Strategy Computation for Robotic Missions using Possibilistic Approaches

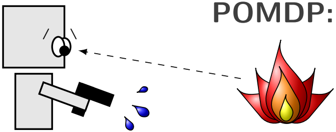
Nicolas Drougard
ONERA–The French Aerospace Lab, DCSD, Toulouse

ONERA
THE FRENCH AEROSPACE LAB

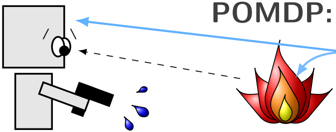
retour sur innovation

Outline

Outline

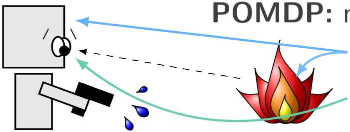


POMDP: model for autonomous robotics



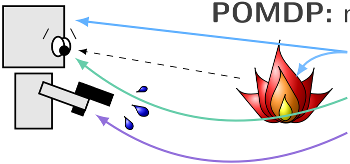
POMDP: model for autonomous robotics

$s \in \mathcal{S}$: **system state**;



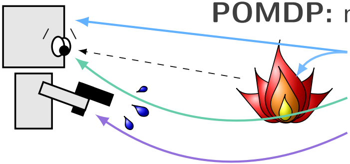
POMDP: model for autonomous robotics

- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;



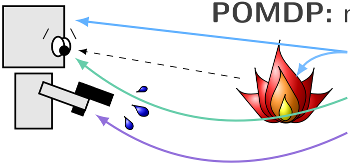
POMDP: model for autonomous robotics

- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;



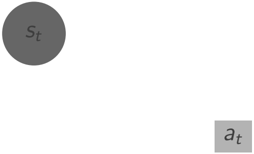
POMDP: model for autonomous robotics

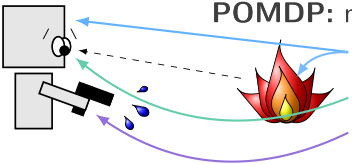
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;



POMDP: model for autonomous robotics

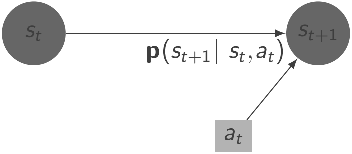
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

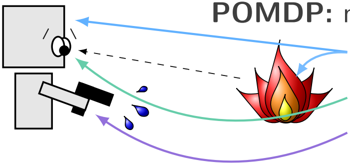




POMDP: model for autonomous robotics

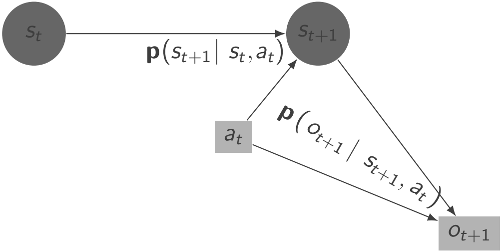
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

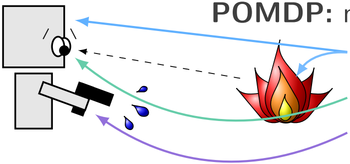




POMDP: model for autonomous robotics

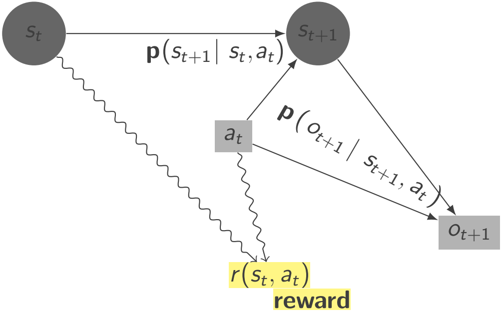
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

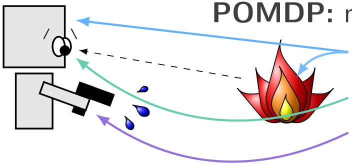




POMDP: model for autonomous robotics

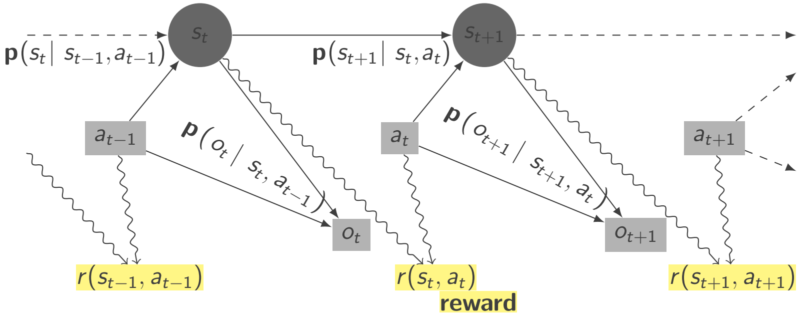
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

