Exploiting Imprecise Information Sources in Sequential Decision Making Problems under Uncertainty.

Ph.D defense **N.Drougard**

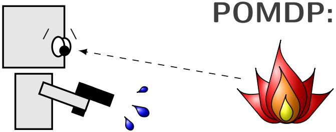
doctoral school: EDSYS,
institution: ISAE–SUPAERO,
laboratory: ONERA–The French Aerospace Lab.

ONERA
THE FRENCH AEROSPACE LAB

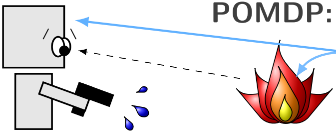
retour sur innovation

Plan

Plan

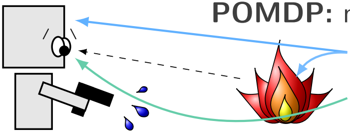


POMDP: model for autonomous robotics



POMDP: model for autonomous robotics

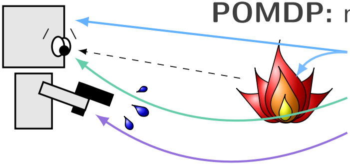
$s \in \mathcal{S}$: **system state**;



POMDP: model for autonomous robotics

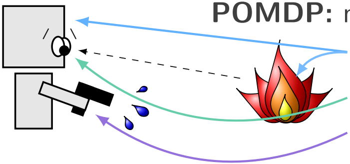
$s \in \mathcal{S}$: **system state**;

$o \in \mathcal{O}$: **system observations**;



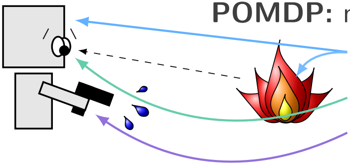
POMDP: model for autonomous robotics

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



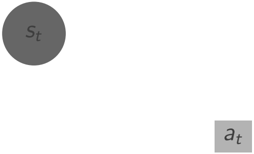
POMDP: model for autonomous robotics

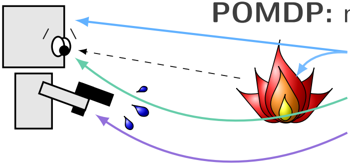
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;



POMDP: model for autonomous robotics

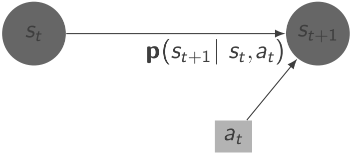
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

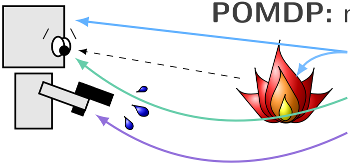




POMDP: model for autonomous robotics

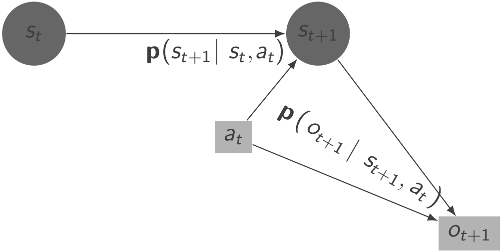
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;





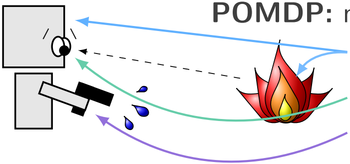
POMDP: model for autonomous robotics

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



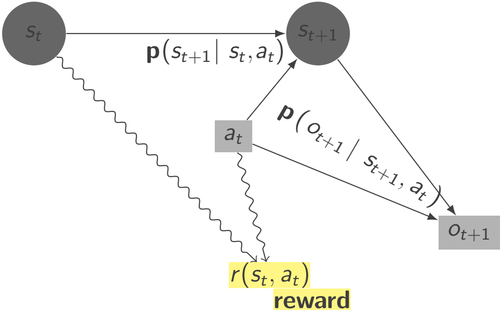
Context

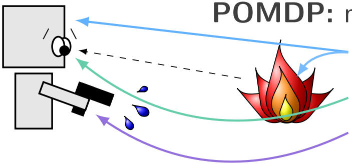
Partially Observable Markov Decision Processes (POMDPs)



POMDP: model for autonomous robotics

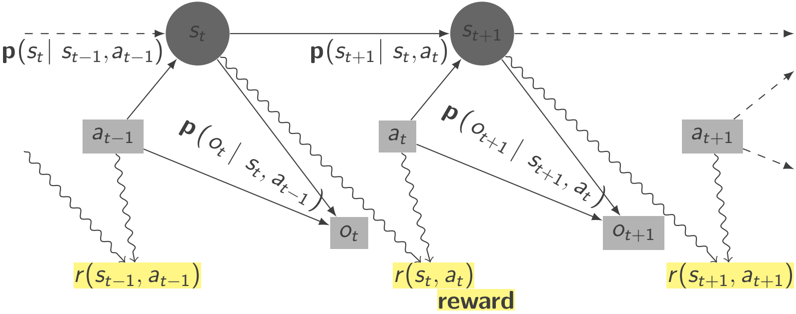
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

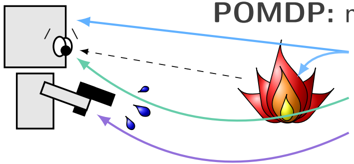




POMDP: model for autonomous robotics

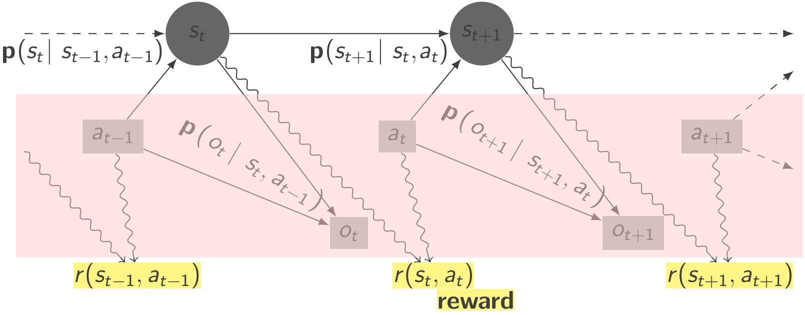
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;

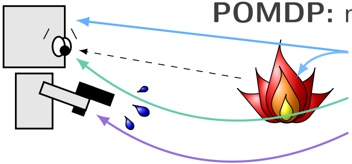




POMDP: model for autonomous robotics

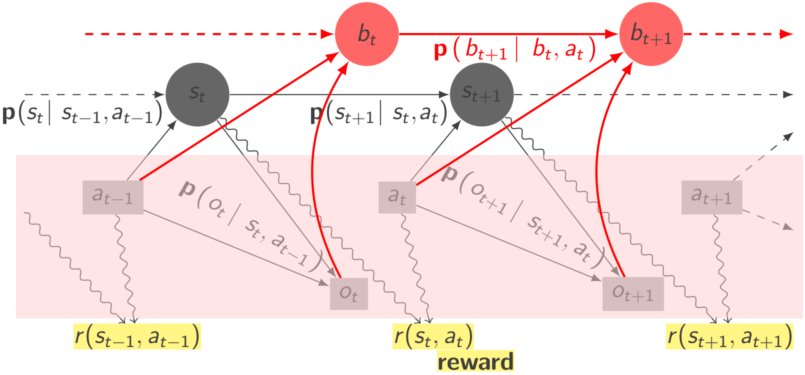
- $s \in \mathcal{S}$: **system state**;
- $o \in \mathcal{O}$: **system observations**;
- $a \in \mathcal{A}$: **agent's actions**;





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.**



POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$,

- **transition** function $T(s,a,s') = \mathbf{p}\left(s' \mid s, a \right)$;
- **observation** function $O(s',a,o') = \mathbf{p}\left(o' \mid s', a \right)$.

POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$,

- **transition** function $T(s, a, s') = \mathbf{p}\left(s' \mid s, a \right)$;
- **observation** function $O(s', a, o') = \mathbf{p}\left(o' \mid s', a \right)$.

belief state: $b_t(s) = \mathbb{P}\left(s_t = s \mid a_0, o_1, \ldots, a_{t-1}, o_t \right)$

POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$,

- **transition** function $T(s,a,s') = \mathbf{p}\left(s' \mid s,a\right)$;
- **observation** function $O(s',a,o') = \mathbf{p}\left(o' \mid s',a\right)$.

belief state: $b_t(s) = \mathbb{P}\left(s_t = s \mid a_0, o_1, \ldots, a_{t-1}, o_t\right)$

probabilistic belief update

$$b_{t+1}(s') \propto \mathbf{p}\left(o' \mid s',a\right) \cdot \sum_{s \in \mathcal{S}} \mathbf{p}\left(s' \mid s,a\right) \cdot b_t(s)$$

POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$,

- **transition** function $T(s,a,s') = \mathbf{p}\left(s' \mid s, a \right)$;
- **observation** function $O(s',a,o') = \mathbf{p}\left(o' \mid s', a \right)$.

belief state: $b_t(s) = \mathbb{P}\left(s_t = s \mid a_0, o_1, \ldots, a_{t-1}, o_t \right)$

probabilistic belief update

$$b_{t+1}(s') \propto \mathbf{p}\left(o' \mid s', a \right) \cdot \sum_{s \in \mathcal{S}} \mathbf{p}\left(s' \mid s, a \right) \cdot b_t(s)$$

action choices: strategy $\delta(b_t) = a_t \in \mathcal{A}$

$$\text{maximizing } \mathbb{E}_{s_0 \sim b_0} \left[\sum_{t=0}^{+\infty} \gamma^t \cdot r\Big(s_t, \delta(b_t) \Big) \right], \ 0 < \gamma < 1.$$

- optimal strategy computation $\geqslant$ **PSPACE**;
- probabilities are **imprecisely known** in practice;
- agent's **ignorance** not taken into account.

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

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

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

Context

practical issues: Complexity, Vision and Initial Belief.

- **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.

- initial belief deterministic $s_0 = s_A$.

