Computing Contingent Plans via Fully Observable Non-Deterministic Planning

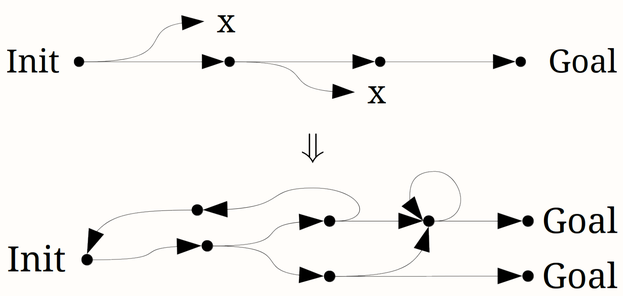
Christian Muise Vaishak Belle Sheila A. McIlraith
Department of Computer Science,
University of Toronto, Toronto, Canada.
{cjmuise,vaishak,sheila}@cs.toronto.edu

Motivation

- Many hard problems in AI have partially observable non-deterministic (POND) planning at its core.
- PRP enabled us to scale FOND planning significantly.
- Encodings have been proposed to solve contingent planning problems via non-deterministic techniques.
- Offline methods have become out-of-fashion due to the overwhelming size of the contingent plans generated.

PRP [1] $\Rightarrow$ PRP + Cond Effects [2] $\Rightarrow$ **PO-PRP** [3]

PO-PRP Algorithm



UpdatePolicy($\langle \mathcal{F}', s, \mathcal{G}', \mathcal{A}' \rangle$, P)

1. $\mathcal{A}'' = \textsc{Determinize}(\mathcal{A}')$; // Using all-outcomes
2. $[a_1, \cdots a_n] = \textsc{ComputePlan}(\langle \mathcal{F}', s, \mathcal{G}', \mathcal{A}'' \rangle)$;
3. For every suffix $[a_i, \cdots a_n]$ of the plan,
   1. $p_i = \textsc{Primf}^*(\mathcal{G}', [a_i, \cdots a_n])$;
   2. $c_i = cost([a_i, \cdots a_n])$;
   3. Add $\langle p_i, a_i, c_i \rangle$ to P;

Modifications to ComputePlan

- Invariants / ramifications left as actions.
- Priority is given to *useful* ramifications.
- Arbitrary ordering of ramifications enforced.

Pre-Image Filtering

Intuition: The partial state characterizing what must hold prior to executing a in order for the partial state p to hold, assuming a was executed in the context state s_c.

$$Added = \{l \mid \langle c, l \rangle \in \textsc{Eff} \text{ and } s_c \models c\}$$
$$Support = \textsc{Pre}_a \cup_{i=1\ldots n} \cup_{\langle c,l \rangle \in \textsc{Eff}_i} c$$
$$\textsc{Primf}(p, a, \textsc{Eff}, s_c) = (p \setminus Added) \cup Support$$