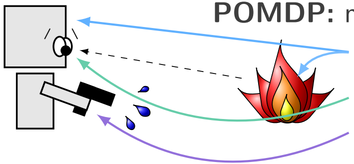




POMDP: model for autonomous robotics

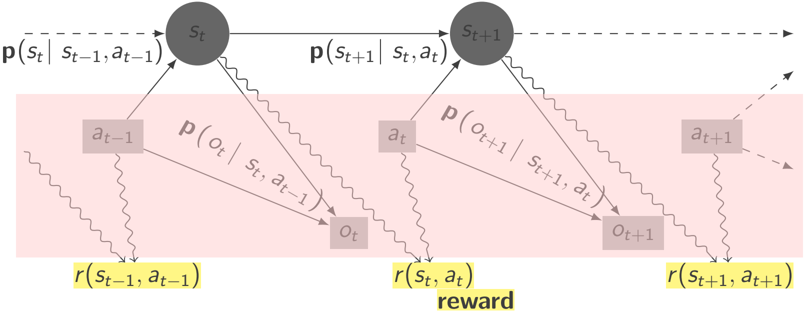
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

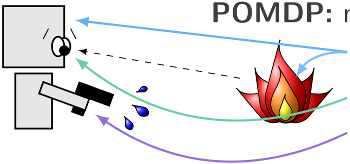




POMDP: model for autonomous robotics

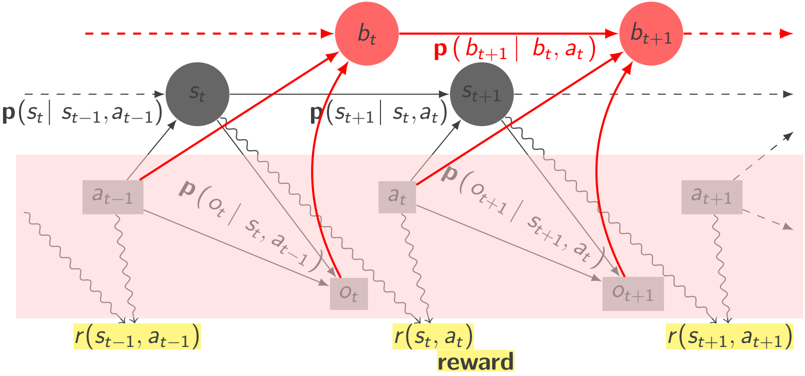
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;





POMDP: model for autonomous robotics

- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;
- b **belief.**

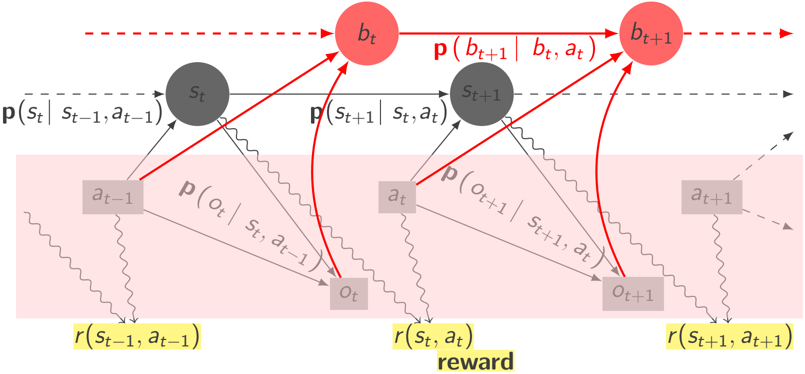


ONERA
THE FRENCH AEROSPACE LAB

Context and Background

Bayes rule, Strategy, Criterion.

$$b_{t+1}(s') = nextBelief\left(b_t, a, \tilde{o}\right) = \frac{\mathbf{p}(\tilde{o}|s',a).\sum_s \mathbf{p}(s'|s,a)b_t(s)}{\sum_{\underline{s},\bar{s}} \mathbf{p}(\tilde{o}|\bar{s},a).\mathbf{p}(\bar{s}|\underline{s},a)b_t(\underline{s})}$$