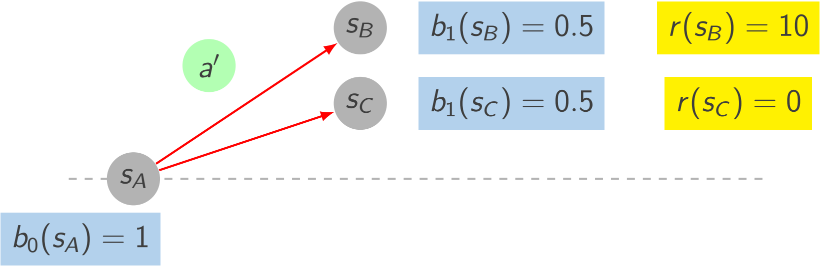
s_A

$b_0(s_A) = 1$

Why model ignorance?

knowledge is not always encouraged with POMDPs

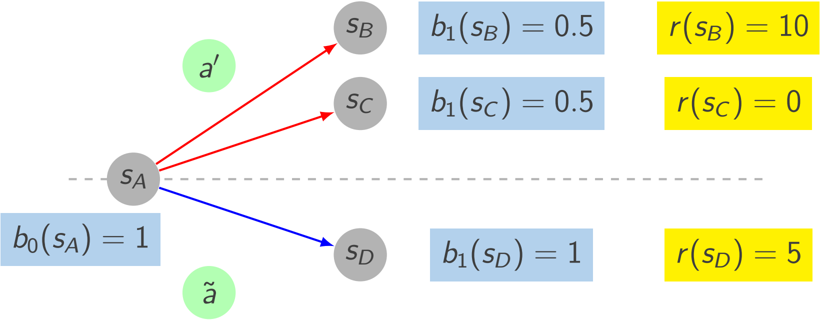
- initial belief deterministic $s_0 = s_A$.



Why model ignorance?

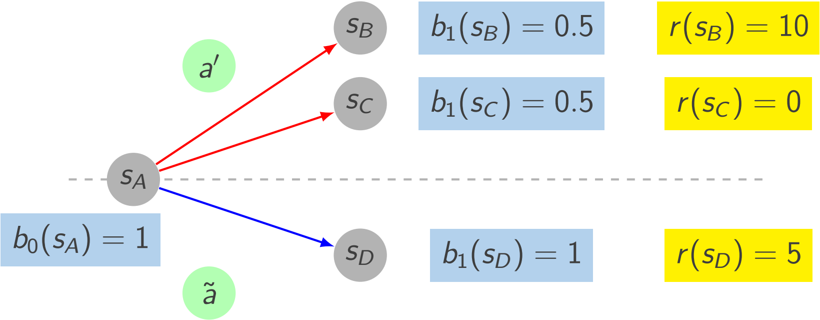
knowledge is not always encouraged with POMDPs

- initial belief deterministic $s_0 = s_A$.



Why model ignorance?
knowledge is not always encouraged with POMDPs

- initial belief deterministic $s_0 = s_A$.

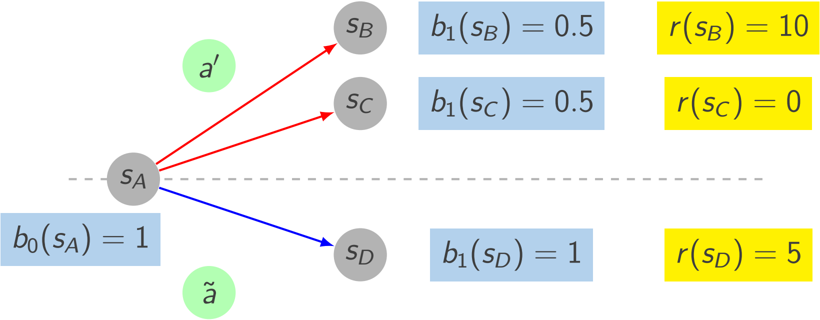


- $\{s_B, s_C, s_D\} \xrightarrow{\text{deterministic}} s_E,\ r(s_E) = 0.$

Why model ignorance?

knowledge is not always encouraged with POMDPs

- initial belief deterministic $s_0 = s_A$.



- $\{ s_B, s_C, s_D \} \xrightarrow{\text{deterministic}} s_E,\ r(s_E) = 0.$

$$\mathbb{E}_{s_0 \sim b_0} \left[\sum_{t=0}^{+\infty} \gamma^t \cdot r(s_t) \,\middle|\, a_0 = \tilde{a} \text{ or } a' \right] = r(s_0) + 5\gamma.$$

the safe action is not prefered.

Qualitative Possibility Theory

- **simplification/imprecision** taken into account, BUT frequentist information lost;
- **ignorance** modeling;
- possibilistic belief states already studied: π-POMDP (Sabbadin UAI98, Drougard UAI13, AAAI14).

Qualitative Possibility Theory

- **simplification/imprecision** taken into account,
  BUT frequentist information lost;
- **ignorance** modeling;
- possibilistic belief states already studied: π-POMDP
  (Sabbadin UAI98, Drougard UAI13, AAAI14).

- **POMDP with possibilistic belief states**
  $\rightarrow$ heuristic for solving POMDPs;
  $\rightarrow$ results in a standard MDP.

Qualitative Possibility Theory

- **simplification/imprecision** taken into account,
  BUT frequentist information lost;
- **ignorance** modeling;
- possibilistic belief states already studied: π-POMDP
  (Sabbadin UAI98, Drougard UAI13, AAAI14).

- **POMDP with possibilistic belief states**
  $\rightarrow$ heuristic for solving POMDPs;
  $\rightarrow$ results in a standard MDP.

- defined distributions π:
  $\mathbb{P} \rightarrow \pi$ transformations: pignistic, specific, ...

$1 = l_1 > l_2 > \ldots > l_{\#\mathcal{L}} = 0$ form the **finite scale** $\mathcal{L}$.

events $e \subset \Omega$ (universe)

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

$\neq$

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

$1 = l_1 > l_2 > \ldots > l_{\#\mathcal{L}} = 0$ form the **finite scale** $\mathcal{L}$.

events $e \subset \Omega$ (universe)

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

$\neq$

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

$e_1 \neq e_2$, 2 events $\subset \Omega$

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

$1 = l_1 > l_2 > \ldots > l_{\#\mathcal{L}} = 0$ form the **finite scale** $\mathcal{L}$.

events $e \subset \Omega$ (universe)

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

$\neq$

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

$e_1 \neq e_2$, 2 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	$\mathbf{p}(e_1) + \mathbf{p}(e_2 \cap \overline{e_1})$	$\max\{\pi(e_1), \pi(e_2)\}$
e_1 and e_2	$\mathbf{p}(e_1).\mathbf{p}(e_2 \mid e_1)$	$\min\{\pi(e_1), \pi(e_2 \mid e_1)\}$

Plan

Possibilistic models: π-MOMDPs

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

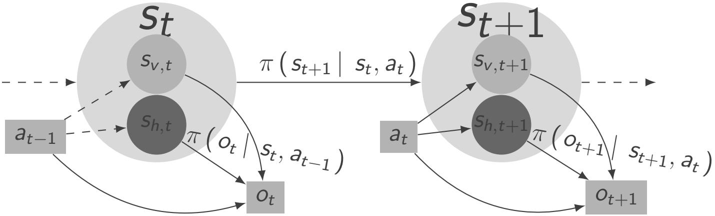
- finite belief space $\#\mathcal{B} = \#\mathcal{L}^{\#\mathcal{S}} - (\#\mathcal{L} - 1)^{\#\mathcal{S}}$

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 $s \in \mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$

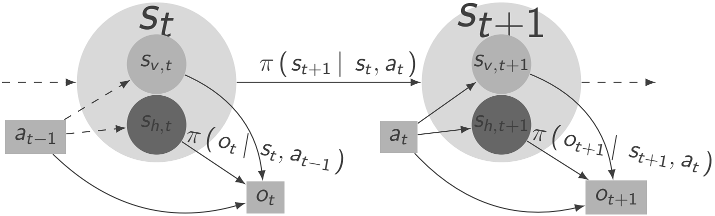
i.e. state s = visible component s_v & hidden component s_h.

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

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

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**,

contribution (UAI13): Infinite Horizon

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

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

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

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 $< \#\mathcal{X}$.

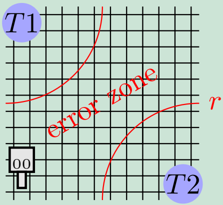
$\forall x \in \mathcal{X}$, $V_0(x) = \mu(x)$ **preference**, and, until convergence,

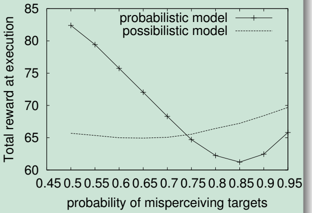
• $V_{i+1}(x) = \max\limits_{a \in \mathcal{A}} \max\limits_{x' \in \mathcal{X}} \min \left\{ \pi\left(x' \mid x, a \right), V_i(x') \right\}$, and

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

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: $\mathbf{p}\left(o' \mid s', a \right)$.

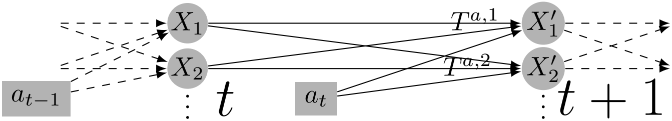




Actually, misperception in the error zone is: $P_{bad} > \frac{1}{2}$.

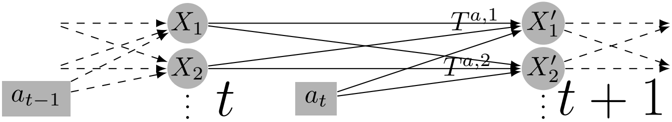
Plan

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.

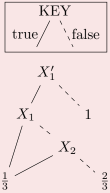


Factorization and symbolic solver

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.



- transition functions
  $T_i^a = \pi\left(X'_i \mid parents(X'_i), a\right)$
  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) ;

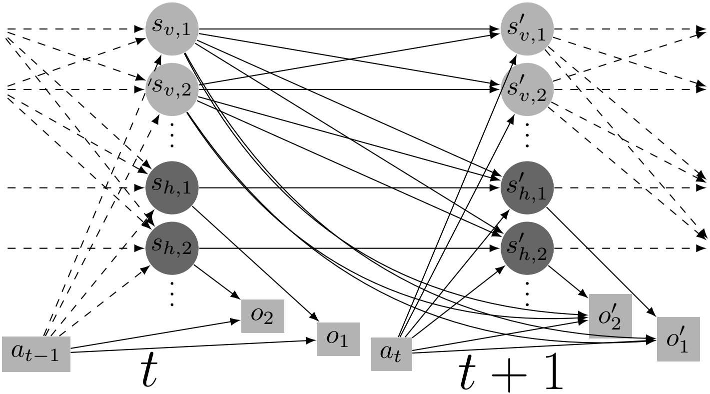
factorization
$\Rightarrow$ dynamic programming
divided into *n* stages

used ADDs smaller
$\rightarrow$ **faster computations.**

computations on trees: *CU Decision Diagram Package*.

Natural factorisation: belief independence.

