Structured Possibilistic Planning using Decision Diagrams

Nicolas Drougard, Florent Teichteil-Königsbuch, Jean-Loup Farges and Didier Dubois.

AAAI-14 – *July 31st 2014*

Mixed-Observable Markov Decision Process (MOMDP)

(*Ong et al., RSS-09*)

model for **sequential decision making under uncertainty**.

Mixed-Observable Markov Decision Process (MOMDP)

(*Ong et al., RSS-09*)

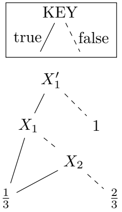
model for **sequential decision making under uncertainty**.

- real problems may be huge $\longrightarrow$ need for compact representation;
- probabilities may not be known $\longrightarrow$ need for flexible modelling.

We adopt a possibilistic model: π-MOMDP,
counterpart of MOMDP from *Qualitative Possibility Theory*;

We adopt a possibilistic model: π-MOMDP,
counterpart of MOMDP from *Qualitative Possibility Theory*;

Algebraic Decision Diagrams (ADDs)
symbolically represent transition functions.

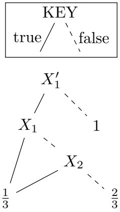


example of ADD

We adopt a possibilistic model: π-MOMDP,
counterpart of MOMDP from *Qualitative Possibility Theory*;

Algebraic Decision Diagrams (ADDs)
symbolically represent transition functions.

- probabilistic model: $+$ and $\times$ used (e.g. for $\mathbb{E}$)
  $\Rightarrow$ new values created,
  number of ADD leaves **potentially huge**.



example of ADD

- possibilistic model: $\min$ and $\max$ used
  $\Rightarrow$ finite number of values manipulated, number of leaves bounded,
  ADDs smaller in practice.

Qualitative Possibility theory

Events $e \subset \Omega$ **sorted** using **possibility degrees** $\pi(e) \in \mathcal{L}$,
where $\mathcal{L}$ finite scale such as $\left\{ 0, \frac{1}{n}, \frac{2}{n}, \ldots, 1 \right\}$

$$\neq$$

quantified with **frequencies** $\mathbf{p}(e) \in [0,1]$ (probabilities)

- $\pi(e_1) \leqslant \pi(e_2)$ only means "e_1 is less plausible than e_2";

Qualitative Possibility theory

Events $e \subset \Omega$ **sorted** using **possibility degrees** $\pi(e) \in \mathcal{L}$,
where $\mathcal{L}$ finite scale such as $\left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, 1 \right\}$

$$\neq$$

quantified with **frequencies** $\mathbf{p}(e) \in [0,1]$ (probabilities)

- $\pi(e_1) \leqslant \pi(e_2)$ only means "e_1 is less plausible than e_2";

Qualitative Possibilistic translation

$$\Updownarrow$$

approximation that requires less knowledge.

Qualitative Possibility theory

Events $e \subset \Omega$ **sorted** using **possibility degrees** $\pi(e) \in \mathcal{L}$, where $\mathcal{L}$ finite scale such as $\left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, 1 \right\}$

$$\neq$$

quantified with **frequencies** $\mathbf{p}(e) \in [0,1]$ (probabilities)

- $\pi(e_1) \leqslant \pi(e_2)$ only means "e_1 is less plausible than e_2";

Qualitative Possibilistic translation

$$\Updownarrow$$

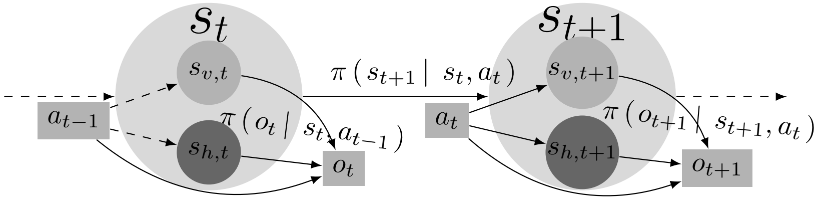
approximation that requires less knowledge.

Probability ($\mathbb{P}$) / possibility (Π) theories. $e_1 \neq e_2$, two events $\subset \Omega$:

e_1 or e_2 (union)	$\mathbb{P}$: $\mathbf{p}(e_1) + \mathbf{p}(e_2 \cap \overline{e_1})$	Π: $\max\left\{ \pi(e_1), \pi(e_2) \right\}$
e_1 and e_2 (intersect.)	$\mathbb{P}$: $\mathbf{p}(e_1).\mathbf{p}\left(e_2 \mid e_1\right)$	Π: $\min\left\{ \pi(e_1), \pi(e_2 \mid e_1) \right\}$

π-MOMDPs

system $=$ $\{$ an agent selecting actions $a \in \mathcal{A}$, its environment $\}$;

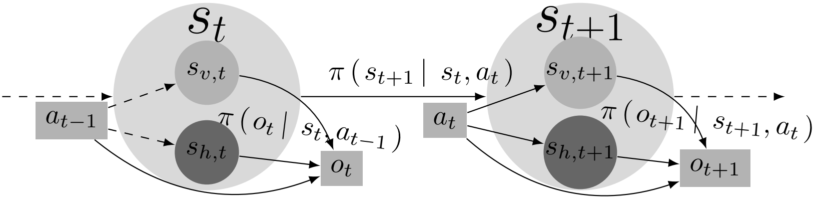


π-MOMDPs

system $=$ $\{$ an agent selecting actions $a \in \mathcal{A}$, its environment $\}$;

system state at step t: $s_t = (s_{v,t}, s_{h,t})$

- $s_v \in \mathcal{S}_v$ visible to the agent;
- $s_h \in \mathcal{S}_h$ hidden but producing observation $o_t \in \mathcal{O}$ with possibility degree $\pi\left(o_t \mid s_t, a_{t-1}\right) \in \mathcal{L}$.



π-MOMDPs

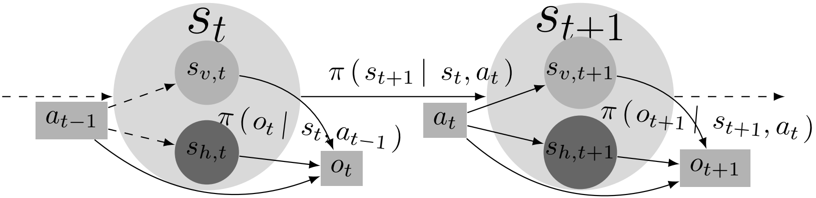
system $=$ $\{$ an agent selecting actions $a \in \mathcal{A}$, its environment $\}$;

system state at step t: $s_t = (s_{v,t}, s_{h,t})$

- $s_v \in \mathcal{S}_v$ visible to the agent;
- $s_h \in \mathcal{S}_h$ hidden but producing observation $o_t \in \mathcal{O}$ with possibility degree $\pi\left(o_t \mid s_t, a_{t-1}\right) \in \mathcal{L}$.