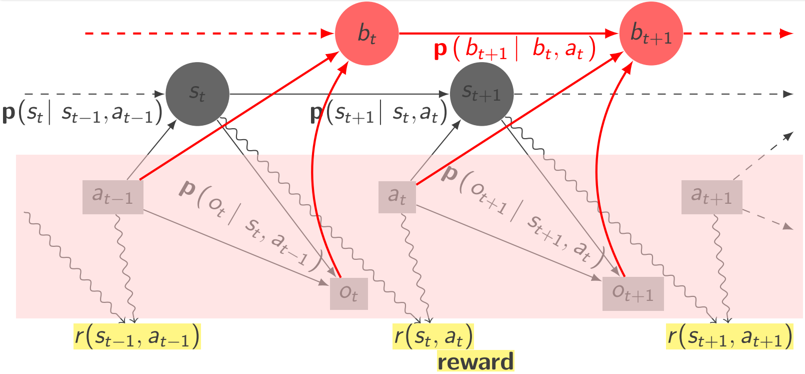
Context and Background

Bayes rule, Strategy, Criterion.

$$b_{t+1}(s') = nextBelief\left(b_t, a, \tilde{o}\right) = \frac{\mathbf{p}(\tilde{o}|s',a).\sum_s \mathbf{p}(s'|s,a)b_t(s)}{\sum_{\underline{s},\bar{s}} \mathbf{p}(\tilde{o}|\bar{s},a).\mathbf{p}(\bar{s}|\underline{s},a)b_t(\underline{s})}$$

Actions choice: strategy $\delta(b_t) = a_t \in \mathcal{A}$

maximizing $\mathbb{E}[\sum_{t=0}^{+\infty} \gamma^t r(s_t, \delta(b_t))|b_0]$, $0 < \gamma < 1$.



- **strategy computation $\geqslant$ PSPACE-complete:**

- **strategy computation $\geqslant$ PSPACE-complete:**

$\rightarrow$ optimality for "small" ou "structured" POMDPs;

$\rightarrow$ approximate computations (no optimality garantee).

- **strategy computation $\geqslant$ PSPACE-complete:**

$\rightarrow$ optimality for “small” ou “structured” POMDPs;

$\rightarrow$ approximate computations (no optimality garantee).

- **Computer Vision**, statistical learning: $\mathbf{p}\left(o' \mid s', a\right)$

- **strategy computation $\geqslant$ PSPACE-complete:**

$\rightarrow$ optimality for "small" ou "structured" POMDPs;

$\rightarrow$ approximate computations (no optimality garantee).

- **Computer Vision**, statistical learning: $\mathbf{p}\left(o' \mid s', a\right)$

$\rightarrow$ wide picture variability? $\Rightarrow$ hard extraction of a classifier;

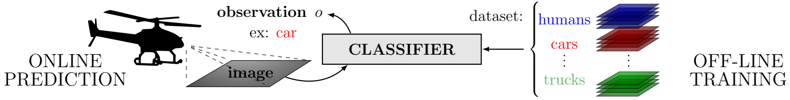
$\rightarrow$ more observations? $\Rightarrow$ more complex POMDP.

- **strategy computation $\geqslant$ PSPACE-complete:**

$\rightarrow$ optimality for "small" ou "structured" POMDPs;
$\rightarrow$ approximate computations (no optimality garantee).

- **Computer Vision**, statistical learning: $\mathbf{p}\left(o' \mid s', a \right)$

$\rightarrow$ wide picture variability? $\Rightarrow$ hard extraction of a classifier;
$\rightarrow$ more observations? $\Rightarrow$ more complex POMDP.



- **Initial belief** b_0 (*prior* information on the system state).

- **strategy computation $\geqslant$ PSPACE-complete:**

$\rightarrow$ optimality for "small" ou "structured" POMDPs;

$\rightarrow$ approximate computations (no optimality garantee).

- **Computer Vision**, statistical learning: $\mathbf{p}\left(o' \mid s', a\right)$

$\rightarrow$ wide picture variability? $\Rightarrow$ hard extraction of a classifier;

$\rightarrow$ more observations? $\Rightarrow$ more complex POMDP.