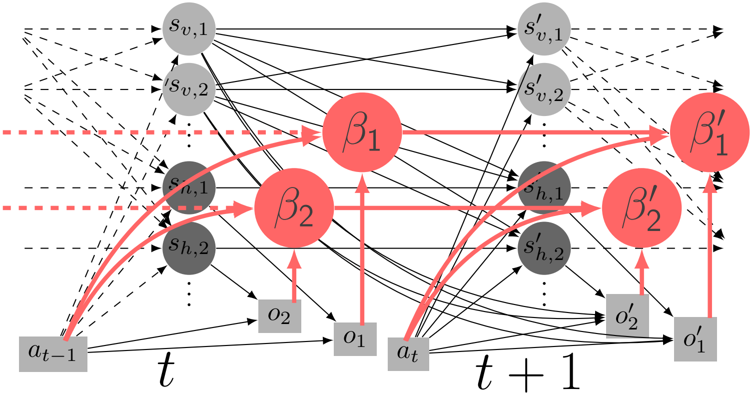
contribution (AAAI14): π-MOMDP following independence assumptions of the graphical model:
$\Rightarrow$ $(s_v, \beta) = (s_{v,1}, \ldots, s_{v,m}, \beta_1, \ldots, \beta_l)$, β_i belief over $s_{h,i}$.



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 $s_{h,i}$.

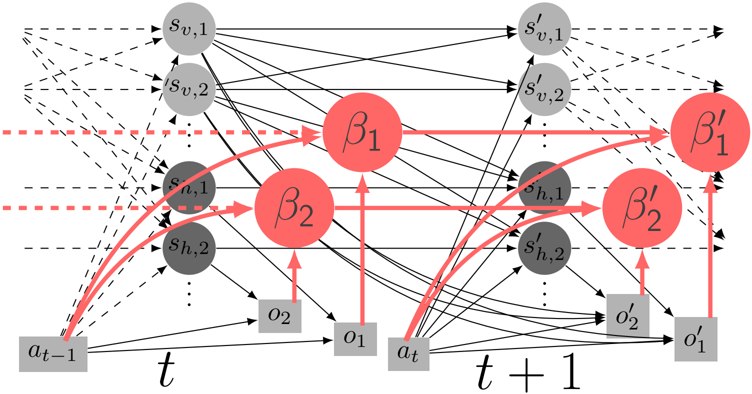


Natural factorisation: belief independence.

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...



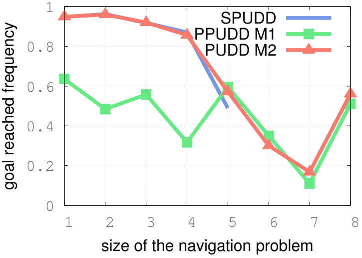
Experiments: Navigation problem – agent = robot.

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

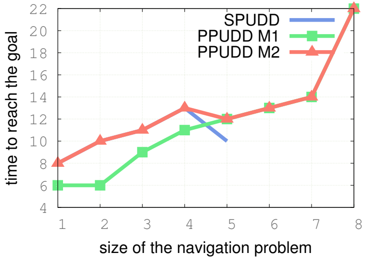
PPUDD vs SPUDD (*Hoey et al.*)
Navigation benchmark: reach a goal; spots with accident risk.
2 possibilistic translations: M1 (optimistic) et M2 (cautious).

Performances, function of the instance size

reached goal frequency

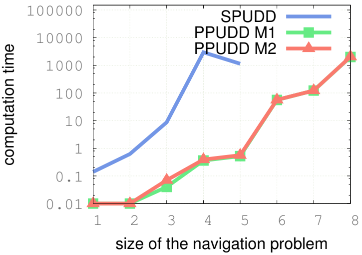


time to reach the goal

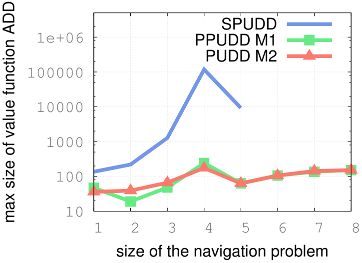


Experiments: Navigation problem – agent = robot.

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.

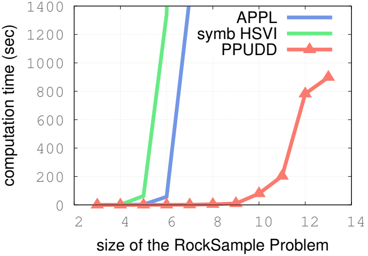
Experiments: RockSample problem – agent = robot.

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

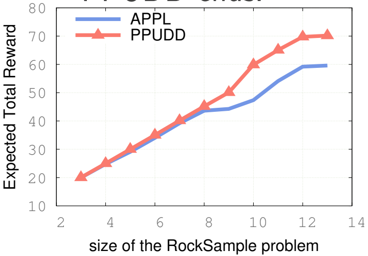
Experiments: RockSample problem – agent = robot.

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.

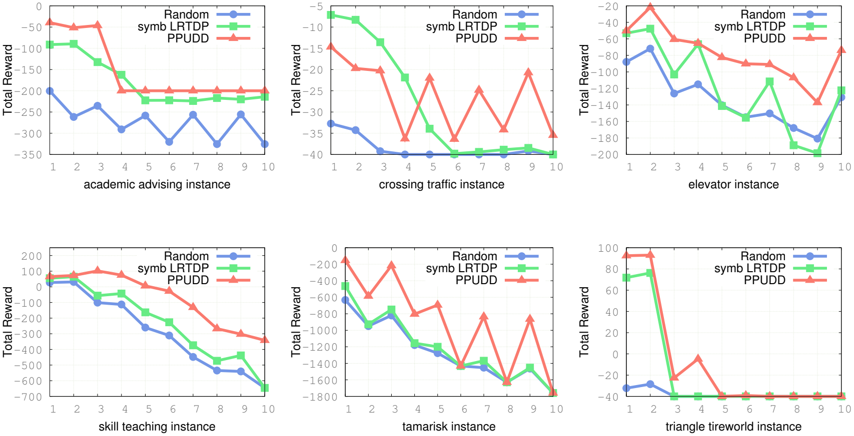


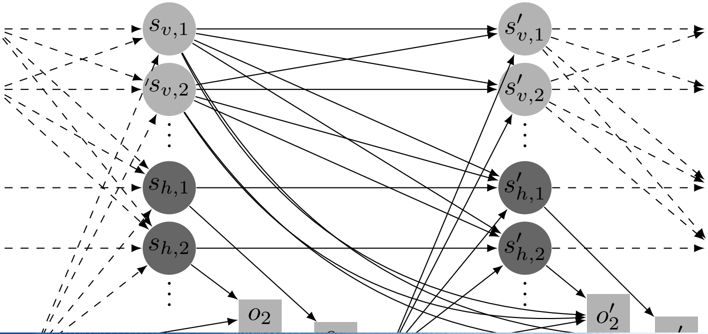
Figure : mean of rewards over simulations.

Plan

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 $s_{h,i}$.

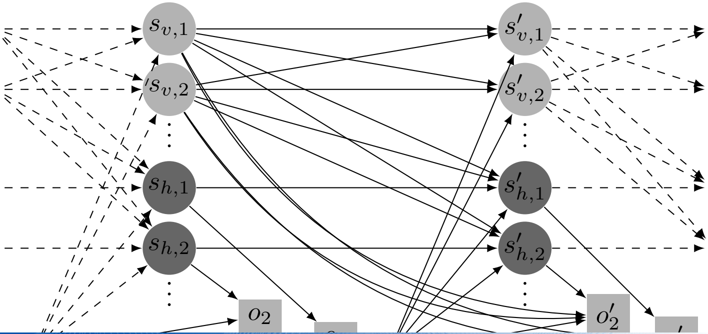


Independent beliefs

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

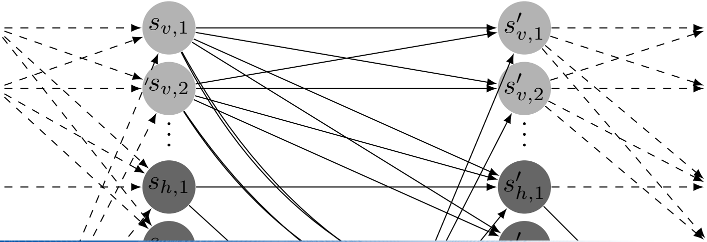
$(s_v, \beta) = (s_{v,1}, \ldots, s_{v,m}, \beta_1, \ldots, \beta_l)$, with β_i belief about $s_{h,i}$.

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



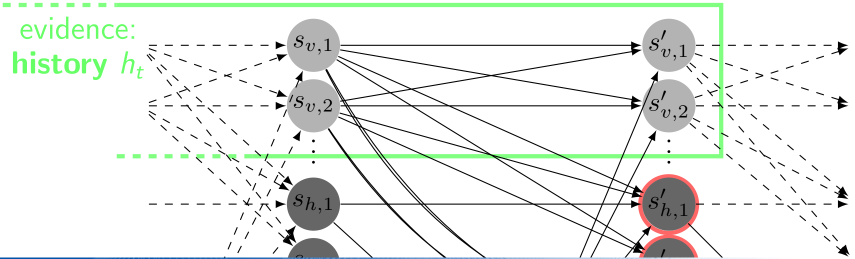
beginning of the proof:

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



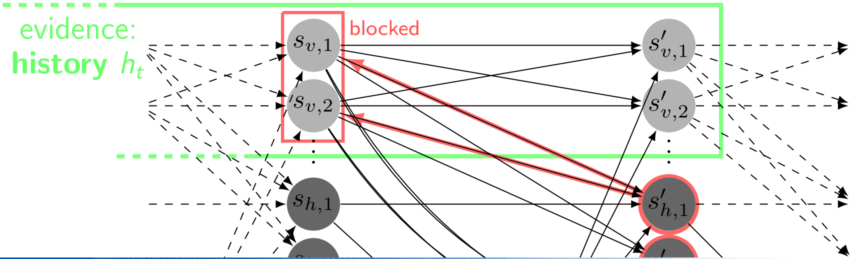
beginning of the proof :

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



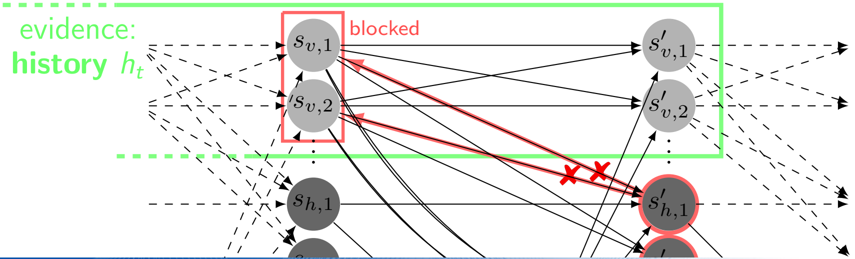
beginning of the proof :