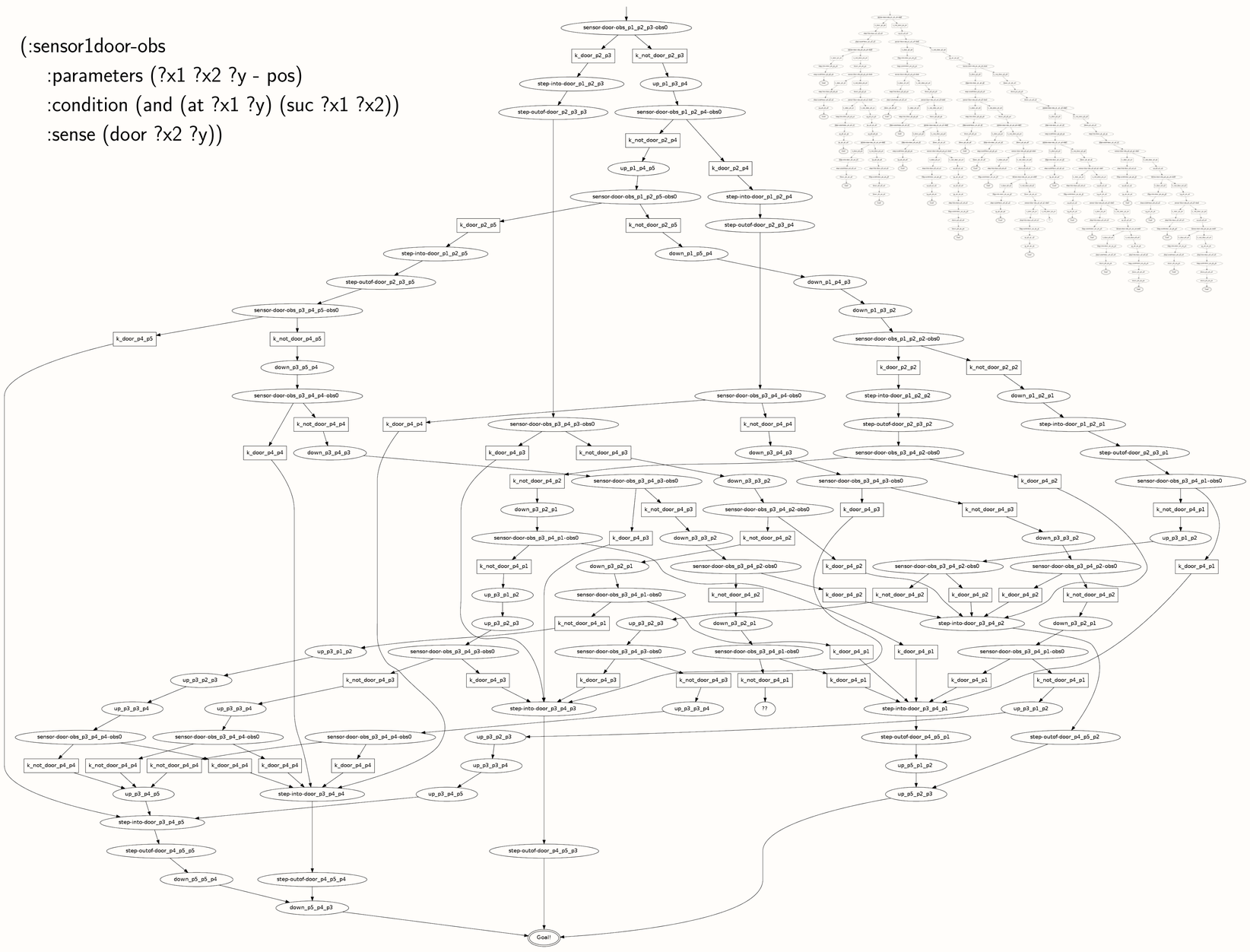
Properties:

- $s_c \models \textsc{Primf}(s, a, Eff, s_c)$
- $\textsc{Primf}(s, a, Eff, s_c) \models \textsc{Regress}(s, a, Eff)$
- When there are no conditional effects:
  $$\textsc{Primf}(s, a, Eff, s_c) = \textsc{Regress}(s, a, Eff)$$

Contingent $\longrightarrow$ FOND
(extension of Bonet & Geffner 2011)