- **Initial belief** b_0 (*prior* information on the system state).

uniform = subjectif, mix up with frequencies!

- **strategy computation $\geqslant$ PSPACE-complete:**

$\rightarrow$ optimality for "small" ou "structured" POMDPs;
$\rightarrow$ approximate computations (no optimality garantee).

- **Computer Vision**, statistical learning: $\mathbf{p}\left(o' \mid s', a\right)$

$\rightarrow$ wide picture variability? $\Rightarrow$ hard extraction of a classifier;
$\rightarrow$ more observations? $\Rightarrow$ more complex POMDP.

- **Initial belief** b_0 (*prior* information on the system state).

uniform = subjectif, mix up with frequencies!

Qualitative Possibility Theory:
$\rightarrow$ simplification, ignorance and imprecision modeling.

Qualitative Possibility Theory

$\mathcal{L}$ finite scale, ex: $\left\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\right\}$.

events $e \subset \Omega$ (sample space)

sorted with possibility **degrees** $\pi(e) \in \mathcal{L}$,

$\neq$

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

Qualitative Possibility Theory

$\mathcal{L}$ finite scale, ex: $\left\{ 0, \frac{1}{n}, \frac{2}{n}, \dots, 1 \right\}$.

events $e \subset \Omega$ (sample space)

sorted with possibility **degrees** $\pi(e) \in \mathcal{L}$,

$\neq$

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

$e_1 \neq e_2$, two events $\subset \Omega$

- $\pi(e_1) < \pi(e_2) \Leftrightarrow$ "e_1 is less plausible than e_2";

Qualitative Possibility Theory

$\mathcal{L}$ finite scale, ex: $\left\{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\right\}$.

events $e \subset \Omega$ (sample space)

sorted with possibility **degrees** $\pi(e) \in \mathcal{L}$,

$\neq$

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

$e_1 \neq e_2$, two events $\subset \Omega$

- $\pi(e_1) < \pi(e_2) \Leftrightarrow$ "e_1 is less plausible than e_2";

Probability ($\mathbb{P}$) / Possibility (Π):

e_1 or e_2	$\mathbf{p}(e_1) + \mathbf{p}(e_2 \cap \overline{e_1})$	$\max\left\{\pi(e_1), \pi(e_2)\right\}$
e_1 and e_2	$\mathbf{p}(e_1).\mathbf{p}\left(e_2 \mid e_1\right)$	$\min\left\{\pi(e_1), \pi(e_2 \mid e_1)\right\}$

Outline

Possibilistic models: π-MOMDPs

possibilistic POMDPs (π-POMDPs): *Sabbadin UAI-98*.

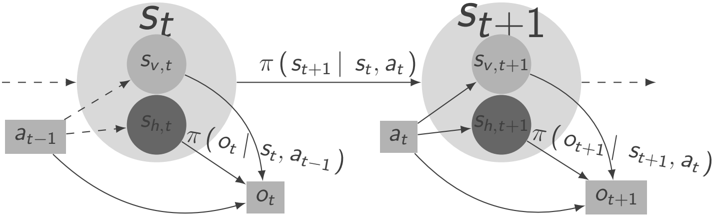
- finite belief space $\#\mathcal{B} = \#\mathcal{L}^{\#\mathcal{S}} - (\#\mathcal{L} - 1)^{\#\mathcal{S}}$

Possibilistic models: π-MOMDPs

possibilistic POMDPs (π-POMDPs): *Sabbadin UAI-98*.

- finite belief space $\#\mathcal{B} = \#\mathcal{L}^{\#\mathcal{S}} - (\#\mathcal{L} - 1)^{\#\mathcal{S}}$

contribution (UAI13):



Mixed-Observability: system state $s \in \mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$

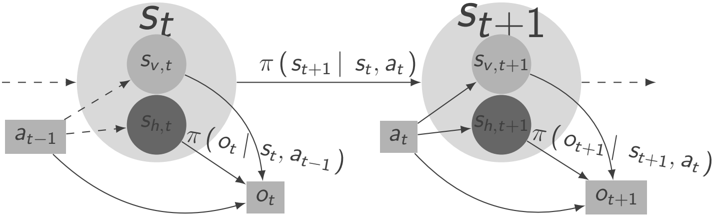
i.e. state s = visible component s_v & hidden component s_h.