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



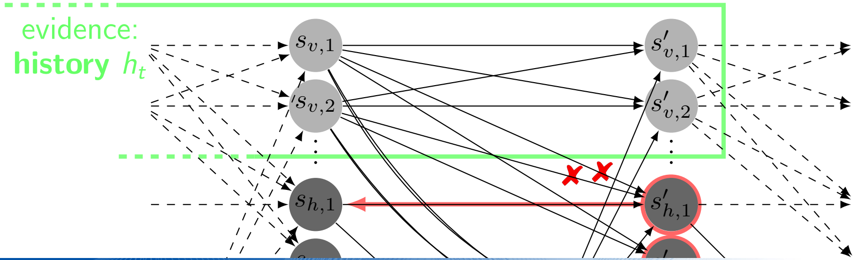
beginning of the proof :

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



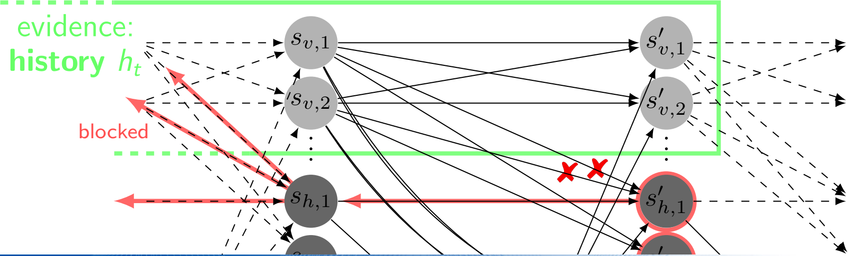
beginning of the proof:

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



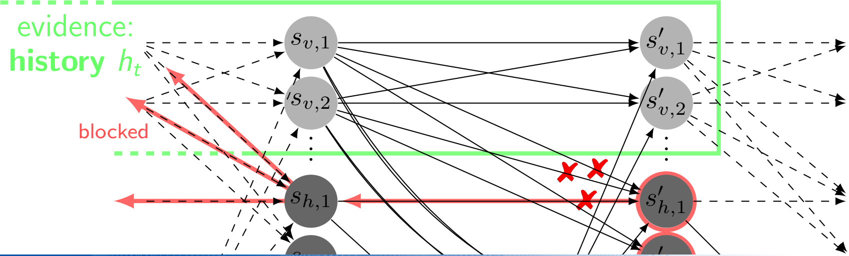
beginning of the proof :

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



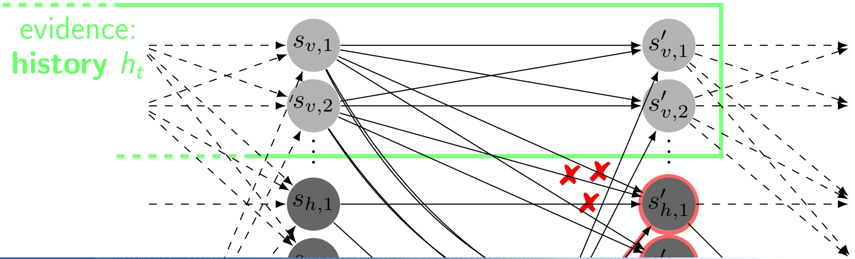
beginning of the proof :

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



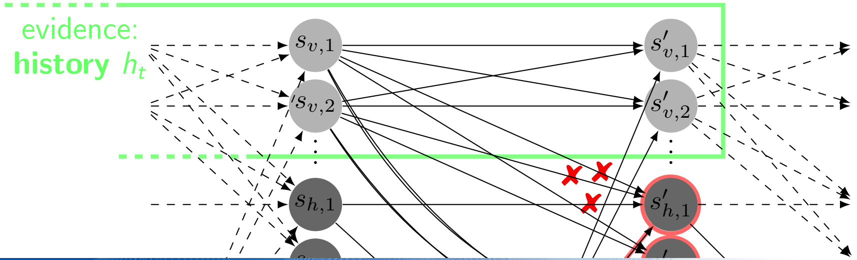
beginning of the proof:

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



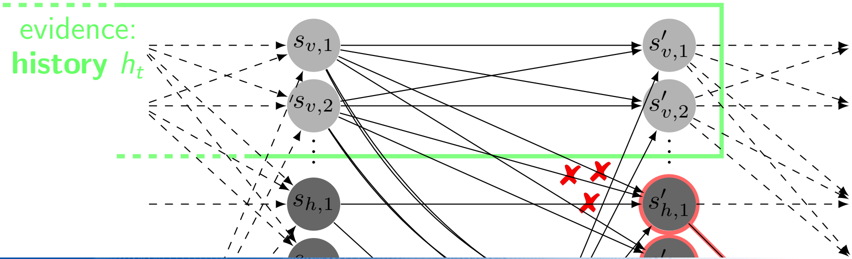
beginning of the proof:

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



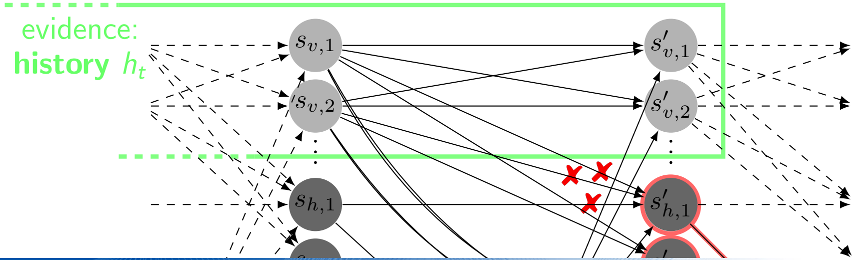
beginning of the proof:

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



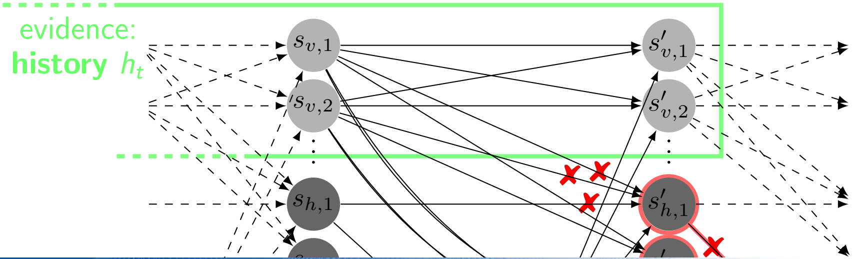
beginning of the proof :

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ 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)

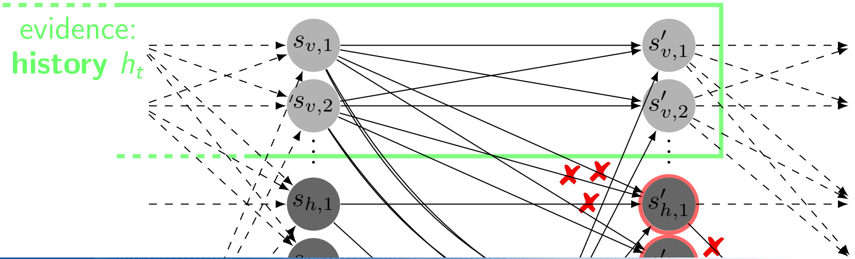


beginning of the proof :

- $\forall 1 \leqslant i < j \leqslant l$, $s_{h,i}$ 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

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



Possibility Theory:

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

Belief factorization

towards a hybrid POMDP

Possibility Theory:

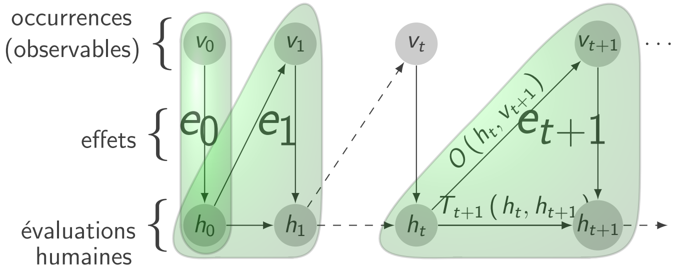
- **granulated** belief space representation (discretization),
- efficient problem **simplification** (PPUDD $2\times$ better than LRTDP with ADDs);
- **ignorance and imprecision** modeling.

- 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.

Plan

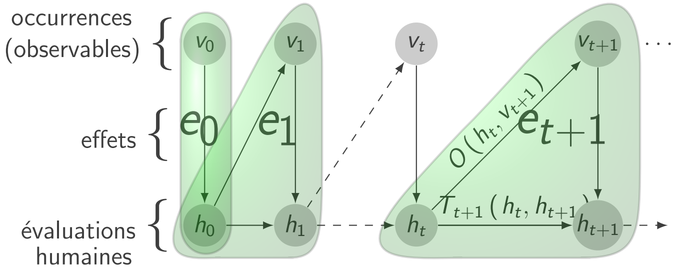
Processus π-MOMDPs, outils de diagnostic pour l'Intéraction Homme-Machine (avec Sergio Pizziol)

- **occurrences:** états de la machine et actions humaines;
- **évaluation humaine** (de l'état de la machine);
- **effets:** transitions, classées par degrés de possibilité.



Processus π-MOMDPs, outils de diagnostic pour l'Intéraction Homme-Machine (avec Sergio Pizziol)

- **occurrences:** états de la machine et actions humaines;
- **évaluation humaine** (de l'état de la machine);
- **effets:** transitions, classées par degrés de possibilité.



- **estimation** de l'état selon l'opérateur humain;
- **détection** des erreurs humaines d'évaluation de l'état;
- causes plausibles de ces erreurs (**diagnostique**).

Plan

$$\Pi_S = \left\{ \text{ possibility distributions } \right\} \colon \#\Pi_S < +\infty$$

$\rightarrow$ **belief space discretization**.

