From FOND to Probabilistic Planning: Guiding search for quality policies

Alberto Camacho[1] Christian Muise[2] Akshay Ganeshen[1] Sheila A. McIlraith[1]

[1]Department of Computer Science
University of Toronto, Canada
{acamacho,akshay,sheila}@cs.toronto.edu

[2]Department of Computing and Information Systems
University of Melbourne, Australia
christian.muise@unimelb.edu.au

HSDIP, ICAPS.
8 June 2015

UNIVERSITY OF TORONTO

THE UNIVERSITY OF MELBOURNE

Take away message

We leverage...

- core similarities between FOND and Probabilistic Planning.
- significant advances in state-of-the-art FOND Planning.

...to design Prob-PRP, a Probabilistic Planner that:

- is sound, and optimal in domains without unavoidable dead ends.
- is robust to small deviations in the probability distribution.
- is robust to action orderings.
- finds good quality solutions.

Motivation: No hasty decisions!

River problem [Little and Thiébaux, 2007]

The agent has two options to cross a river:

1. traverse a path of slippery rocks, with:
   - a 25% chance of success.
   - a 25% chance of slipping and falling into the river.
   - a 50% chance of reaching a small island. In this case she can swim, with:
     - an 80% chance of success.
     - a 20% chance of drowning.
2. swim from one side of the river to the other, with:
   - a 50% chance of success.
   - a 50% chance of falling in.

Motivation: No hasty decisions!

River problem [Little and Thiébaux, 2007]

The agent has two options to cross a river:

1. traverse a path of slippery rocks, with:
   - a 25% chance of success.
   - a 25% chance of slipping and falling into the river.
   - a 50% chance of reaching a small island. In this case she can swim, with:
     - an 80% chance of success.
     - a 20% chance of drowning.
2. swim from one side of the river to the other, with:
   - a 50% chance of success.
   - a 50% chance of falling in.

Option 2: 50% success.

Motivation: No hasty decisions!

River problem [Little and Thiébaux, 2007]

The agent has two options to cross a river:

1. traverse a path of slippery rocks, with:
   - a 25% chance of success.
   - a 25% chance of slipping and falling into the river.
   - a 50% chance of reaching a small island. In this case she can swim, with:
     - an 80% chance of success.
     - a 20% chance of drowning.
2. swim from one side of the river to the other, with:
   - a 50% chance of success.
   - a 50% chance of falling in.

Option 2: 50% success.
Option 1: 65% success.

Outline

Probabilistic Planning

We use a variant of the SAS$^+$ notation from [Mattmüller et al., 2010].

Probabilistic Planning Problem

A SAS$^+$ **probabilistic planning problem** is a tuple $\mathcal{P} = \langle \mathcal{V}, s_0, s_\star, \mathcal{A} \rangle$

- $\mathcal{V}$ finite set of variables v, each with domain $\mathcal{D}_\mathcal{V}$.
  - **complete state:** assignment $s : \mathcal{V} \to \mathcal{D}_\mathcal{V}$
  - $\mathcal{D}^+_\mathcal{V} = \mathcal{D}_\mathcal{V} \cup \{\bot\}$.
    partial state: (or simply, a state) assignment $s : \mathcal{V} \to \mathcal{D}^+_\mathcal{V}$
- s_0 is the **initial state**
- $s_\star$ is the **goal state**
- $\mathcal{A}$ set of actions $a = \langle Pre_a, Eff_a \rangle$
  - **Non determinism:** $Eff_a = \langle (p_1; Eff^1_a), \ldots, (p_n; Eff^n_a) \rangle$, $\Sigma_i p_i = 1$.
  - **Conditional effects:** $Eff^i_a = \{\langle cond_1, v_1, d_1 \rangle, \ldots, \langle cond_k, v_k, d_k \rangle\}$

Solutions to probabilistic planning problems are **policies**, or mappings $\pi(s)$ from states into actions.

MAXPROB

MAXPROB Probabilitic Planning

MAXPROB is the class of probabilistic planning problems where the **objective is to maximize** the **probability of success** (i.e., the probability of reaching eventually a state that satisfies the goal condition).