Possibilistic models:
π-MOMDPs

possibilistic POMDPs (π-POMDPs): *Sabbadin UAI-98*.

- finite belief space $\#\mathcal{B} = \#\mathcal{L}^{\#\mathcal{S}} - (\#\mathcal{L}-1)^{\#\mathcal{S}}$

contribution (UAI13):



Mixed-Observability: system state $s \in \mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$

i.e. state s = visible component s_v & hidden component s_h.

- beliefs are only over $\mathcal{S}_h$ (component s_v observed),
- computations on $\mathcal{X} = \mathcal{S}_v \times \mathcal{B}_h$ whose size is

$$\#\mathcal{X} = \#\mathcal{S}_v \cdot \left(\#\mathcal{L}^{\#\mathcal{S}_h} - (\#\mathcal{L}-1)^{\#\mathcal{S}_h}\right) \ll \#\mathcal{B}.$$

π-MOMDPs in practice

contribution (UAI13): Infinite Horizon

Dynamic Programming scheme: $\#$ iterations $< \#\mathcal{X}$.

$\forall x \in \mathcal{X}, \; V_0(x) = \mu(x)$ **preference**,

π-MOMDPs in practice

contribution (UAI13): Infinite Horizon

Dynamic Programming scheme: $\#$ iterations $< \#\mathcal{X}$.

$\forall x \in \mathcal{X}$, $V_0(x) = \mu(x)$ **preference**, and, until convergence,

• $V_{i+1}(x) = \max_{a \in \mathcal{A}} \max_{x' \in \mathcal{X}} \min \left\{ \pi\left(x' \mid x, a\right), V_i(x') \right\}$,

π-MOMDPs in practice

contribution (UAI13): Infinite Horizon

Dynamic Programming scheme: $\#$ iterations $< \#\mathcal{X}$.

$\forall x \in \mathcal{X},\ V_0(x) = \mu(x)$ **preference**, and, until convergence,

$\bullet V_{i+1}(x) = \max\limits_{a \in \mathcal{A}} \max\limits_{x' \in \mathcal{X}} \min \left\{ \pi\left(x' \mid x, a \right), V_i(x') \right\}$, and

if $V_{i+1}(x) > V_i(x), \delta(x) = \operatorname{argmax}\limits_{a \in \mathcal{A}} \max\limits_{x' \in \mathcal{X}} \min\{\pi\left(x' \mid x, a \right), V_i(x')\}$.

π-MOMDPs in practice

contribution (UAI13): Infinite Horizon

Dynamic Programming scheme: $\#$ iterations $< \#\mathcal{X}$.

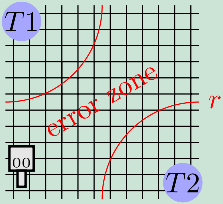
$\forall x \in \mathcal{X}$, $V_0(x) = \mu(x)$ **preference**, and, until convergence,

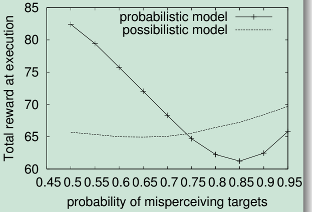
- $V_{i+1}(x) = \max\limits_{a \in \mathcal{A}} \max\limits_{x' \in \mathcal{X}} \min \left\{ \pi\left(x' \mid x, a\right), V_i(x') \right\}$, and

if $V_{i+1}(x) > V_i(x)$, $\delta(x) = \underset{a \in \mathcal{A}}{\operatorname{argmax}} \max\limits_{x' \in \mathcal{X}} \min\{\pi\left(x' \mid x, a\right), V_i(x')\}$.

Recognition mission: robot on a grid $g \times g$, 2 targets $T1$, $T2$.

- **goal:** reach the object $A = T1$ or $T2$;
- noisy observations of the targets natures: $\mathbf{p}\left(o' \mid s', a\right)$.



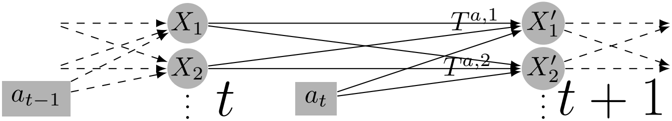


Actually, misperception in the error zone is: $P_{bad} > \frac{1}{2}$.