- Next state s_{t+1} arises with possibility degree $\pi\left(s_{t+1} \mid s_t, a_t\right)$.
- Goal of the problem modeled with preference $\mu(s) \in \mathcal{L}$.
- History: $h_t = \{o_1, \ldots, o_t, s_{v,0}, \ldots, s_{v,t}, a_0, \ldots, a_{t-1}\}$.

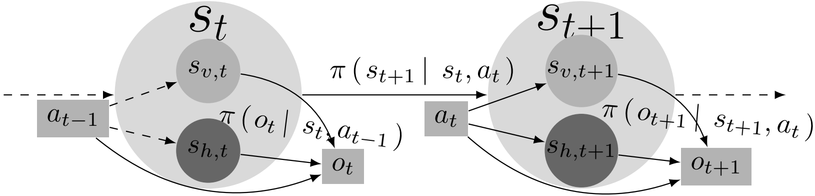
problem solving $=$ build action **strategy** to achieve a mission.



Bayes' rule + current observation $o \in \mathcal{O}$

$\rightarrow$ process **visible to the agent** $x = (s_v, \beta) \in \mathcal{X} = \mathcal{S}_v \times B^\pi$

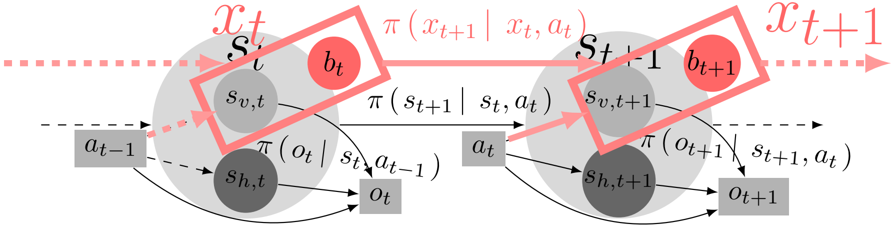
$\beta \in B^\pi$ *belief i.e.* possibility distribution over s_h: $\beta_t(s_h) = \pi\left(s_h \mid h_t\right)$.



Bayes' rule + current observation $o \in \mathcal{O}$
$\rightarrow$ process **visible to the agent** $x = (s_v, \beta) \in \mathcal{X} = \mathcal{S}_v \times B^\pi$
$\beta \in B^\pi$ *belief i.e.* possibility distribution over s_h: $\beta_t(s_h) = \pi\left(s_h \mid h_t\right)$.

- New transition function $\pi\left(x' \mid x, a\right)$ and preference $\mu(x)$.
- Resolution space $\mathcal{X}$ **finite** thanks to finite possibilistic scale $\mathcal{L}$.

$\Rightarrow$ finite state space π-MDP! (probabilistic MOMDP: continuous space)



Bayes' rule + current observation $o \in \mathcal{O}$
$\rightarrow$ process **visible to the agent** $x = (s_v, \beta) \in \mathcal{X} = \mathcal{S}_v \times B^\pi$
$\beta \in B^\pi$ *belief i.e.* possibility distribution over s_h: $\beta_t(s_h) = \pi\left(s_h \mid h_t\right)$.

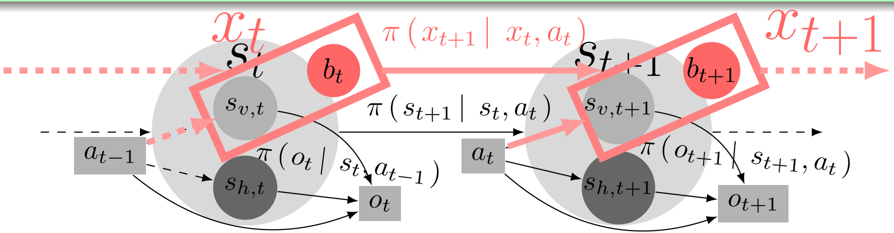
- New transition function $\pi\left(x' \mid x, a\right)$ and preference $\mu(x)$.
- Resolution space $\mathcal{X}$ **finite** thanks to finite possibilistic scale $\mathcal{L}$.

$\Rightarrow$ finite state space π-MDP! (probabilistic MOMDP: continuous space)

Strategy $\delta : \mathcal{X} \rightarrow \mathcal{A}$ computed using **value iteration**

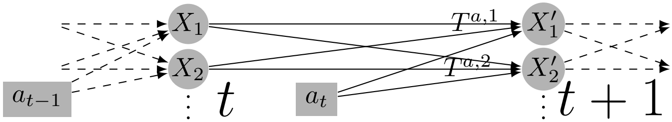
$$V^*_{h+1}(x) = \max_a \max_{x'} \min \left\{ \pi(x' \mid x, a), V_h(x') \right\} \text{ with } V^*_0 = \mu:$$

$\delta(x)$ is the last $a \in \mathcal{A}$ increasing $V^*_h(x)$. (*Drougard et al., UAI-13*)



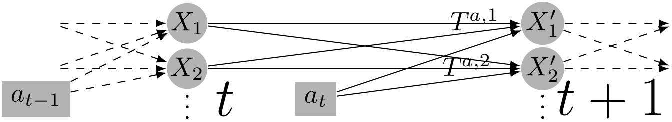
We introduce factored π-MOMDP model

factored π-MOMDP
$\Leftrightarrow$ resolution space $\mathcal{X}$ described with binary variables $(X_1, \ldots, X_n)$
$+$ dynamic fits in with Dynamic Bayesian Network below.

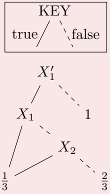


We introduce factored π-MOMDP model

factored π-MOMDP
$\Leftrightarrow$ resolution space $\mathcal{X}$ described with binary variables $(X_1, \dots, X_n)$
$+$ dynamic fits in with Dynamic Bayesian Network below.



- Transition functions $T_i^a = \pi\left(X'_i \mid parents(X_i), a \right)$ **represented with Algebraic Decision Diagrams** on finite $\mathcal{L} \subset \mathbb{N}$.
(see SPUDD – *Hoey et al., UAI-99*).



example of ADD

Proposition: value function qualitative aggregation

$V_h^* : \{0,1\}^p \rightarrow \mathcal{L}$ value function, $a \in \mathcal{A}$ action.

define: 1) $q_0^a = V_h^*(X'_1, \cdots, X'_n)$;

2) for $i : 1..n$, $q_i^a = \max\limits_{X'_i \in \{0,1\}} \min \left\{ \left(X'_i \mid parents(X'_i), a \right), q_{i-1}^a \right\}$;

Then, $V_{h+1}^* = \max\limits_{a \in \mathcal{A}} q_n^a$.