A possibilistic belief state
belief space discretization

$$\Pi_S = \left\{ \text{ possibility distributions } \right\} : \ \#\Pi_S < +\infty$$

$\rightarrow$ **belief space discretization**.

$$b_t^{\pi}(s) = \pi \left(s_t = s \,\middle|\, a_0, o_1, \ldots, a_{t-1}, o_t \right)$$

$\Pi_S = \left\{ \text{ possibility distributions } \right\}$: $\#\Pi_S < +\infty$

$\rightarrow$ **belief space discretization**.

$$b_t^{\pi}(s) = \pi\left(s_t = s \mid a_0, o_1, \dots, a_{t-1}, o_t \right)$$

update – possibilistic belief state

$$b_{t+1}^{\pi}(s') = \begin{cases} 1 & \text{if } \pi\left(o', s' \mid b_t^{\pi}, a\right) = \pi\left(o' \mid b_t^{\pi}, a\right) \\ \pi\left(o', s' \mid b_t^{\pi}, a\right) & \text{otherwise.} \end{cases}$$

denoted by $b_{t+1}^{\pi}(s') \propto^{\pi} \pi\left(o', s' \mid b_t^{\pi}, a\right)$

A possibilistic belief state
belief space discretization

$$\Pi_S = \left\{ \text{ possibility distributions } \right\}: \#\Pi_S < +\infty$$

$\rightarrow$ **belief space discretization.**

$$b_t^\pi(s) = \pi\left(s_t = s \,\middle|\, a_0, o_1, \dots, a_{t-1}, o_t \right)$$

update – possibilistic belief state

$$b_{t+1}^\pi(s') = \begin{cases} 1 & \text{if } \pi\left(o', s' \,\middle|\, b_t^\pi, a\right) = \pi\left(o' \,\middle|\, b_t^\pi, a\right) \\ \pi\left(o', s' \,\middle|\, b_t^\pi, a\right) & \text{otherwise.} \end{cases}$$

denoted by $b_{t+1}^\pi(s') \propto^\pi \pi\left(o', s' \,\middle|\, b_t^\pi, a\right)$

- $\pi(o', s' | b_t^\pi, a) = \max_{s \in \mathcal{S}} \min \left\{ \pi(o' | s', a), \pi(s' | s, a), b_t^\pi(s) \right\};$
- $\pi\left(o' \,\middle|\, s', a\right) = \max_{s' \in \mathcal{S}} \pi\left(o', s' \,\middle|\, b_t^\pi, a\right).$

A possibilistic belief state
belief space discretization

$$\Pi_S = \left\{ \text{ possibility distributions } \right\} : \ \#\Pi_S < +\infty$$

$\rightarrow$ **belief space discretization**.

$$b_t^\pi(s) = \pi\left(s_t = s \mid a_0, o_1, \dots, a_{t-1}, o_t \right)$$

update – possibilistic belief state

$$b_{t+1}^\pi(s') = \begin{cases} 1 & \text{if } \pi\left(o', s' \mid b_t^\pi, a\right) = \pi\left(o' \mid b_t^\pi, a\right) \\ \pi\left(o', s' \mid b_t^\pi, a\right) & \text{otherwise.} \end{cases}$$

denoted by $b_{t+1}^\pi(s') \propto^\pi \pi\left(o', s' \mid b_t^\pi, a\right)$

- the update **only depends on** o' **and** a.

numbering of the $n = \#\mathcal{S}$ system states:
$1 = b^\pi(s_1) \geqslant \ldots \geqslant b^\pi(s_n) \geqslant b^\pi(s_{n+1}) = 0.$

pignistic transformation – $P : \Pi_{\mathcal{S}} \rightarrow \mathbb{P}_{\mathcal{S}}$

$$\overline{b^\pi}(s_i) = \sum_{j=i}^{\#\mathcal{S}} \frac{b^\pi(s_j) - b^\pi(s_{j+1})}{j}.$$

- probability distribution $\overline{b^\pi}$ = **gravity center** of the represented probabilistic distributions;
- **Laplace principle**: ignorance $\rightarrow$ uniform probability.

Pignistic transformation

Examples of pignistic transformations (red) of possibility distributions (blue)

Approximation of the probabilities over the observations:

- $\mathbf{p}\left(o' \mid s, a\right) = \sum_{s' \in \mathcal{S}} O(s', a, o') \cdot T(s, a, s')$;
- $\mathbf{p}\left(o' \mid b^{\pi}, a\right) := \sum_{s \in \mathcal{S}} \mathbf{p}\left(o' \mid s, a\right) \cdot \overline{b^{\pi}}(s)$.

$$\Rightarrow \mathbf{p}\Big((b^{\pi})' \Big| b^{\pi}, a\Big) = \sum_{\substack{o' \text{ t.q.} \\ u(b^{\pi}, a, o') = (b^{\pi})'}} \mathbf{p}\left(o' \mid b^{\pi}, a\right).$$

notation: if $a \in \mathcal{A}$ selected, $o' \in \mathcal{O}$ received,

$$b^{\pi}_{t+1} = u\left(o', a, b^{\pi}_t\right) = \text{ update of } b^{\pi}_t.$$

imprecision of b^{π} = agent ignorance + discretization:
pessimistic reward about these imprecisions.

Choquet integral and rewards
pessimistic evaluation of the rewards – necessity measure

imprecision of b^π = agent ignorance + discretization:
pessimistic reward about these imprecisions.

Dual measure of $\Pi : 2^{\mathcal{S}} \rightarrow \mathcal{L}$

necessity $\mathcal{N}$ such that $\forall A \subseteq \mathcal{S}$, $\mathcal{N}(A) = 1 - \Pi(\overline{A})$.

imprecision of b^π = agent ignorance + discretization:
pessimistic reward about these imprecisions.

Dual measure of $\Pi : 2^{\mathcal{S}} \rightarrow \mathcal{L}$

necessity $\mathcal{N}$ such that $\forall A \subseteq \mathcal{S}$, $\mathcal{N}(A) = 1 - \Pi(\overline{A})$.

$r_1 > r_2 > \ldots > r_{k+1} = 0$ represents elements of $\left\{ r(s,a) \middle| s \in \mathcal{S} \right\}$.

Choquet integral of r with respect to $\mathcal{N}$

$$Ch(r,\mathcal{N}) \ = \ \sum_{i=1}^{k} (r_i - r_{i+1}) \cdot \mathcal{N}\left(\left\{ r(s) \geqslant r_i \right\} \right) \tag{1}$$

$$\tag{2}$$

Choquet integral and rewards

pessimistic evaluation of the rewards – necessity measure

imprecision of b^π = agent ignorance + discretization: **pessimistic reward** about these imprecisions.

Dual measure of $\Pi : 2^{\mathcal{S}} \to \mathcal{L}$

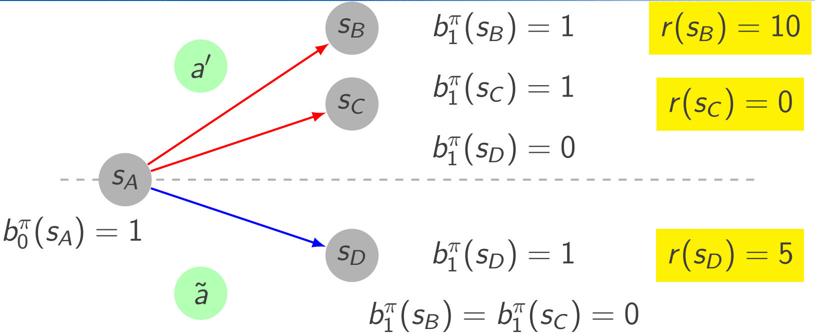
necessity $\mathcal{N}$ such that $\forall A \subseteq \mathcal{S}$, $\mathcal{N}(A) = 1 - \Pi(\overline{A})$.

$r_1 > r_2 > \ldots > r_{k+1} = 0$ represents elements of $\left\{ r(s,a) \middle| s \in \mathcal{S} \right\}$.

Choquet integral of r with respect to $\mathcal{N}$

$$\begin{aligned} Ch(r, \mathcal{N}) &= \sum_{i=1}^{k} (r_i - r_{i+1}) \cdot \mathcal{N}\left(\left\{ r(s) \geqslant r_i \right\} \right) && (1) \\ &= \sum_{i=1}^{\#\mathcal{L}-1} \left(l_i - l_{i+1} \right) \cdot \min_{\substack{s \in \mathcal{S} \text{ s.t.} \\ b^\pi(s) \geqslant l_i}} r(s). && (2) \end{aligned}$$

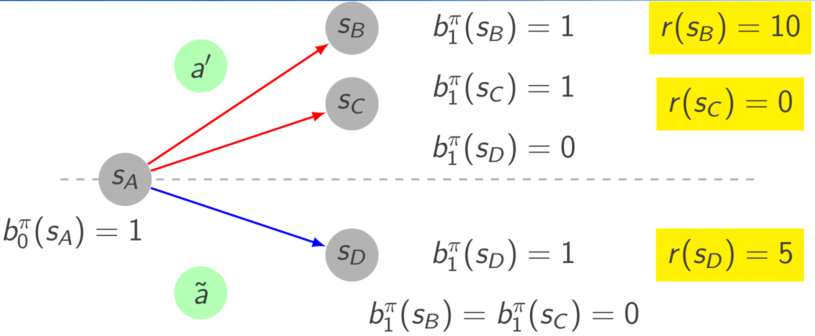
notation $\mathcal{L} = \left\{ l_1 = 1, l_2, l_3, \ldots, 0 \right\}$.



s_B
$b_1^\pi(s_B) = 1$
$r(s_B) = 10$
a'
s_C
$b_1^\pi(s_C) = 1$
$r(s_C) = 0$
$b_1^\pi(s_D) = 0$
s_A
$b_0^\pi(s_A) = 1$
s_D
$b_1^\pi(s_D) = 1$
$r(s_D) = 5$
$\tilde{a}$
$b_1^\pi(s_B) = b_1^\pi(s_C) = 0$

Choquet integral and rewards

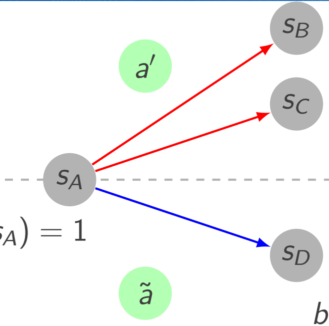
back to the example about ignorance



- $Ch\left(r, N_{b_1^\pi} \;\middle|\; a_0 = \tilde{a}\right) = r(s_D, \tilde{a}) = 5,$
- $Ch\left(r, N_{b_1^\pi} \;\middle|\; a_0 = a'\right) = \min_{s \in \mathcal{S}} r(s, a') = 0.$

the safe action is prefered! **dispersion reduced**

back to the example about ignorance



$b_1^\pi(s_B) = 1$ $\quad$ $r(s_B) = 10$

$b_1^\pi(s_C) = 1$ $\quad$ $r(s_C) = 0$

$b_1^\pi(s_D) = 0$

$b_0^\pi(s_A) = 1$

$b_1^\pi(s_D) = 1$ $\quad$ $r(s_D) = 5$

$b_1^\pi(s_B) = b_1^\pi(s_C) = 0$

- $Ch\left(r, N_{b_1^\pi} \;\middle|\; a_0 = \tilde{a}\right) = r(s_D, \tilde{a}) = 5,$
- $Ch\left(r, N_{b_1^\pi} \;\middle|\; a_0 = a'\right) = \min_{s \in \mathcal{S}} r(s, a') = 0.$

the safe action is prefered! **dispersion reduced**

if $\mathcal{N}_{b_1^\pi}$ replaced by $b_1 \Rightarrow Ch(r, b_1) = \mathbb{E}_{s \sim b_1}\left[r(s,a)\right].$

resulting MDP
translation summary

input: a POMDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$;
output: the MDP $\langle \tilde{\mathcal{S}}, \mathcal{A}, \tilde{T}, \tilde{r}, \gamma \rangle$:

input: a POMDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$;
output: the MDP $\langle \tilde{\mathcal{S}}, \mathcal{A}, \tilde{T}, \tilde{r}, \gamma \rangle$:

- **state space** $\tilde{\mathcal{S}} = \Pi_{\mathcal{S}}$,
  the set of the possibility distributions over $\mathcal{S}$;

input: a POMDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$;
output: the MDP $\langle \tilde{\mathcal{S}}, \mathcal{A}, \tilde{T}, \tilde{r}, \gamma \rangle$:

- **state space** $\tilde{\mathcal{S}} = \Pi_{\mathcal{S}}$,
  the set of the possibility distributions over $\mathcal{S}$;
- $\forall b^\pi, (b^\pi)'$ possibilistic belief states $\in \Pi_{\mathcal{S}}$, $\forall a \in \mathcal{A}$,
  transitions $\tilde{T}\Big(b^\pi, a, (b^\pi)'\Big) = \mathbf{p}\Big((b^\pi)' \Big| b^\pi, a\Big)$;

input: a POMDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$;
output: the MDP $\langle \tilde{\mathcal{S}}, \mathcal{A}, \tilde{T}, \tilde{r}, \gamma \rangle$:

- **state space** $\tilde{\mathcal{S}} = \Pi_{\mathcal{S}}$,
  the set of the possibility distributions over $\mathcal{S}$;
- $\forall b^\pi, (b^\pi)'$ possibilistic belief states $\in \Pi_{\mathcal{S}}$, $\forall a \in \mathcal{A}$,
  transitions $\tilde{T}\Big(b^\pi, a, (b^\pi)'\Big) = \mathbf{p}\Big((b^\pi)' \Big| b^\pi, a\Big)$;
- **reward** $\tilde{r}(a, b^\pi) = Ch\Big(r(a,.), \mathcal{N}_{b^\pi}\Big)$,
  $\mathcal{N}_{b^\pi}$ necessity measure computed from b^π.

input: a POMDP $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$;
output: the MDP $\langle \tilde{\mathcal{S}}, \mathcal{A}, \tilde{T}, \tilde{r}, \gamma \rangle$:

- **state space** $\tilde{\mathcal{S}} = \Pi_{\mathcal{S}}$,
  the set of the possibility distributions over $\mathcal{S}$;
- $\forall b^\pi, (b^\pi)'$ possibilistic belief states $\in \Pi_{\mathcal{S}}$, $\forall a \in \mathcal{A}$,
  transitions $\tilde{T}\Big(b^\pi, a, (b^\pi)'\Big) = \mathbf{p}\Big((b^\pi)' \Big| b^\pi, a\Big)$;
- **reward** $\tilde{r}(a, b^\pi) = Ch\Big(r(a,.), \mathcal{N}_{b^\pi}\Big)$,
  $\mathcal{N}_{b^\pi}$ necessity measure computed from b^π.

criterion: $\mathbb{E}_{(b_t^\pi)\sim\tilde{T}}\left[\sum_{t=0}^{+\infty} \gamma^t \cdot \tilde{r}\left(b_t^\pi, d_t\right)\right]$.

hybrid POMDP and π-POMDP

differences with possibilistic models

	hybrid POMDP	π-POMDP
transitions	probabilities	qualitative possibility
rewards	quantitative $\in \mathbb{R}$	qualitative $\in \mathcal{L}$
situation	-some imprecisions -large POMDP	few quantitative
issues	π definition	commensurability
in practice	MDP	π-MDP

	hybrid POMDP	π-POMDP
transitions	probabilities	qualitative possibility
rewards	quantitative $\in \mathbb{R}$	qualitative $\in \mathcal{L}$
situation	-some imprecisions -large POMDP	few quantitative
issues	π definition	commensurability
in practice	MDP	π-MDP

hybrid model:

- only belief states are possibilistic:

$\rightarrow$ agent knowledge $=$ **possibility** distribution;

- probabilistic dynamics:

$\rightarrow$ **approximated** (prob.) transition between epistemic states.

Plan

factorized POMDP
definition

- $\mathcal{S}$ described by $\mathbb{S} = \{ s_1, \ldots, s_m \}$: $\mathcal{S} = s_1 \times \ldots \times s_m$.
  Notation: $\mathbb{S}' = \{ s'_1, \ldots, s'_m \}$;

- $\mathcal{S}$ described by $\mathbb{S} = \{s_1, \ldots, s_m\}$: $\mathcal{S} = s_1 \times \ldots \times s_m$.
  Notation: $\mathbb{S}' = \{s'_1, \ldots, s'_m\}$;

- **transition** function of s'_j,
  $T^a_j(\mathbb{S}, s'_j) = \mathbf{p}\left(s'_j \,\middle|\, \mathbb{S}, a\right)$, $\forall j \in \{1, \ldots, m\}$ et $\forall a \in \mathcal{A}$;

factorized POMDP
definition

- $\mathcal{S}$ described by $\mathbb{S} = \{ s_1, \ldots, s_m \}$: $\mathcal{S} = s_1 \times \ldots \times s_m$.
  Notation: $\mathbb{S}' = \{ s'_1, \ldots, s'_m \}$;
- **transition** function of s'_j,
  $T^a_j(\mathbb{S}, s'_j) = \mathbf{p}\left(s'_j \,\middle|\, \mathbb{S}, a \right)$, $\forall j \in \{1, \ldots, m\}$ et $\forall a \in \mathcal{A}$;
- $\mathcal{O}$ described by $\mathbb{O} = \{ o_1, \ldots, o_n \}$: $\mathcal{O} = o_1 \times \ldots \times o_n$;

factorized POMDP
definition

- $\mathcal{S}$ described by $\mathbb{S} = \{ s_1, \ldots, s_m \}$: $\mathcal{S} = s_1 \times \ldots \times s_m$.
  Notation: $\mathbb{S}' = \{ s'_1, \ldots, s'_m \}$;
- **transition** function of s'_j,
  $T^a_j(\mathbb{S}, s'_j) = \mathbf{p}\left(s'_j \mid \mathbb{S}, a \right)$, $\forall j \in \{1, \ldots, m\}$ et $\forall a \in \mathcal{A}$;
- $\mathcal{O}$ described by $\mathbb{O} = \{ o_1, \ldots, o_n \}$: $\mathcal{O} = o_1 \times \ldots \times o_n$;
- **observation** function of o'_i,
  $O^a_i(\mathbb{S}', o'_i) = \mathbf{p}\left(o'_i \mid \mathbb{S}', a \right)$, $\forall i \in \{1, \ldots, n\}$ et $\forall a \in \mathcal{A}$.

- $\mathcal{S}$ described by $\mathbb{S} = \{ s_1, \ldots, s_m \}$: $\mathcal{S} = s_1 \times \ldots \times s_m$.
  Notation: $\mathbb{S}' = \{ s'_1, \ldots, s'_m \}$;

- **transition** function of s'_j,
  $T^a_j(\mathbb{S}, s'_j) = \mathbf{p}\left(s'_j \;\middle|\; \mathbb{S}, a \right)$, $\forall j \in \{1, \ldots, m\}$ et $\forall a \in \mathcal{A}$;

- $\mathcal{O}$ described by $\mathbb{O} = \{ o_1, \ldots, o_n \}$: $\mathcal{O} = o_1 \times \ldots \times o_n$;

- **observation** function of o'_i,
  $O^a_i(\mathbb{S}', o'_i) = \mathbf{p}\left(o'_i \mid \mathbb{S}', a \right)$, $\forall i \in \{1, \ldots, n\}$ et $\forall a \in \mathcal{A}$.

independences:

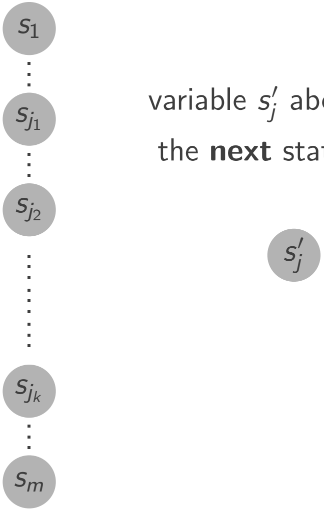
$\rightarrow \forall s'_i, s'_j \in \mathbb{S}', \qquad s'_i \perp\!\!\!\perp s'_j \mid \{\mathbb{S}, a \in \mathcal{A}\}$,

$\rightarrow \forall o'_i, o'_j \in \mathbb{O}', \qquad o'_i \perp\!\!\!\perp o'_j \mid \{\mathbb{S}', a \in \mathcal{A}\}$.

Notations

some variables does not interact with each other

variables about the **current** system state,



variable s'_j about the **next** state.

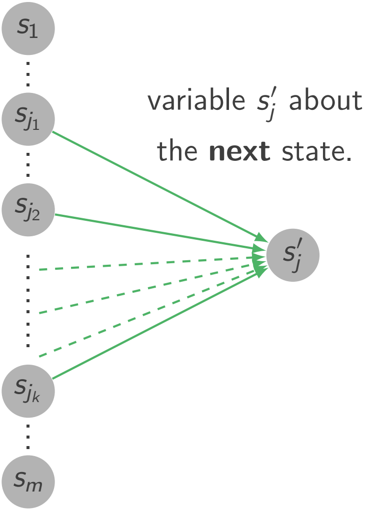
Notations

some variables does not interact with each other

variables about the **current** system state,

$$s_k \to s'_j$$
$$\Updownarrow$$
$\exists a \in \mathcal{A}$, such that

$T^a_j(\mathbb{S}, s'_j)$ depends on s_k.

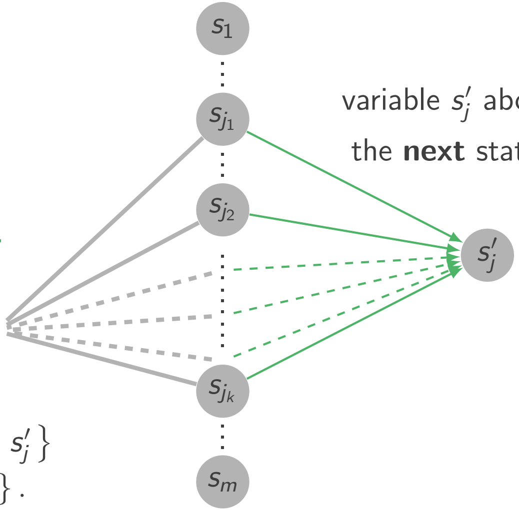


variables about the **current** system state,

$$s_k \to s'_j$$
$$\Updownarrow$$
$$\exists a \in \mathcal{A}, \text{ such that}$$
$$T^a_j(\mathbb{S}, s'_j) \text{ depends on } s_k.$$

variable s'_j about the **next** state.