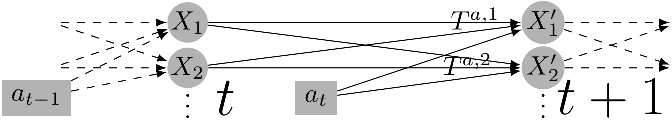
Outline

Factorization and symbolic solver

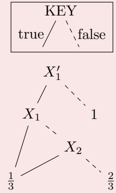
contribution (AAAI14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \mathcal{B}_h =$ Boolean variables $(X_1, \ldots, X_n)$
$+$ independence assumptions $\Leftarrow$ graphical model.



contribution (AAAI14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \mathcal{B}_h =$ Boolean variables $(X_1, \ldots, X_n)$
$+$ independence assumptions $\Leftarrow$ graphical model.



- transition functions
  $T_i^a = \pi\left(X_i' \mid \textit{parents}(X_i'), a \right)$
  represented by Algebraic Decision Diagrams (ADD).
  (SPUDD – *Hoey et al., UAI-99*).



example of ADD

Solver π-MOMDP résultant: PPUDD

- probabilistic model: $+$ and $\times$ $\Rightarrow$ new values created, number of ADDs leaves **potentially huge**.
- possibilistic model: min and max $\Rightarrow$ values $\in \mathcal{L}$ finite, number of leaves bounded, **ADDs smaller**.

Solver π-MOMDP résultant: PPUDD

- probabilistic model: $+$ and $\times$ $\Rightarrow$ new values created, number of ADDs leaves **potentially huge**.
- possibilistic model: min and max $\Rightarrow$ values $\in \mathcal{L}$ finite, number of leaves bounded, **ADDs smaller**.

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) ;

computations on trees: *CU Decision Diagram Package*.

Solver π-MOMDP résultant: PPUDD

- probabilistic model: $+$ and $\times$ $\Rightarrow$ new values created, number of ADDs leaves **potentially huge**.
- possibilistic model: min and max $\Rightarrow$ values $\in \mathcal{L}$ finite, number of leaves bounded, **ADDs smaller**.

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
$\Rightarrow$ dynamic programming

computations on trees: *CU Decision Diagram Package*.

Solver π-MOMDP résultant: PPUDD

- probabilistic model: $+$ and $\times$ $\Rightarrow$ new values created, number of ADDs leaves **potentially huge**.
- possibilistic model: min and max $\Rightarrow$ values $\in \mathcal{L}$ finite, number of leaves bounded, **ADDs smaller**.

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

$\Rightarrow$ dynamic programming

divided into *n* stages

used ADDs smaller

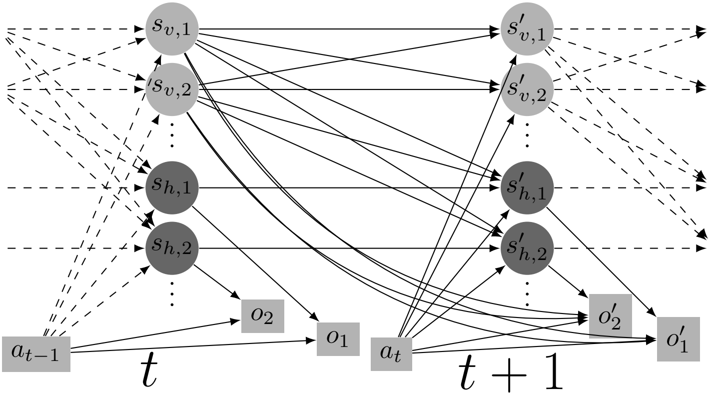
$\rightarrow$ **faster computations.**

computations on trees: *CU Decision Diagram Package*.

Natural factorisation: belief independence.

contribution (AAAI14): π-MOMDP following independence assumptions of the graphical model:

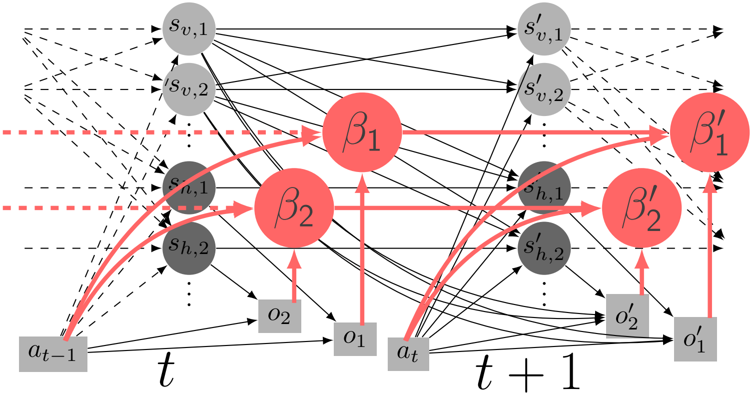
$\Rightarrow$ $(s_v, \beta) = (s_{v,1}, \ldots, s_{v,m}, \beta_1, \ldots, \beta_l)$, β_i belief over $s_{h,i}$.