Proposition: value function qualitative aggregation

$V_h^*: \{0,1\}^p \rightarrow \mathcal{L}$ value function, $a \in \mathcal{A}$ action.

define: 1) $q_0^a = V_h^*(X_1', \cdots, X_n')$;

2) for $i : 1..n$, $q_i^a = \max\limits_{X_i' \in \{0,1\}} \min\left\{ \left(X_i' \mid parents(X_i'), a \right), q_{i-1}^a \right\}$;

Then, $V_{h+1}^* = \max\limits_{a \in \mathcal{A}} q_n^a$.

Algorithm PPUDD: Possibilistic Planning Using Decision Diagrams

$V^* \leftarrow 0$; $V^c \leftarrow \mu$; $\delta \leftarrow \overline{a}$;
while $V^* \neq V^c$ **do**
  $V^* \leftarrow V^c$;
  for $a \in \mathcal{A}$ **do**
    $q^a \leftarrow$ swap each X_i variable in V^* with X_i' ;
    for $1 \leqslant i \leqslant n$ **do**
      $q^a \leftarrow \boxed{\min} \{ q^a, \pi(X_i' \mid parents(X_i'), a) \}$;
      $q^a \leftarrow \boxed{\max}_{X_i'} q^a$;
    $V^c \leftarrow \boxed{\max} \{ q^a, V^c \}$;
    update δ to a where $q^a = V^c$ and $V^c > V^*$;
return (V^*, δ) ;

Proposition: value function qualitative aggregation

$V_h^* : \{0,1\}^p \rightarrow \mathcal{L}$ value function, $a \in \mathcal{A}$ action.

define: 1) $q_0^a = V_h^*(X_1', \cdots, X_n')$;

2) for $i : 1..n$, $q_i^a = \max\limits_{X_i' \in \{0,1\}} \min \left\{ \left(X_i' \mid parents(X_i'), a \right), q_{i-1}^a \right\}$;

Then, $V_{h+1}^* = \max\limits_{a \in \mathcal{A}} q_n^a$.

Algorithm PPUDD: Possibilistic Planning Using Decision Diagrams

$V^* \leftarrow 0$; $V^c \leftarrow \mu$; $\delta \leftarrow \overline{a}$;
while $V^* \neq V^c$ **do**
    $V^* \leftarrow V^c$;
    for $a \in \mathcal{A}$ **do**
        $q^a \leftarrow$ swap each X_i variable in V^* with X_i' ;
        for $1 \leqslant i \leqslant n$ **do**
            $q^a \leftarrow \boxed{\min}\{q^a, \pi(X_i' \mid parents(X_i'), a)\}$;
            $q^a \leftarrow \boxed{\max}_{X_i'} q^a$;
        $V^c \leftarrow \boxed{\max}\{q^a, V^c\}$;
        update δ to a where $q^a = V^c$ and $V^c > V^*$;
return (V^*, δ) ;

factorization
$+$
previous proposition,

$\Rightarrow$ value iterations

Proposition: value function qualitative aggregation

$V^*_h : \{0,1\}^p \rightarrow \mathcal{L}$ value function, $a \in \mathcal{A}$ action.

define: 1) $q^a_0 = V^*_h(X'_1, \cdots, X'_n)$;

2) for $i : 1..n$, $q^a_i = \max\limits_{X'_i \in \{0,1\}} \min \left\{ \left(X'_i \mid parents(X'_i), a \right), q^a_{i-1} \right\}$;

Then, $V^*_{h+1} = \max\limits_{a \in \mathcal{A}} q^a_n$.

Algorithm PPUDD: Possibilistic Planning Using Decision Diagrams

$V^* \leftarrow 0$; $V^c \leftarrow \mu$; $\delta \leftarrow \overline{a}$;
while $V^* \neq V^c$ **do**
  $V^* \leftarrow V^c$;
  for $a \in \mathcal{A}$ **do**
    $q^a \leftarrow$ swap each X_i variable in V^* with X'_i ;
    for $1 \leqslant i \leqslant n$ **do**
      $q^a \leftarrow \boxed{\min}\{q^a, \pi(X'_i \mid parents(X'_i), a)\}$;
      $q^a \leftarrow \boxed{\max}_{X'_i} q^a$;
    $V^c \leftarrow \boxed{\max}\{q^a, V^c\}$;
    update δ to a where $q^a = V^c$ and $V^c > V^*$;
return (V^*, δ) ;

factorization
+
previous proposition,

$\Rightarrow$ value iterations

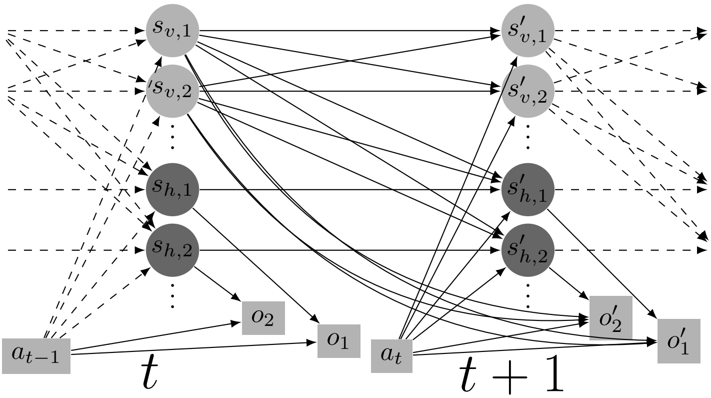
divided into n steps

handled ADDs smaller

$\rightarrow$ **speeds up computation.**

Independent beliefs

A π-MOMDP fulfilling assumptions of the Dynamic Bayesian Network below has a **natural factorization**:
$(s_v, \beta) = (s_{v,1}, \ldots, s_{v,m}, \beta_1, \ldots, \beta_l)$, with β_i belief about $s_{h,i}$.

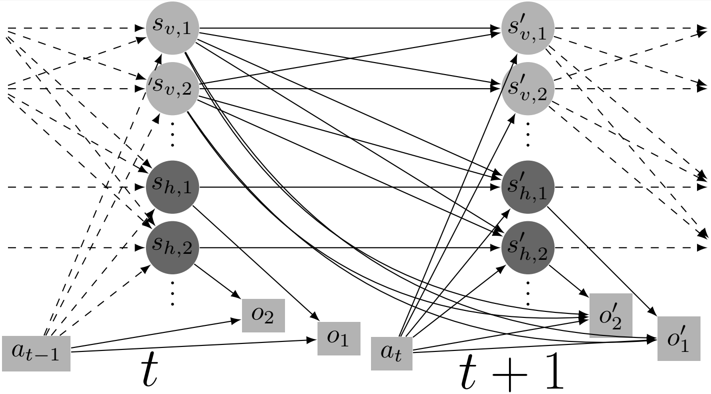


Independent beliefs

A π-MOMDP fulfilling assumptions of the Dynamic Bayesian Network below has a **natural factorization**:

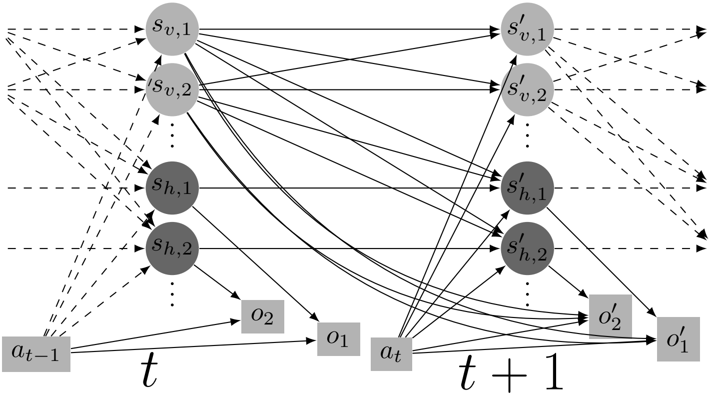
$(s_v, \beta) = (s_{v,1}, \dots, s_{v,m}, \beta_1, \dots, \beta_l)$, with β_i belief about $s_{h,i}$.

some assumptions: one observation variable for each hidden state variable, hidden state variables independent on other hidden state variables ...



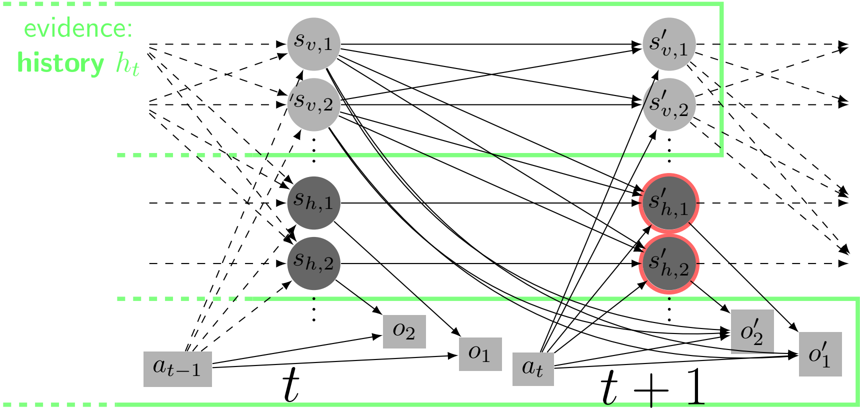
beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



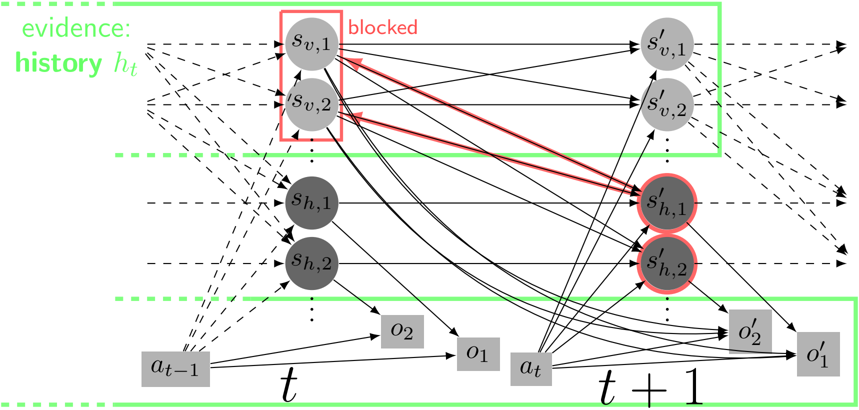
beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



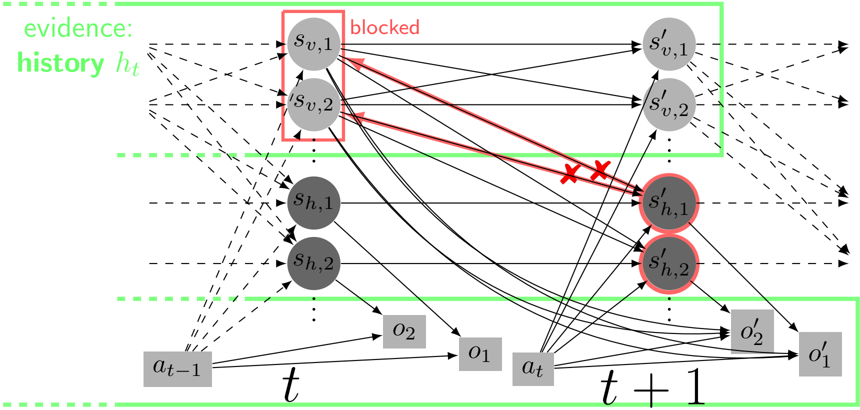
beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



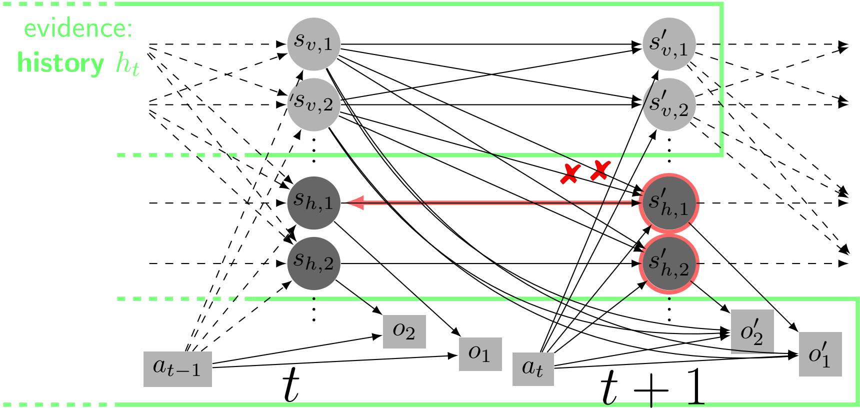
beginning of the proof :

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



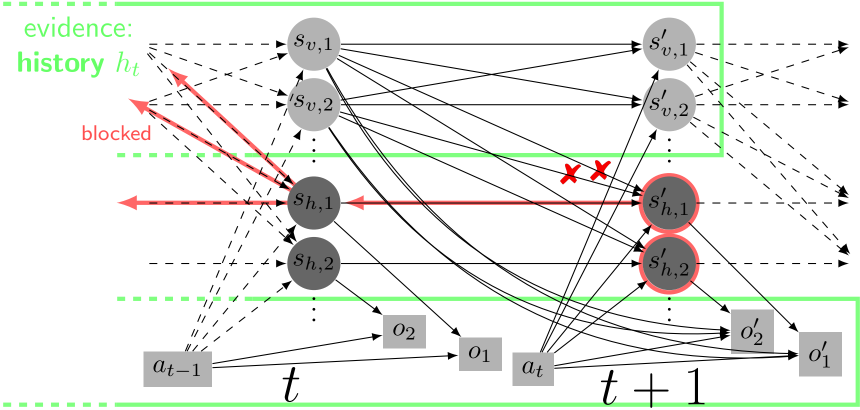
beginning of the proof :