Optimal MAXPROB solution

A policy π is an **optimal solution** for the MAXPROB problem $\mathcal{P}$ if the probability of success by π is maximal.

FOND Planning

Fully-Observable Non-Deterministic (FOND) planning as a special case of MAXPROB Probabilistic planning:

- Exist solutions with probability of success 1.
- Transition probabilities are ignored.

Solutions are either **strong** o **strong cyclic** policies:
Strong Solutions guarantee of reaching the goal in a bounded number of actions.
Strong Cyclic Solutions guarantee of reaching the goal eventually.

Traditionally, FOND planning has preference for small-sized solutions.

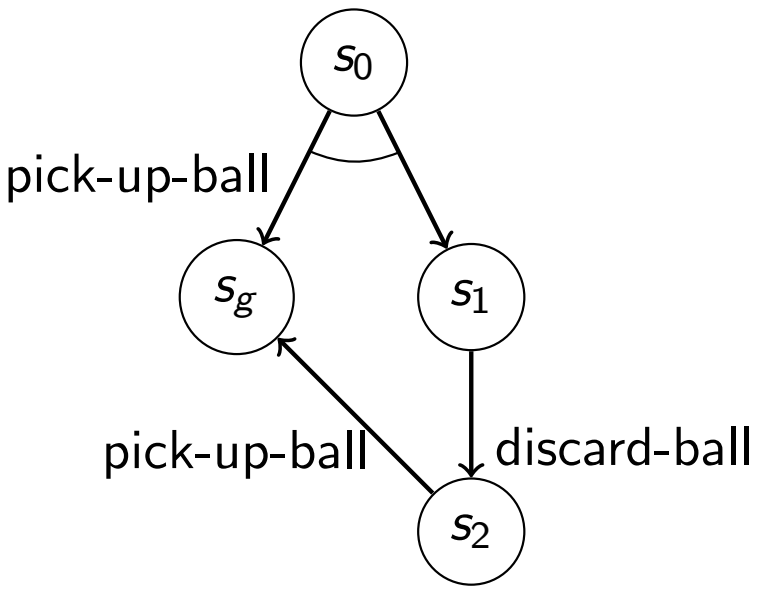
Desired Properties of the Solutions

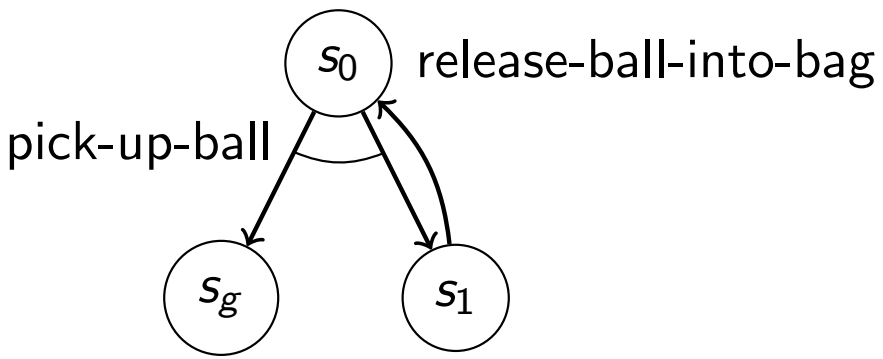
Optimal solutions maximize the probability of success. Other desired properties include:

- Small policy size
- Small expected plan length

Example

A bag with two balls: a red one, and a blue one. A robot can pick up one ball at a time, at random, from the bag. Afterwards, it can discard it or get it back into the bag. The goal for the robot is to hold the red ball.





Online vs. Offline Solutions

	Online	Offline
Reasoning	Incomplete Lookahead	Probabilistic Reasoning
Optimality Guarantees	Weak	Strong
Run Time	Fast	Slow

Outline

International Probabilistic Planning Competition (IPPC)

IPPC Edition	Problems	Winner
2004 & 2006	MAXPROB probabilistic problems	**FF-Replan** [Yoon et al., 2007]
2008	MAXREWARD probabilistic problems	**Robust-FF** [Teichteil-Königsbuch et al., 2010]
2011 & 2014	MAXREWARD MDPs	**PROST** [Keller and Eyerich, 2012]

All winners of the previous editions of the IPPC are online planners.