Natural factorisation: belief independence.

contribution (AAAI14): π-MOMDP following independence assumptions of the graphical model:

$\Rightarrow$ $(s_v, \beta) = (s_{v,1}, \ldots, s_{v,m}, \beta_1, \ldots, \beta_l)$, β_i belief over $s_{h,i}$.

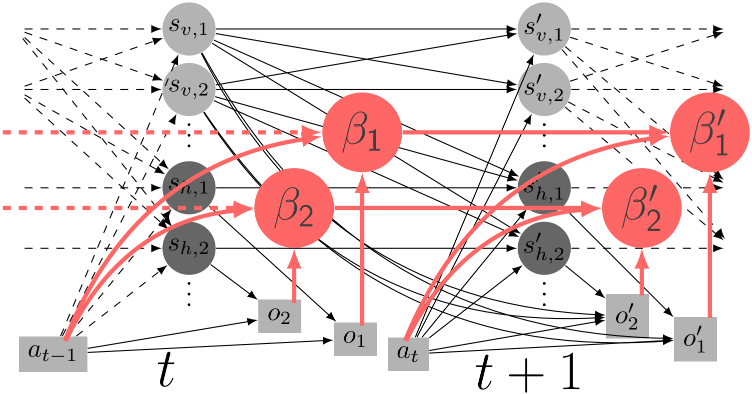


Natural factorisation: belief independence.

contribution (AAAI14): π-MOMDP following independence assumptions of the graphical model:

$\Rightarrow$ $(s_v, \beta) = (s_{v,1}, \ldots, s_{v,m}, \beta_1, \ldots, \beta_l)$, β_i belief over $s_{h,i}$.

assumptions: independent captors, hidden states...



Experiments: Navigation problem – agent = robot.

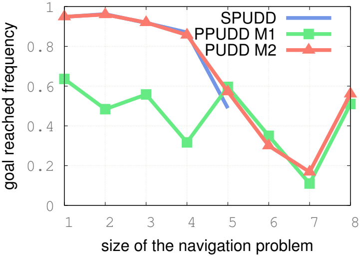
PPUDD vs SPUDD (*Hoey et al.*)
Navigation benchmark: reach a goal; spots with accident risk.
2 possibilistic translations: M1 (optimistic) et M2 (cautious).

Experiments: Navigation problem – agent = robot.

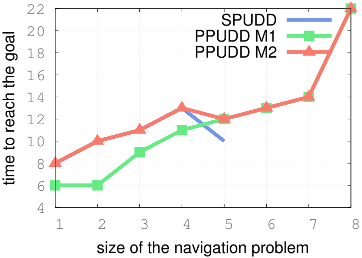
PPUDD vs SPUDD (*Hoey et al.*)
Navigation benchmark: reach a goal; spots with accident risk.
2 possibilistic translations: M1 (optimistic) et M2 (cautious).

Performances, function of the instance size

reached goal frequency

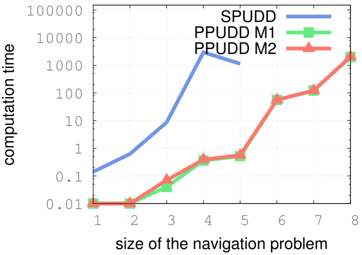


time to reach the goal

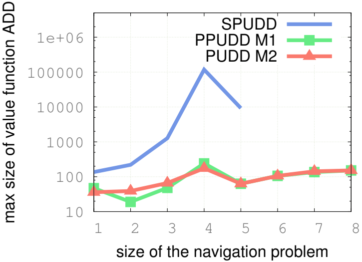


Experiments: Navigation problem – agent = robot.

computation time



max size of ADDs



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

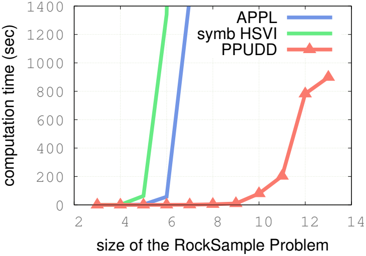
Experiments: RockSample problem – agent = robot.