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



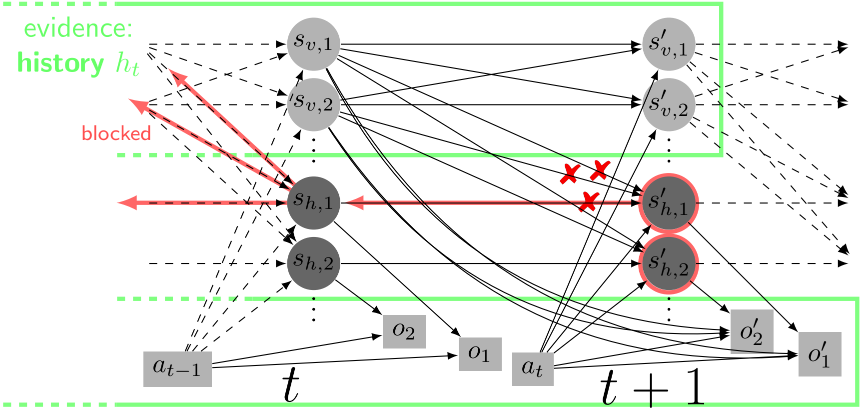
beginning of the proof :

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



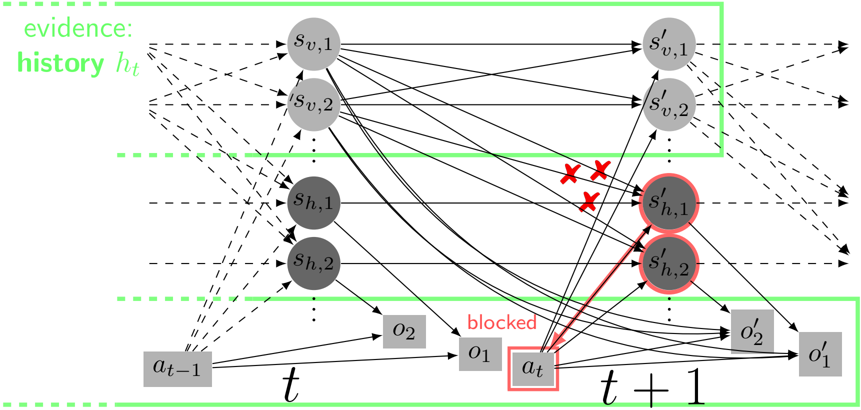
beginning of the proof :

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



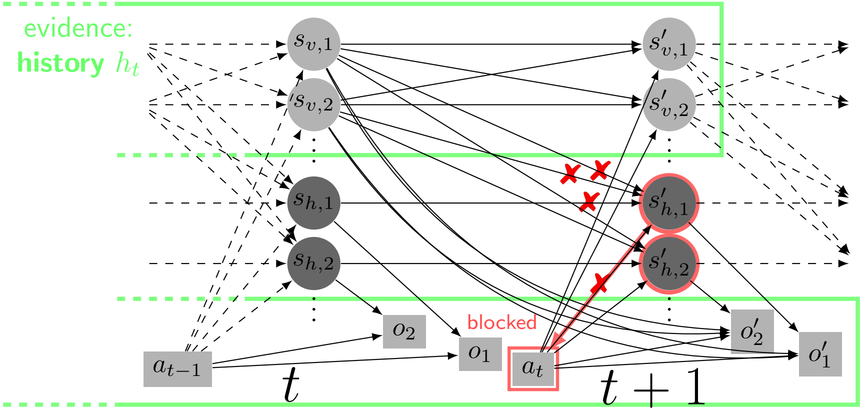
beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



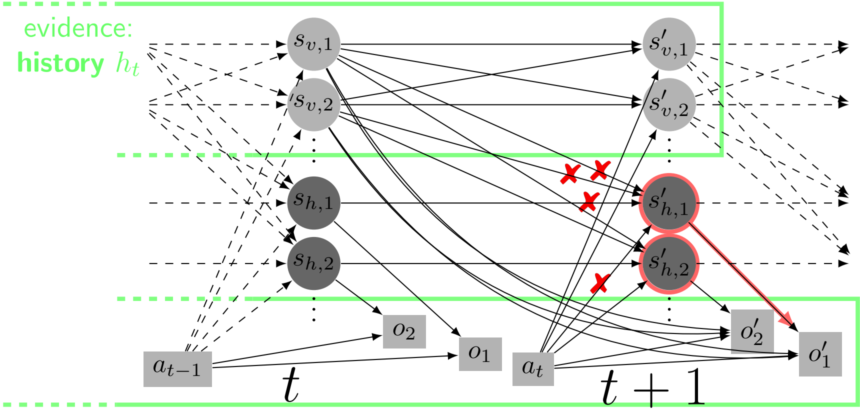
beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



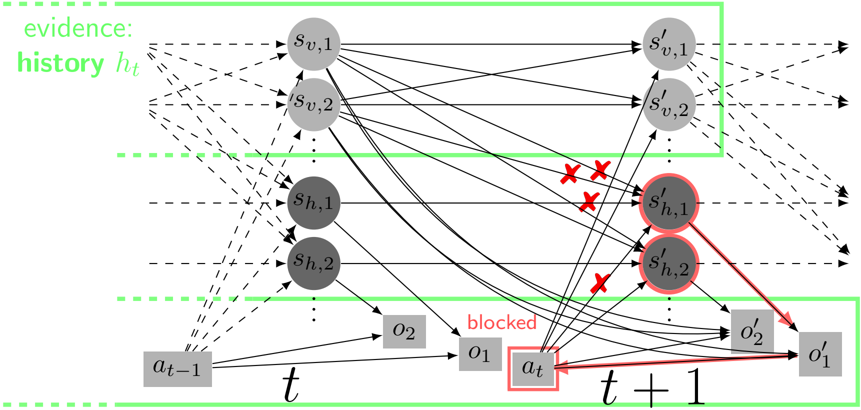
beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