Robust-FF (RFF) [Teichteil-Konigsbuch, 2010]

RFF is The state of the art in MAXPROB probabilistic planning.

- Construct an **envelope** incrementally until Prob. failure $< \rho$.
- Replan when a state unhandled by the policy envelope is reached.

RFF offers different configurations:

1. **all-outcomes** determinization
   - probability of failure of the envelope guaranteed to be $< \rho$ in domains without deadends.
   - eventual goal reachability via replanning in domains without deadends.
   - Bad scalability.
2. **most-probable** determinization
   - probability of failure of the envelope not guaranteed.
   - goal reachability not guaranteed.

PRP [Muise et al., 2012]

PRP is the state of the art in FOND planning.

1. Compute a **deterministic plan** P to the goal.
2. **Regress** P to compute the **relevant** part of the states.
3. **forbidden state-action pairs** computed via regression from **dead ends**.

Theorem [Muise et al., 2012]

PRP is **sound**, and **complete** in problems without unavoidable deadends.

- **Compact** representation of the states via regression.
- Dead end detection and **forbidden state-action** mechanism to avoid deadends
- Fast, small solutions thanks to compact states representation.

Outline

From FOND to Probabilistic Planning

Given a MAXPROB planning problem, $\mathcal{P}$:

- FOND($\mathcal{P}$) is the FOND problem like $\mathcal{P}$ w/o probabilities.

Properties:

- π well-defined in FOND($\mathcal{P}$) $\Rightarrow$ π well-defined in $\mathcal{P}$.
- π strong cyclic for FOND($\mathcal{P}$) $\Rightarrow$ π optimal for $\mathcal{P}$.
- $\mathcal{P}$ has optimal solution that reaches the goal with probability 1 $\Leftrightarrow$ FOND($\mathcal{P}$) has a strong cyclic solution.

Well-Defined Policy

π is **well defined** when $\pi(s)$ is applicable in s.

From PRP to Prob-PRP

Plain Prob-PRP

Plain **Prob-PRP** is a call to PRP on the mapped problem FOND($\mathcal{P}$).

Theorem (Soundness and Completeness)

Plain Prob-PRP is **sound**, and **optimal** in problems with avoidable or no dead ends.

We enhance the plain version of Prob-PRP with two extensions.

1 Full Exploration in Last Iteration

Rationale:

- Forbidden state-action pairs $\langle s, a \rangle$ cut the search.
- There may still be a chance to reach the goal.

When plain Prob-PRP converges, the best incumbent policy found so far is explored in an extra iteration of the algorithm, with the forbidden state-action mechanism disabled.
The resulting policy potentially increases the probability of success.

2 Max. Likelihood Plan Exploration

In many domains:

- the non-determinism comes from faulty, low probable action effects.
- the preferred plans are short and have high likelihood.

Formally, the **likelihood** of a state-action plan $P = s_0, a_0, s_1, a_1, \ldots, a_{n-1}, s_n$ is defined as the product:

$$L_P = \Pi_{i=0}^{n-1} T(s_i, a_i, s_{i+1})$$

Prob-PRP skews the search of deterministic plans towards the ones that have maximum likelihood.