PPUDD vs APPL (*Kurniawati et al.*, solver MOMDP);
symbolic HSVI (*Sim et al., solver POMDP*).
RockSample benchmark: recognize and sample "good" rocks;

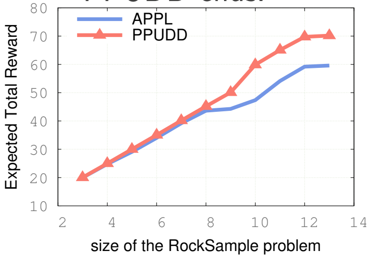
Experiments: RockSample problem – agent = robot.

PPUDD vs APPL (*Kurniawati et al.*, solver MOMDP);
symbolic HSVI (*Sim et al., solver POMDP*).
RockSample benchmark: recognize and sample "good" rocks;

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



average of rewards
APPL stopped when
PPUDD ends.



- **approximate model + exact resolution solver**
  → can improve of computation time and performances.

IPPC 2014 – MDP track. ADDs-based approaches: PPUDD vs symbolic LRTDP (*Bonet et al.*)

PPUDD + BDD mask over reachable states.

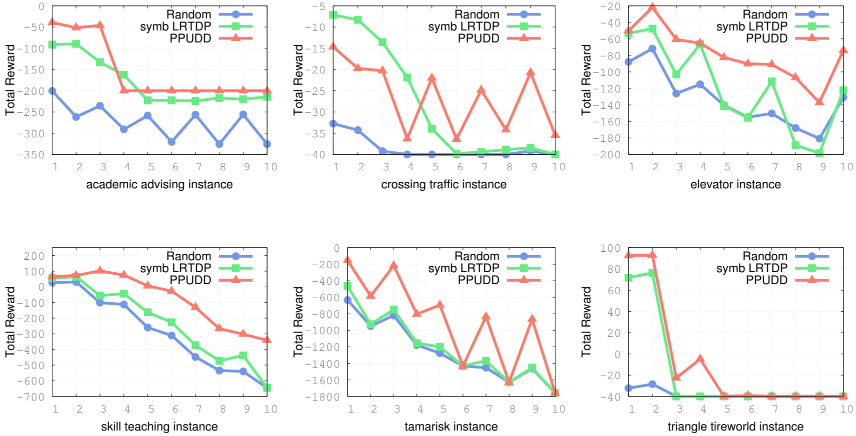


Figure : mean of rewards over simulations.

Outline

Possibility Theory:

- **granulated** belief space representation (discretization),
- efficient problem **simplification** (PPUDD $2\times$ better than LRTDP with ADDs);
- **ignorance and imprecision** modeling.

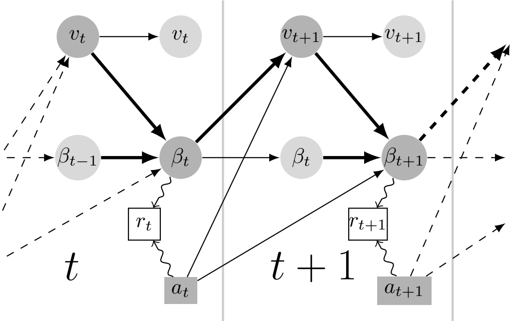
Conclusions/Perspectives

Possibility Theory:

- **granulated** belief space representation (discretization),
- efficient problem **simplification** (PPUDD $2\times$ better than LRTDP with ADDs);
- **ignorance and imprecision** modeling.

- ADD methods $\prec$ state space search methods: winners of IPPC 2014, (PROST & GOURMAND, $2\times$ better than PPUDD).
- choice of the qualitative criterion (optimistic/pessimistic);
- non additive utility degrees, from the same scale as possibility degrees.

Work in progress...
towards a hybrid POMDP



POMDP $\rightarrow$ POMDP with possibilistic beliefs

- transition probability distributions over possibilistic beliefs;
- reward aggregation: Choquet integral;
- factored POMDPs leads to factorized MPDs;
- resolution with any MDP solver.

Thank you.