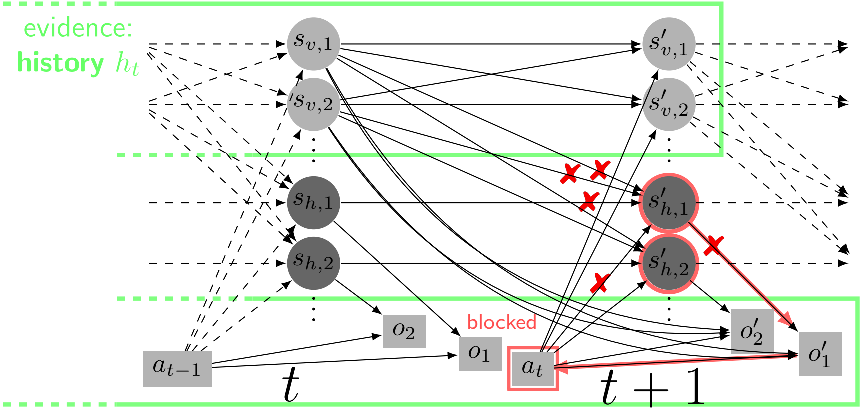
beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)



beginning of the proof:

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)

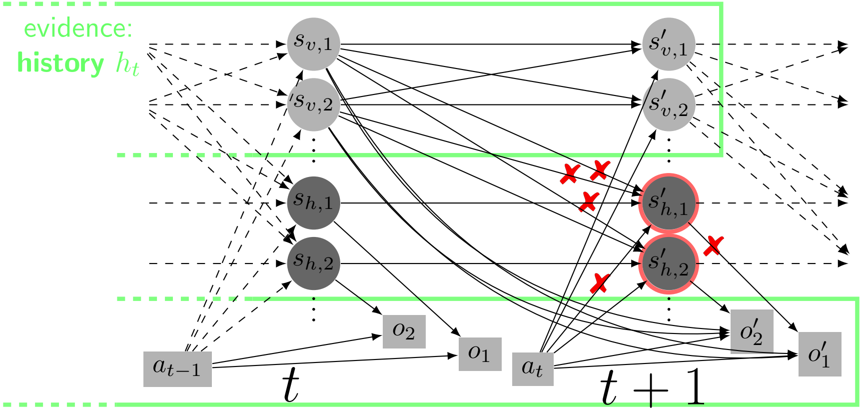


beginning of the proof :

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ and $s_{h,j}$ are d-**separated** by evidence h_t (history)

$\rightarrow$ for each time t, hidden state variables $s_{h,i}$ are independent given h_t

$$\text{i.e. } \beta_t(s_h) = \pi\left(s_h \mid h_t\right) = \min_i \pi\left(s_{h,i} \mid h_t\right) = \min_i \beta_{t,i}(s_{h,i})$$



Experimentations

PPUDD against SPUDD (*Hoey et al., UAI-99*)

Navigation benchmark: robot has to reach a goal;
may disappear in some places.

2 possibilistic translations: M1 (optimistic) and M2 (pessimistic).

Experimentations

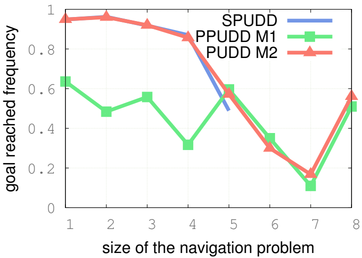
PPUDD against SPUDD (*Hoey et al., UAI-99*)

Navigation benchmark: robot has to reach a goal;
may disappear in some places.

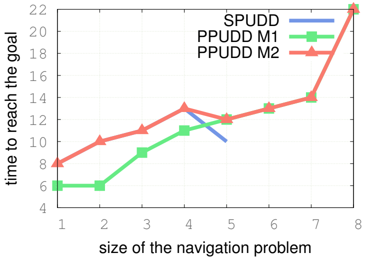
2 possibilistic translations: M1 (optimistic) and M2 (pessimistic).

Performances for increasing size of benchmark instances:

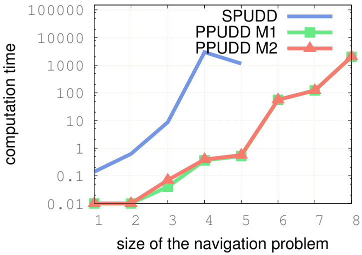
Goal reached frequency



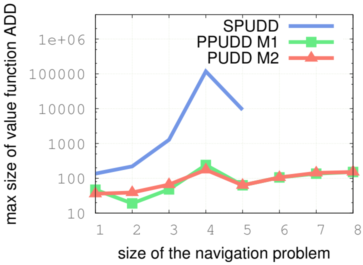
Average of times to reach the goal



Computation time



Max size of ADD "value function"



- PPUDD + M2 (pessimistic translation) **faster with same quality performances** as SPUDD;
- SPUDD only solve the 5th first instances;
- verified intuition: ADDs are smaller.

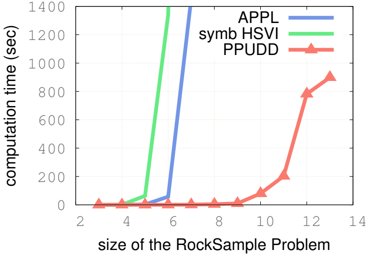
PPUDD against APPL (*Kurniawati et al., RSS-08*, MOMDP solver);
symbolic HSVI (*Sim et al., AAAI-09, POMDP solver*).

RockSample benchmark: a robot has to recognize interesting rocks;
sample only interesting ones.

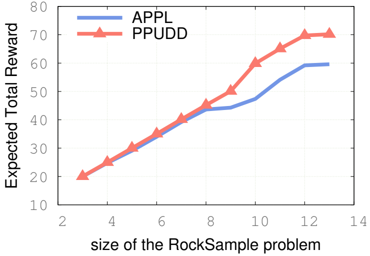
PPUDD against APPL (*Kurniawati et al., RSS-08*, MOMDP solver);
symbolic HSVI (*Sim et al., AAAI-09, POMDP solver*).

RockSample benchmark: a robot has to recognize interesting rocks;
sample only interesting ones.

Computation time:
probabilistic solvers, prec. 1;
PPUDD, exact resolution.



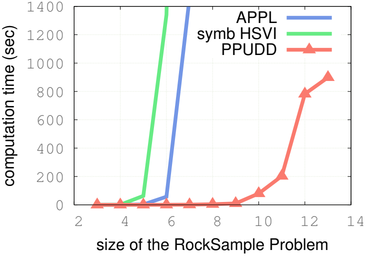
Expected total reward
APPL stopped at
PPUDD computation time.



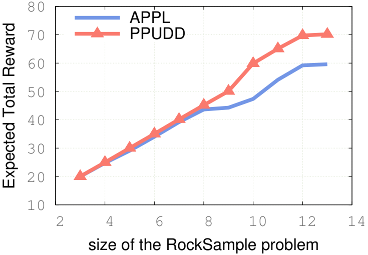
PPUDD against APPL (*Kurniawati et al., RSS-08*, MOMDP solver);
symbolic HSVI (*Sim et al., AAAI-09, POMDP solver*).

RockSample benchmark: a robot has to recognize interesting rocks;
sample only interesting ones.

Computation time:
probabilistic solvers, prec. 1;
PPUDD, exact resolution.