Outline

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (I/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
blocksworld-p01	**100**	23,2	18,0	0,02	1,00	**100**	**20,9**	**17**	**0,00**
blocksworld-p06	**100**	64,3	59,9	0,69	1,00	**100**	**50,6**	**43**	**0,16**
blocksworld-p12	**100**	**41,7**	**38,4**	0,68	1,01	**100**	68,4	61	**0,46**
blocksworld-p13	0	∞	117	16,5	1,00	**100**	125	**107**	**1,38**
blocksworld-p14	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,40**
blocksworld-p15	0	∞	117	16,6	1,00	**100**	**125**	**107**	**1,38**
boxworld-p01	**100**	**29,4**	**49,3**	0,43	1,27	**100**	31,3	57	**0,06**
boxworld-p06	**100**	**64,9**	**166**	13,0	1,35	**100**	68,1	266	**2,34**
boxworld-p09	**100**	64,7	**132**	7,56	1,38	**100**	**62,5**	207	**1,84**
boxworld-p12	**100**	**74,1**	**199**	23,3	1,36	**100**	102	415	**18,0**
boxworld-p13	0	∞	344	35,6	1,94	**100**	178	906	130
boxworld-p14	0	∞	325	34,9	1,83	**100**	177	906	157
boxworld-p15	0	∞	347	35,2	1,96	**100**	177	906	160
triangle-tireworld-p02	**100**	13,2	80,7	0,17	1,32	**100**	**12,0**	**23**	**0,00**
triangle-tireworld-p04	**100**	29,6	248	1,76	1,20	**100**	**25,1**	**55**	**0,06**
triangle-tireworld-p06	**100**	45,6	490	7,98	1,14	**100**	**37,9**	**95**	**0,22**
triangle-tireworld-p08	**100**	61,5	958	36,5	1,19	**100**	**50,9**	**143**	**0,72**
triangle-tireworld-p10	**100**	77,6	1595	111	1,21	**100**	**64,0**	**199**	**2,38**

Quality of the policies (II/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
schedule-p02	**100**	59,2	**5,00**	**0,01**	1,00	**100**	**51,0**	7	0,04
schedule-p03	**100**	100	**5,00**	**0,01**	1,00	**100**	**95,0**	7	0,12
schedule-p04	96	57,8	14,3	0,02	1,12	**100**	**46,9**	21	0,14
schedule-p05	89	116	14,5	0,03	1,15	**100**	**92,0**	16	0,18
schedule-p06	**45**	**364**	**141**	**1,42**	3,01	–	–	m	–
schedule-p07	**36**	**390**	**146**	**1,34**	3,11	–	–	m	–
schedule-p08	**34**	**354**	**146**	**3,94**	3,17	–	–	m	–
schedule-p09	**4**	**402**	**317**	**3,17**	4,31	–	–	m	–
ex-blocksworld-p01	60	8,0	20,8	0,05	1,06	**100**	**8,0**	**9**	**0,00**
ex-blocksworld-p05	**100**	**6,0**	11,6	**0,01**	1,09	**100**	**6,0**	**11**	0,02
ex-blocksworld-p06	90	12,6	62,2	0,10	1,35	**96**	20,7	**28**	0,34
ex-blocksworld-p09	13	25,2	95,7	1,07	1,23	–	–	–	t
ex-blocksworld-p10	2	36,0	76,8	0,97	1,24	**3**	116	105	14,3
ex-blocksworld-p11	**13**	**32,0**	92,7	**1,59**	1,31	**13**	93,4	**82**	7,42
ex-blocksworld-p12	1	38,0	96,6	2,15	1,21	**2**	91,5	**78**	6,28
ex-blocksworld-p13	**10**	**59,2**	**451**	**5,76**	1,45	–	–	–	t
ex-blocksworld-p14	**0**	**0**	**130**	**116**	1,24	–	–	–	t
ex-blocksworld-p15	**9**	**43,6**	**172**	**8,91**	1,28	–	–	–	t