- $\mathcal{F}'$: $\{Kl \mid l \in F\} \cup \{K\neg l \mid l \in F\}$
- $\mathcal{I}'$: $\{Kl \mid l \in \mathcal{I}\}$ $\mathcal{G}'$: $\{Kl \mid l \in \mathcal{G}\}$
- $\mathcal{A}'$: $\mathcal{A}'_{\mathcal{A}} \cup \mathcal{A}'_{\mathcal{O}} \cup \mathcal{A}'_{\mathcal{V}}$
  - $\mathcal{A}'_{\mathcal{A}}$: a' for every $a \in \mathcal{A}$ where, $\textsc{Pre}(a') = \{Kl \mid l \in \textsc{Pre}(a)\}$ and $\textsc{Eff}(a') = \{\langle Kc, Kl \rangle \mid \langle c, l \rangle \in \textsc{Eff}(a)\}$.
  - $\mathcal{A}'_{\mathcal{O}}$: for $o = \langle c, l \rangle \in \mathcal{O}$, there is an action $a' \in \mathcal{A}'_{\mathcal{O}}$ such that $\textsc{Pre}(a') = Kc \wedge \neg Kl \wedge \neg K\neg l$ with two possible non-deterministic effects: $\{\langle \top, Kl \rangle\}$ and $\{\langle \top, K\neg l \rangle\}$.
  - $\mathcal{A}'_{\mathcal{V}}$: for every $(c \supset l) \in \mathcal{I}^*$, there is an action $a' \in \mathcal{A}'_{\mathcal{V}}$ such that $\textsc{Pre}(a') = Kc$ and $\textsc{Eff}(a') = \{\langle \top, Kl \rangle\}$.

```
(:sensor1door-obs
    :parameters (?x1 ?x2 ?y - pos)
    :condition (and (at ?x1 ?y) (suc ?x1 ?x2))
    :sense (door ?x2 ?y))
```



Problem	Time (seconds)		Size (actions + sensing)	
	CLG	PO-PRP	CLG	PO-PRP
cballs-4-1	0.20	**0.02**	343	**261**
cballs-4-2	19.86	**0.67**	22354	**13887**
cballs-4-3	1693.02	**171.28**	1247512	**671988**
cballs-10-1	211.66	**1.57**	4829	**4170**
cballs-10-2	T	M	T	M
ctp-ch-1	**0.00**	**0.00**	5	**4**
ctp-ch-5	0.02	**0.00**	125	**16**
ctp-ch-10	2.2	**0.02**	4093	**31**
ctp-ch-15	133.24	**0.07**	131069	**46**
ctp-ch-20	T	**0.22**	T	**61**
doors-5	0.12	**0.01**	169	**82**
doors-7	3.50	**0.04**	2492	**1295**
doors-9	187.60	**1.07**	50961	**28442**
doors-11	T	M	T	M
wumpus-5	0.44	**0.16**	854	**233**
wumpus-7	9.28	**1.54**	7423	**770**
wumpus-10	1379.62	**11.17**	362615	**2669**
wumpus-15	T	**86.16**	T	**15628**
wumpus-20	T	M	T	M

Table: Compilation time and conditional plan size for CLG and PO-PRP. Bold indicates top performance, while T and M represent time limit or memory exceeded.

Key Considerations

- **Fairness Assumption:** How do we handle the assumption that the non-deterministic outcomes of a sensing action should be "fair"?
- **Reachable Incoherent States:** How do we handle belief states that will never occur in practice? (e.g., considering a sensing outcome that is not possible)
- **Solution Correspondence:** How do we interpret the policy that PRP generates for the original problem?

Summary

- Introduced PO-PRP: a planner capable of solving *simple* contingent planning problems.
- Identified key considerations for using FOND technology to solve encodings of contingent problems.
- PO-PRP pushes the limit of what is possible with offline contingent planning.

References

[1] **Improved Non-deterministic Planning by Exploiting State Relevance** Muise, C.; McIlraith, S. A.; and Beck, Christopher, J. In the 22nd International Conference on Automated Planning and Scheduling, 2012.

[2] **Non-Deterministic Planning With Conditional Effects** Muise, C.; McIlraith, S. A.; and Bell, V. In the 24th International Conference on Automated Planning and Scheduling, 2014.

[3] **Computing Contingent Plans via Fully Observable Non-Deterministic Planning** Muise, C.; Bell, V.; and McIlraith, S. A. In the 28th AAAI Conference on Artificial Intelligence, 2014.

http://www.haz.ca/research/poprp