Expected total reward
APPL stopped at
PPUDD computation time.



- **approximate model + exact efficient algorithm**
  → beneficial in computation times and performances.

IPPC 2014 results – MDP track

PPUDD + BDD mask for computation over reachable states only

IPPC 2014 results – MDP track

PPUDD + BDD mask for computation over reachable states only

results:

PPUDD $2\times$ better than LRTDP (*Bonet et al.*, ICAPS-03) with ADDs;
$2\times$ worse than LRTDP with enumerated states.

IPPC 2014 results – MDP track

PPUDD + BDD mask for computation over reachable states only

results:

PPUDD $2\times$ better than LRTDP (*Bonet et al.*, ICAPS-03) with ADDs;
$2\times$ worse than LRTDP with enumerated states.

$\Rightarrow$ ADDs based methods: possibilistic approach (*without* heuristics) **better than** probabilistic approach (*with* heuristic).

IPPC 2014 results – MDP track

PPUDD + BDD mask for computation over reachable states only

results:

PPUDD $2\times$ better than LRTDP (*Bonet et al.*, ICAPS-03) with ADDs;
$2\times$ worse than LRTDP with enumerated states.

$\Rightarrow$ ADDs based methods: possibilistic approach (*without* heuristics)
better than probabilistic approach (*with* heuristic).

open problems:

- symbolic methods $\prec$ state space search
  $\longrightarrow$ possibilistic approach with enumerated states and heuristics like winners of IPPC 2014 (PROST and GOURMAND)?
- **naive** $\mathbb{P} \rightarrow \Pi$ **automated translation**
  automatic and efficient translation $\mathbb{P} \rightarrow \Pi$?

IPPC 2014 results – MDP track

PPUDD + BDD mask for computation over reachable states only

results:

PPUDD $2\times$ better than LRTDP (*Bonet et al.*, ICAPS-03) with ADDs;
$2\times$ worse than LRTDP with enumerated states.

$\Rightarrow$ ADDs based methods: possibilistic approach (*without* heuristics)
better than probabilistic approach (*with* heuristic).

open problems:

- symbolic methods $\prec$ state space search
  $\rightarrow$ possibilistic approach with enumerated states and heuristics like winners of IPPC 2014 (PROST and GOURMAND)?
- **naive $\mathbb{P} \rightarrow \Pi$ automated translation**
  automatic and efficient translation $\mathbb{P} \rightarrow \Pi$?
- time management: PPUDD's total computation time $= 15h \ll 24h$;
- ADD instanciation does not fit into memory for 15 in 80 benchmarks.

Conclusion and perspectives

Works called by this one:

- test POMDP track of IPPC 2014,
- bring belief factorization to probabilistic framework (see Boyen, Koller, McAllester, Poupart, Shani . . .).

Conclusion and perspectives

Works called by this one:

- test POMDP track of IPPC 2014,
- bring belief factorization to probabilistic framework (see Boyen, Koller, McAllester, Poupart, Shani . . .).

PPUDD: possibilistic algorithm taking advantage of π-MOMDP structure.

- computations stay in finite scale $\mathcal{L}$
  $\Rightarrow$ used **ADDs smaller** than for probabilistic framework;
- **natural factorization** for π-MOMDP modelling system with few independent sensors;
- approximate model and exact resolution can win against exact model and approximate resolution.

Thank you for your attention!

Any questions?

let $e_1 \neq e_2$ be two events $\subset \Omega$:

meaning or operation	THEORY Probability ($\mathbb{P}$)	Possibility (Π)
normalization	$\sum_\Omega \mathbf{p}(e) = 1$	$\max_\Omega \pi(e) = 1$
e_1 impossible	$\mathbf{p}(e_1) = 0$	$\pi(e_1) = 0$
e_1 certain	$\mathbf{p}(e_1) = 1$	$\max_{e \notin e_1} \pi(e) = 0$
e_1 or e_2 (union)	$\mathbf{p}(e_1) + \mathbf{p}(e_2)$	$\max\left\{\pi(e_1), \pi(e_2)\right\}$
e_1 and e_2 (intersect.)	$\mathbf{p}(e_1).\mathbf{p}\left(e_2 \mid e_1\right)$	$\min\left\{\pi(e_1), \pi(e_2 \mid e_1)\right\}$
Bayes' rule: $e_1 \| e_2$	$\frac{\mathbf{p}(e_2 \mid e_1).\mathbf{p}(e_1)}{\mathbf{p}(e_2)}$	1 if $e_1 \in \underset{e \subset \Omega}{\operatorname{argmax}} \pi(e_1, e)$ $\pi\left(e_1, e_2\right)$ otherwise.

d-separation

$\Rightarrow$ each belief state β_i updated using corresponding observation only;

$\Rightarrow s_{v,1}, \dots s_{v,m}, \beta_1, \dots, \beta_l$ are independent post-action variables.

$\Rightarrow m + l$ ADDs (for each subset) instead of a bigger one (over $\mathcal{X}$).

d-separation

$\Rightarrow$ each belief state β_i updated using corresponding observation only;

$\Rightarrow s_{v,1}, \dots s_{v,m}, \beta_1, \dots, \beta_l$ are independent post-action variables.

$\Rightarrow m + l$ ADDs (for each subset) instead of a bigger one (over $\mathcal{X}$).

$$\Rightarrow \mathcal{X} = \mathcal{S}_{v,1} \times \dots \times \mathcal{S}_{v,m} \times B^{\pi}_1 \times \dots \times B^{\pi}_l$$

Size of B^{π}_j: $\#B^{\pi}_j = \#\mathcal{L}^{\#\mathcal{S}_{h,j}} - (\#\mathcal{L} - 1)^{\#\mathcal{S}_{h,j}}$.

d-separation

$\Rightarrow$ each belief state β_i updated using corresponding observation only;

$\Rightarrow$ $s_{v,1}, \ldots s_{v,m}, \beta_1, \ldots, \beta_l$ are independent post-action variables.

$\Rightarrow m + l$ ADDs (for each subset) instead of a bigger one (over $\mathcal{X}$).

$$\Rightarrow \mathcal{X} = \mathcal{S}_{v,1} \times \ldots \times \mathcal{S}_{v,m} \times B_1^\pi \times \cdots \times B_l^\pi$$

Size of B_j^π: $\#B_j^\pi = \#\mathcal{L}^{\#\mathcal{S}_{h,j}} - (\#\mathcal{L} - 1)^{\#\mathcal{S}_{h,j}}$.