Quality of the policies (II/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
schedule-p02	**100**	59,2	**5,00**	**0,01**	1,00	**100**	**51,0**	7	0,04
schedule-p03	**100**	100	**5,00**	**0,01**	1,00	**100**	**95,0**	7	0,12
schedule-p04	96	57,8	14,3	0,02	1,12	**100**	**46,9**	21	0,14
schedule-p05	89	116	14,5	0,03	1,15	**100**	**92,0**	16	0,18
schedule-p06	**45**	**364**	**141**	**1,42**	3,01	–	–	m	–
schedule-p07	**36**	**390**	**146**	**1,34**	3,11	–	–	m	–
schedule-p08	**34**	**354**	**146**	**3,94**	3,17	–	–	m	–
schedule-p09	**4**	**402**	**317**	**3,17**	4,31	–	–	m	–
ex-blocksworld-p01	60	8,0	20,8	0,05	1,06	**100**	**8,0**	**9**	**0,00**
ex-blocksworld-p05	**100**	**6,0**	11,6	**0,01**	1,09	**100**	**6,0**	**11**	0,02
ex-blocksworld-p06	90	12,6	62,2	0,10	1,35	**96**	20,7	**28**	0,34
ex-blocksworld-p09	13	25,2	95,7	1,07	1,23	–	–	–	t
ex-blocksworld-p10	2	36,0	76,8	0,97	1,24	**3**	116	105	14,3
ex-blocksworld-p11	**13**	**32,0**	92,7	**1,59**	1,31	**13**	93,4	**82**	7,42
ex-blocksworld-p12	1	38,0	96,6	2,15	1,21	**2**	91,5	**78**	6,28
ex-blocksworld-p13	**10**	**59,2**	**451**	**5,76**	1,45	–	–	–	t
ex-blocksworld-p14	**0**	**0**	**130**	**116**	1,24	–	–	–	t
ex-blocksworld-p15	**9**	**43,6**	**172**	**8,91**	1,28	–	–	–	t

Quality of the policies (II/II)

	RFF					Prob-PRP			
Problem	%	L	S	T	R	%	L	S	T
schedule-p02	**100**	59,2	**5,00**	**0,01**	1,00	**100**	**51,0**	7	0,04
schedule-p03	**100**	100	**5,00**	**0,01**	1,00	**100**	**95,0**	7	0,12
schedule-p04	96	57,8	14,3	0,02	1,12	**100**	**46,9**	21	0,14
schedule-p05	89	116	14,5	0,03	1,15	**100**	**92,0**	16	0,18
schedule-p06	**45**	**364**	**141**	**1,42**	3,01	–	–	m	–
schedule-p07	**36**	**390**	**146**	**1,34**	3,11	–	–	m	–
schedule-p08	**34**	**354**	**146**	**3,94**	3,17	–	–	m	–
schedule-p09	**4**	**402**	**317**	**3,17**	4,31	–	–	m	–
ex-blocksworld-p01	60	8,0	20,8	0,05	1,06	**100**	**8,0**	**9**	**0,00**
ex-blocksworld-p05	**100**	**6,0**	11,6	**0,01**	1,09	**100**	**6,0**	**11**	0,02
ex-blocksworld-p06	90	12,6	62,2	0,10	1,35	**96**	20,7	**28**	0,34
ex-blocksworld-p09	13	25,2	95,7	1,07	1,23	–	–	–	t
ex-blocksworld-p10	2	36,0	76,8	0,97	1,24	**3**	116	105	14,3
ex-blocksworld-p11	**13**	**32,0**	92,7	**1,59**	1,31	**13**	93,4	**82**	7,42
ex-blocksworld-p12	1	38,0	96,6	2,15	1,21	**2**	91,5	**78**	6,28
ex-blocksworld-p13	**10**	**59,2**	**451**	**5,76**	1,45	–	–	–	t
ex-blocksworld-p14	**0**	**0**	**130**	**116**	1,24	–	–	–	t
ex-blocksworld-p15	**9**	**43,6**	**172**	**8,91**	1,28	–	–	–	t