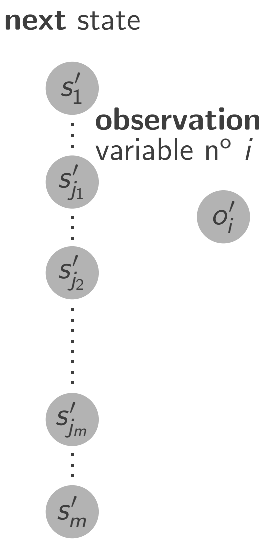
$\mathcal{P}(s'_j)$



the parents of s'_j are:

$$\mathcal{P}(s'_j) = \{\, s \in \mathbb{S} \mid s \to s'_j \,\}$$
$$= \{\, s_{j_1}, s_{j_2}, \ldots, s_{j_k} \,\}.$$

Notations
concerning observation variables



Notations

concerning observation variables

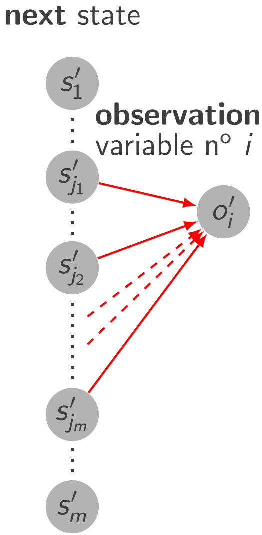
$$s'_j \to o'_i$$

$$\Updownarrow$$

$\exists a \in \mathcal{A}$, such that

$$O_i^a(\mathbb{S}', o'_i)$$

depends on s'_j.



Notations

concerning observation variables

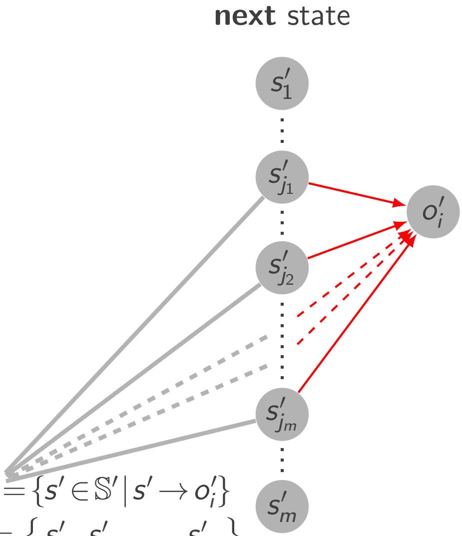
$$s'_j \to o'_i$$

$$\Updownarrow$$

$\exists a \in \mathcal{A}$, such that

$$O_i^a(\mathbb{S}', o'_i)$$

depends on s'_j.



$$\mathcal{P}(o'_i) = \{s' \in \mathbb{S}' \,|\, s' \to o'_i\}$$
$$= \left\{ s'_{j1}, s'_{j2}, \dots, s'_{jm} \right\}$$

Notations

concerning observation variables

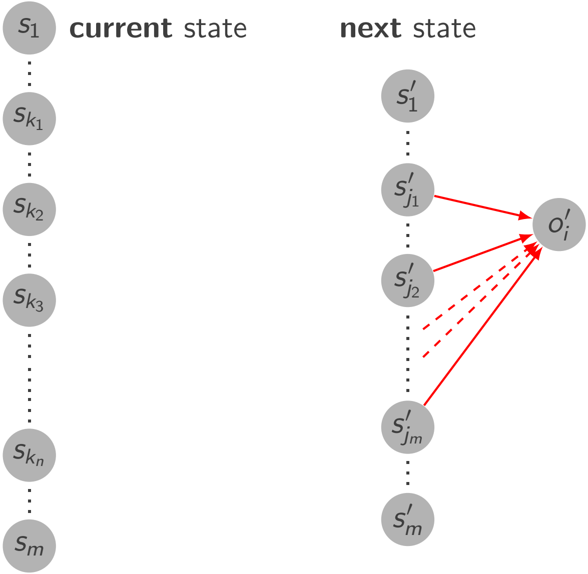
$s'_j \to o'_i$

$\Updownarrow$

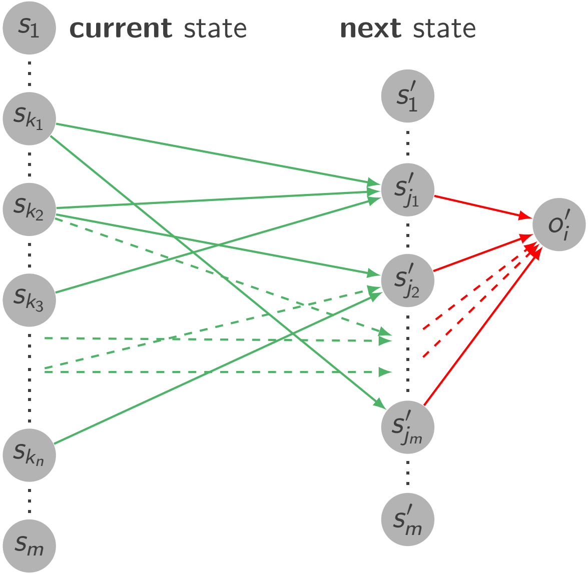
$\exists a \in \mathcal{A}$, such that

$O^a_i(\mathbb{S}', o'_i)$

depends on s'_j.



$$s'_j \to o'_i$$
$$\Updownarrow$$
$\exists a \in \mathcal{A}$, such that
$$O_i^a(\mathbb{S}', o'_i)$$
depends on s'_j.



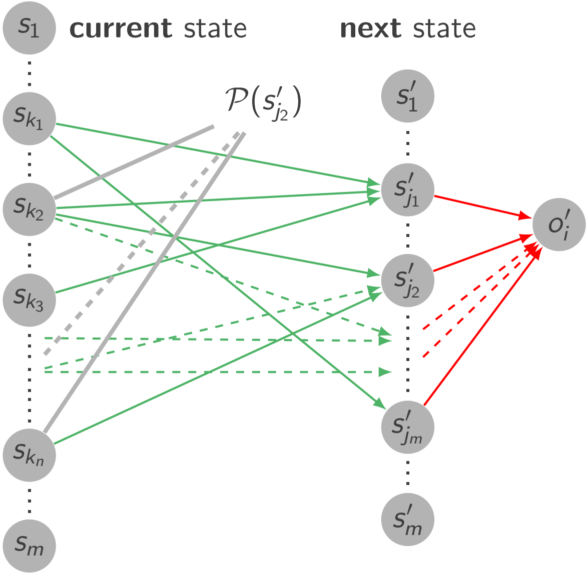
$s'_j \to o'_i$

$\Updownarrow$

$\exists a \in \mathcal{A}$, such that

$O_i^a(\mathbb{S}', o'_i)$

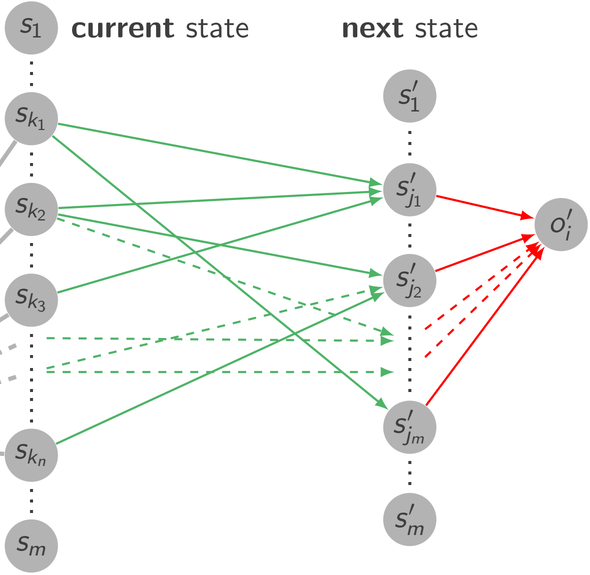
depends on s'_j.



$$s'_j \to o'_i$$
$$\Updownarrow$$
$\exists a \in \mathcal{A}$, such that

$O_i^a(\mathbb{S}', o'_i)$

depends on s'_j.



$$\mathcal{Q}(o'_i) = \bigcup_{s'_j \in \mathcal{P}(o'_i)} P(s'_j)$$
$$= \{ s_{k_1}, s_{k_2}, \ldots, s_{k_n} \}$$

Rewritings of parameters
PROBABILISTIC parameters

- $T_j^a\left(\mathbb{S}, s'_j\right) = T_j^a\left(\mathcal{P}(s'_j), s'_j\right)$;
- $O_i^a\left(\mathbb{S}', o'_i\right) = O_i^a\left(\mathcal{P}(o'_i), o'_i\right)$.

Rewritings of parameters
PROBABILISTIC parameters

- $T_j^a\left(\mathbb{S}, s_j'\right) = T_j^a\left(\mathcal{P}(s_j'), s_j'\right)$;
- $O_i^a\left(\mathbb{S}', o_i'\right) = O_i^a\left(\mathcal{P}(o_i'), o_i'\right)$.

consequences on the joint distribution

$$\begin{aligned}\mathbf{p}\left(o_i', \mathcal{P}(o_i') \mid \mathbb{S}, a\right) &= O_i^a\left(\mathcal{P}(o_i'), o_i'\right) \cdot \prod_{s_j' \in \mathcal{P}(o_i')} T_i^a\left(\mathcal{P}(s_j'), s_j'\right) \\ &= \mathbf{p}\left(o_i', \mathcal{P}(o_i') \mid \mathcal{Q}(o_i'), a\right).\end{aligned}$$

Rewritings of parameters
PROBABILISTIC parameters

- $T_j^a\left(\mathbb{S}, s'_j\right) = T_j^a\left(\mathcal{P}(s'_j), s'_j\right)$;
- $O_i^a\left(\mathbb{S}', o'_i\right) = O_i^a\left(\mathcal{P}(o'_i), o'_i\right)$.

consequences on the joint distribution

$$\begin{aligned}\mathbf{p}\left(o'_i, \mathcal{P}(o'_i) \mid \mathbb{S}, a\right) &= O_i^a\left(\mathcal{P}(o'_i), o'_i\right) \cdot \prod_{s'_j \in \mathcal{P}(o'_i)} T_i^a\left(\mathcal{P}(s'_j), s'_j\right) \\ &= \mathbf{p}\left(o'_i, \mathcal{P}(o'_i) \mid \mathcal{Q}(o'_i), a\right).\end{aligned}$$

observation probabilities

epistemic state

$$b^{\pi}(\mathbb{S}) \xrightarrow{\textbf{marginalization}} b^{\pi}(\mathcal{Q}(o'_i)) \xrightarrow{\substack{\textbf{pignistic} \\ \textbf{transformation}}} \overline{b^{\pi}}(\mathcal{Q}(o'_i))$$

$$\mathbf{p}\left(o'_i \mid b^{\pi}, a\right) = \sum_{2^{\mathcal{P}(o'_i)}, 2^{\mathcal{Q}(o'_i)}} \mathbf{p}\left(o'_i, \mathcal{P}(o'_i) \mid \mathcal{Q}(o'_i), a\right) \cdot \overline{b^{\pi}}(\mathcal{Q}(o'_i))$$

Parameters rewritings
POSSIBILISTIC parameters

- $\pi\left(s'_j \mid \mathbb{S}, a\right) = \pi\left(s'_j \mid \mathcal{P}(s'_j), a\right)$;
- $\pi\left(o'_i \mid \mathbb{S}', a\right) = \pi\left(o'_i \mid \mathcal{P}(o'_i), a\right)$.