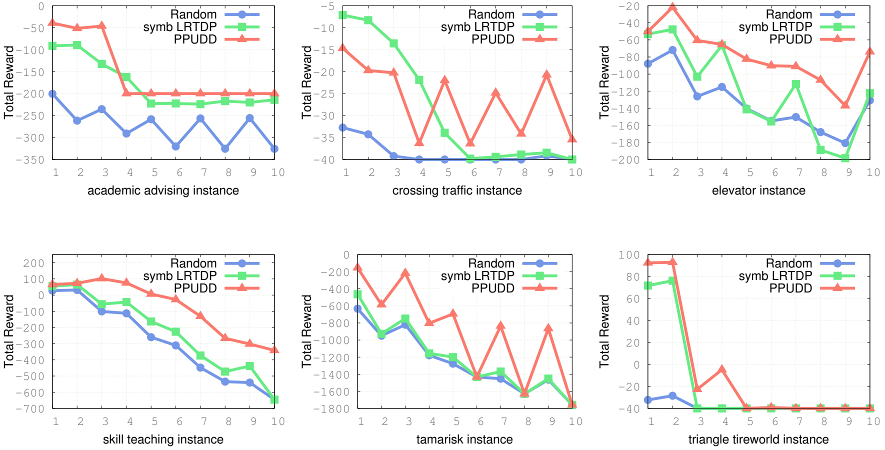
binary state variables $\Rightarrow \#B_j^\pi = 2\#\mathcal{L} - 1$

$$\Rightarrow \#\mathcal{X} = 2^m (2\#\mathcal{L} - 1)^l.$$

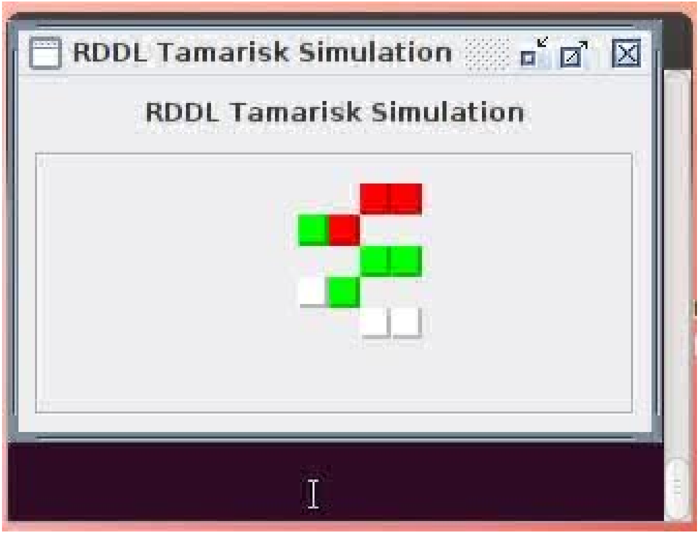
hidden and visible state variables $\rightarrow$ **same impact** on space size: singly-exponential in the number of state variables $\neq$ probabilistic settings.

IPPC 2014 results – PPUDD vs symbolic LRTDP

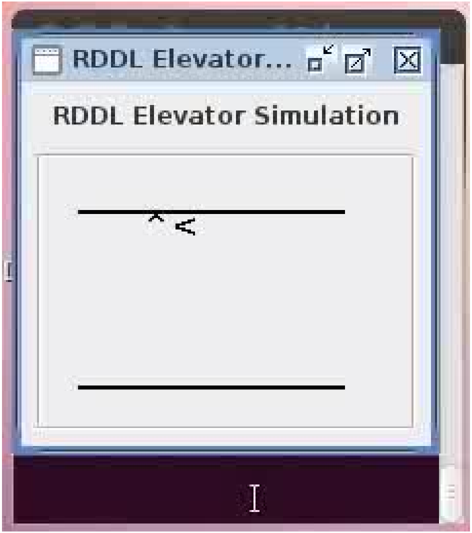
Average of Total Reward for simulated runs



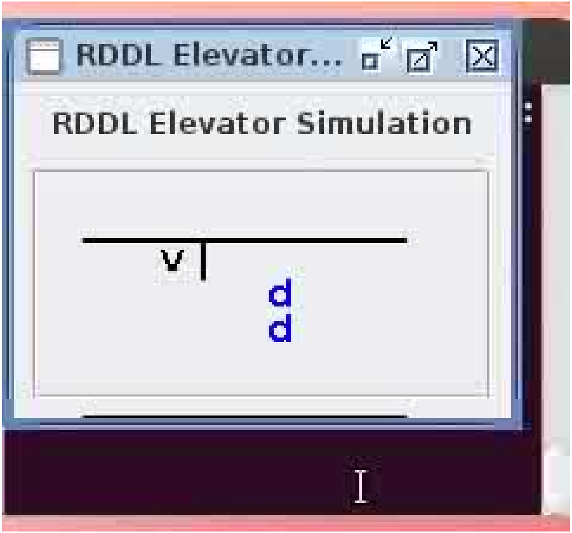
participation in IPPC 2014 – MDP track



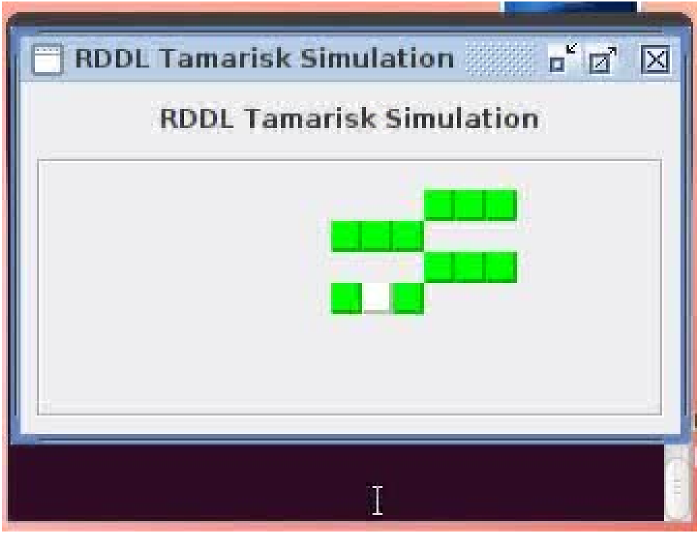
participation in IPPC 2014 – MDP track



participation in IPPC 2014 – MDP track

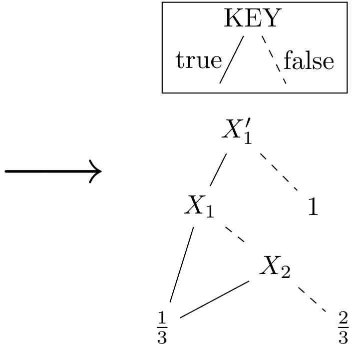


participation in IPPC 2014 – MDP track



$X_1' =$ true		
true	$^1/_3$	$^1/_3$
false	$^1/_3$	$^2/_3$
X_1/X_2	true	false

$X_1' =$ false		
true	1	1
false	1	1
X_1/X_2	true	false



intuition: remains small not creating new leaves/ADDs