Shortening Plan Length

	Prob-PRP$_{uc}$				Prob-PRP			
problem	%	T	S	L	%	T	S	L
blocksworld-p01	**100**	0,02	21	24	**100**	**0,00**	**17**	**19**
blocksworld-p06	**100**	**0,14**	**35**	**39**	**100**	0,16	43	47
blocksworld-p12	**100**	**0,46**	71	75	**100**	0,46	**61**	**65**
blocksworld-p15	**100**	**1,38**	110	119	**100**	1,40	**107**	**115**
boxworld-p01	**100**	0,14	**57**	156	**100**	**0,06**	**57**	**32**
boxworld-p06	**100**	6,56	269	363	**100**	**2,44**	**266**	**69**
boxworld-p12	**100**	**9,86**	**301**	328	**100**	18,0	415	**102**
boxworld-p15	**100**	586	949	1000+	**100**	**159**	**906**	**178**
ex-blocksworld-p01	**100**	**0,00**	**9**	**9**	**100**	**0,00**	**9**	**9**
ex-blocksworld-p06	96,3	0,68	**25**	**22**	**96,8**	**0,32**	28	**22**
ex-blocksworld-p10	**10,2**	**4,08**	**52**	28	4,6	14,0	105	**26**
ex-blocksworld-p11	9,6	33,8	89	29	**19,2**	**7,20**	**82**	**27**
ex-blocksworld-p12	–	–	m	–	**2,4**	**5,90**	**78**	**17**
schedule-p02	**100**	**0,04**	**7**	**48**	**100**	**0,04**	**7**	**48**
schedule-p03	**100**	**0,12**	**7**	**87**	**100**	**0,12**	**7**	**87**
schedule-p04	**100**	**0,08**	**16**	**43**	**100**	**0,14**	21	46
schedule-p05	**100**	**0,18**	**16**	96	**100**	0,20	16	**95**

Robustness to Probability Perturbations

	RFF		Prob-PRP	
problem	% sol	% sim	% sol	% sim
triangle-tireworld-p01	56.7	53.4	**100**	**100**
triangle-tireworld-p02	16.8	12.9	**100**	**100**
triangle-tireworld-p03	4.9	3.2	**100**	**100**
triangle-tireworld-p04	1.8	0.9	**100**	**100**
triangle-tireworld-p05	0.5	0.2	**100**	**100**
triangle-tireworld-p06	0.0	0.0	**100**	**100**
triangle-tireworld-p07	0.0	0.0	**100**	**100**
triangle-tireworld-p08	0.1	0.0	**100**	**100**
triangle-tireworld-p09	0.0	0.1	**100**	**100**
triangle-tireworld-p10	0.0	0.0	**100**	**100**

Outline

Conclusions and Future Work

Conclusions:

- The **core similarities** between the FOND and Probabilistic planning models makes it possible to apply search techniques of FOND to solve MAXPROB probabilistic problems.
- Prob-PRP finds MAXPROB solutions with **improved guarantees** wrt the previous state of the art.
- Prob-PRP finds MAXPROB **offline solutions** that have comparable or better **quality** than the online solutions computed by the previous state of the art.

Future Work:

- Anytime behaviour
- Improved Probabilistic reasoning
- Exploration of new heuristics and search algorithms.

Take away message

We leverage...

- core similarities between FOND and Probabilistic Planning.
- significant advances in state-of-the-art FOND Planning.

...to design Prob-PRP, a Probabilistic Planner that:

- is sound, and optimal in domains without unavoidable dead ends.
- is robust to small deviations in the probability distribution.
- is robust to action orderings.
- finds good quality solutions.

Any Questions?

Benchmarks, code, and slides available soon at:
`http://www.haz.ca/research/probprp/`

References I


Hertle, A., Dornhege, C., Keller, T., Mattmller, R., Ortlieb, M., and Nebel, B. (2014).
An Experimental Comparison of Classical, FOND and Probabilistic Planning.
In *Proc. of 37th International Conference on Artificial Intelligence (KI 2014)*, Prague.


Keller, T. and Eyerich, P. (2012).
PROST: Probabilistic Planning Based on UCT.
ICAPS.


Little, I. and Thiébaux, S. (2007).
Probabilistic planning vs. replanning.
ICAPS Workshop on IPC: Past, Present and Future.


Mattmüller, R., Ortlieb, M., Helmert, M., and Bercher, P. (2010).
Pattern database heuristics for fully observable nondeterministic planning.
In *ICAPS*, pages 105–112.


Muise, C., McIlraith, S. A., and Beck, J. C. (2012).
Improved Non-deterministic Planning by Exploiting State Relevance.
In *ICAPS*, pages 172–180.


Teichteil-Königsbuch, F., Kuter, U., and Infantes, G. (2010).
Incremental plan aggregation for generating policies in MDPs.
In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1*, number 1, pages 1231–1238.

References II

Yoon, S., Fern, A., and Givan, R. (2007).
FF-Replan: A baseline for probabilistic planning.
In *ICAPS*, pages 352–359.