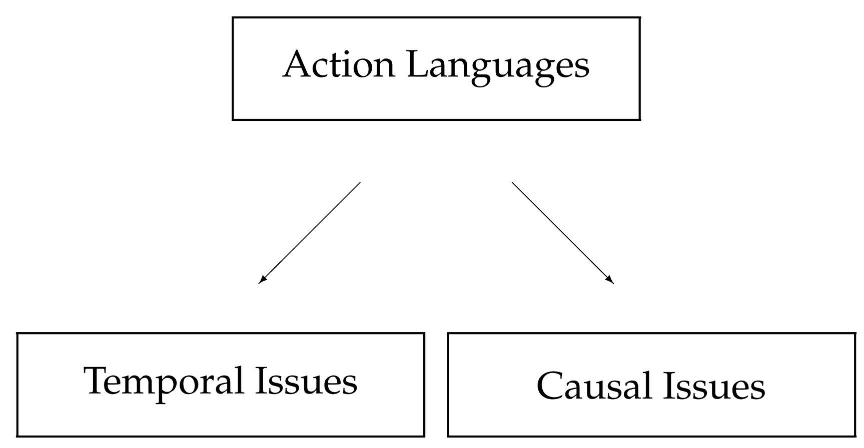
Schematic Example

Let AD=$\{A$ **causes** $\neg\Box f$ **if** $P\}$, and let s be a state in which P holds.



Unresolved Issues with Distance

For $\mathcal{A}$, it is sufficient to consider the Hamming distance. However, we would like to consider the problem in a more general setting. In the proceedings of AI'04, we suggest an abstract formulation of the action language framework that explicitly recognizes the following structure.



This framework can be improved by associating a distance function with each action language. This raises several questions.

- What is an acceptable distance function:
  - metric space?
  - distance between states or structures?
- What is an appropriate update mechanism:
  - add edges to all minimally distant worlds?
  - non-deterministically choose some minimally distance worlds?

Multiple Agents

We have focused on the knowledge of a single agent. We will extend our approach to deal with multi-agent action domains.

- Introduce multiple knowledge operators:
  - no interaction axioms required
  - consider strategy and cooperation
- Apply multi-agent framework to a simple protocol verification problem:
  - focus on providing flexible, intuitive representation
  - represent goals and actions in the same framework

Long Range Goals

We list some additional problems of interest, though we will not have the opportunity to address them all in the dissertation.

- Consider a specific non-epistemic modal extension in detail
  - doxastic logic: no major difference
  - deontic logc: preserving seriality a challenge
- Consider implementing a planner for multi-agent epistemic action languages