- $\pi\left(s'_j \mid \mathbb{S}, a\right) = \pi\left(s'_j \mid \mathcal{P}(s'_j), a\right)$;
- $\pi\left(o'_i \mid \mathbb{S}', a\right) = \pi\left(o'_i \mid \mathcal{P}(o'_i), a\right)$.

marginal possibilistic belief states

$\forall o'_i \in \mathbb{O}$,

$$b^{\pi}_{t+1}\Big(\mathcal{P}(o'_i)\Big) \propto^{\pi} \pi\Big(o'_i, \mathcal{P}(o'_i)\Big| a_0, o_1, \ldots, a_{t-1}, o_t\Big)$$

Parameters rewritings
POSSIBILISTIC parameters

- $\pi\left(s'_j \mid \mathbb{S}, a\right) = \pi\left(s'_j \mid \mathcal{P}(s'_j), a\right)$;
- $\pi\left(o'_i \mid \mathbb{S}', a\right) = \pi\left(o'_i \mid \mathcal{P}(o'_i), a\right)$.

marginal possibilistic belief states

$\forall o'_i \in \mathbb{O}$,

$$b^{\pi}_{t+1}\Big(\mathcal{P}(o'_i)\Big) \propto^{\pi} \pi\Big(o'_i, \mathcal{P}(o'_i)\Big| a_0, o_1, \ldots, a_{t-1}, o_t\Big)$$
$$= \max_{2^{\mathcal{Q}(o'_i)}} \min\left\{\pi\Big(o'_i, \mathcal{P}(o'_i)\Big|\mathcal{Q}(o'_i), a\Big), b^{\pi}_t\Big(\mathcal{Q}(o'_i)\Big)\right\}$$

- $\pi \left(s'_j \mid \mathbb{S}, a \right) = \pi \left(s'_j \mid \mathcal{P}(s'_j), a \right)$;
- $\pi \left(o'_i \mid \mathbb{S}', a \right) = \pi \left(o'_i \mid \mathcal{P}(o'_i), a \right)$.

marginal possibilistic belief states

$\forall o'_i \in \mathbb{O}$,

$$b^{\pi}_{t+1}\Big(\mathcal{P}(o'_i)\Big) \propto^{\pi} \pi\Big(o'_i, \mathcal{P}(o'_i) \Big| a_0, o_1, \ldots, a_{t-1}, o_t\Big)$$

$$= \max_{2^{\mathcal{Q}(o'_i)}} \min \left\{ \pi\Big(o'_i, \mathcal{P}(o'_i) \Big| \mathcal{Q}(o'_i), a\Big), b^{\pi}_t\Big(\mathcal{Q}(o'_i)\Big) \right\}$$

denoted by $\pi\Big(o'_i, \mathcal{P}(o'_i) \Big| b^{\pi}_t, a\Big)$.

Variable classification
3 classes of state variables

variable: visible $s_v \in \mathbb{S}_v$

s'_v

inferred hidden $s_h \in \mathbb{S}_h$

s'_h

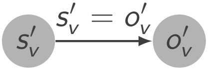
fully hidden $s_f \in \mathbb{S}_f$

s'_f

Variable classification

3 classes of state variables

variable: visible $s_v \in \mathbb{S}_v$



inferred hidden $s_h \in \mathbb{S}_h$

fully hidden $s_f \in \mathbb{S}_f$

Variable classification

3 classes of state variables

variable: visible $s_v \in \mathbb{S}_v$

$$s'_v \xrightarrow{s'_v = o'_v} o'_v$$

$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\Big(\mathcal{P}(s'_v)\Big).$$

inferred hidden $s_h \in \mathbb{S}_h$

s'_h

fully hidden $s_f \in \mathbb{S}_f$

s'_f

<u>variable:</u> visible $s_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable.

$$s'_v \xrightarrow{s'_v = o'_v} o'_v$$

$$\mathbf{p}\left(s'_v \mid b^{\pi}_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^{\pi}_t}\Big(\mathcal{P}(s'_v)\Big).$$

---

inferred hidden $s_h \in \mathbb{S}_h$

s'_h

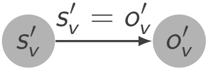
---

fully hidden $s_f \in \mathbb{S}_f$

s'_f

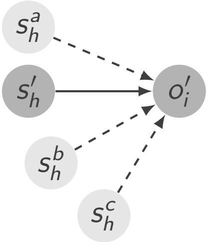
variable: visible $s_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable.



$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\Big(\mathcal{P}(s'_v)\Big).$$

inferred hidden $s_h \in \mathbb{S}_h$

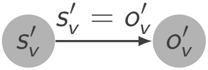


fully hidden $s_f \in \mathbb{S}_f$

<u>variable:</u> visible $s_v \in \mathbb{S}_v$

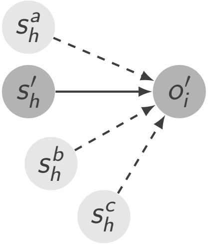
$\Leftrightarrow$ deterministic belief variable.

$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\left(\mathcal{P}(s'_v)\right).$$



inferred hidden $s_h \in \mathbb{S}_h$

$$b^\pi_{t+1}(\mathcal{P}(o'_i)) = b^\pi_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$



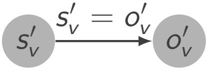
fully hidden $s_f \in \mathbb{S}_f$

ONERA THE FRENCH AEROSPACE LAB

<u>variable:</u> visible $s_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable.

$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\left(\mathcal{P}(s'_v)\right).$$

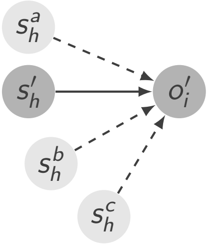


---

inferred hidden $s_h \in \mathbb{S}_h$

$$b^\pi_{t+1}(\mathcal{P}(o'_i)) = b^\pi_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^\pi \pi\left(o'_i, \mathcal{P}(o'_i) \middle| b^\pi_t, a\right).$$

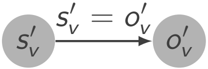


---

fully hidden $s_f \in \mathbb{S}_f$

Variable classification
3 classes of state variables

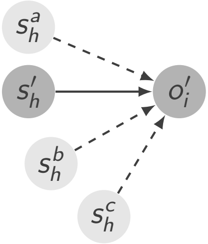
variable: visible $s_v \in \mathbb{S}_v$



$\Leftrightarrow$ deterministic belief variable.

$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\left(\mathcal{P}(s'_v)\right).$$

inferred hidden $s_h \in \mathbb{S}_h$



$$b^\pi_{t+1}(\mathcal{P}(o'_i)) = b^\pi_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^\pi \pi\left(o'_i, \mathcal{P}(o'_i) \middle| b^\pi_t, a\right).$$

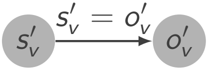
⚠ $\mathcal{P}(o'_i)$ may contain visible variables.

fully hidden $s_f \in \mathbb{S}_f$

<u>variable:</u> visible $s_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable.

$$\mathbf{p}\left(s'_v \mid b^{\pi}_t, a \right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^{\pi}_t}\left(\mathcal{P}(s'_v)\right).$$

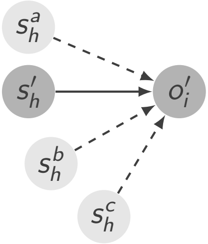


inferred hidden $s_h \in \mathbb{S}_h$

$$b^{\pi}_{t+1}(\mathcal{P}(o'_i)) = b^{\pi}_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^{\pi} \pi\left(o'_i, \mathcal{P}(o'_i) \middle| b^{\pi}_t, a\right).$$

⚠ $\mathcal{P}(o'_i)$ may contain visible variables.



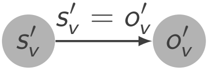
fully hidden $s_f \in \mathbb{S}_f$

Variable classification

3 classes of state variables

<ins>variable:</ins> visible $s_v \in \mathbb{S}_v$

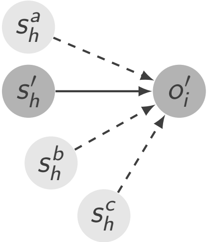
$\Leftrightarrow$ deterministic belief variable.



$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\Big(\mathcal{P}(s'_v)\Big).$$

---

inferred hidden $s_h \in \mathbb{S}_h$



$$b^\pi_{t+1}(\mathcal{P}(o'_i)) = b^\pi_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^\pi \pi\left(o'_i, \mathcal{P}(o'_i) \,\middle|\, b^\pi_t, a\right).$$

⚠ $\mathcal{P}(o'_i)$ may contain visible variables.

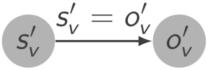
---

fully hidden $s_f \in \mathbb{S}_f$

$$b^\pi_{t+1}(s'_f) = \max_{2^{\mathcal{P}(s'_f)}} \min \left\{ \pi\left(s'_f \middle| \mathcal{P}(s'_f), a\right), b^\pi_t\left(\mathcal{P}(s'_f)\right) \right\}.$$

Variable classification

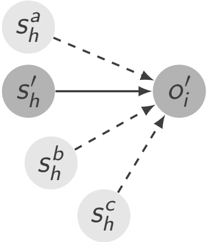
3 classes of state variables



variable: visible $s_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable.

$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\Big(\mathcal{P}(s'_v)\Big).$$



inferred hidden $s_h \in \mathbb{S}_h$

$$b^\pi_{t+1}(\mathcal{P}(o'_i)) = b^\pi_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^\pi \pi\left(o'_i, \mathcal{P}(o'_i) \,\middle|\, b^\pi_t, a\right).$$

⚠ $\mathcal{P}(o'_i)$ may contain visible variables.

---

fully hidden $s_f \in \mathbb{S}_f$

$\rightarrow$ observations don't inform belief state on s'_f.

$$b^\pi_{t+1}(s'_f) = \max_{2^{\mathcal{P}(s'_f)}} \min \left\{ \pi\left(s'_f \middle| \mathcal{P}(s'_f), a\right), b^\pi_t\left(\mathcal{P}(s'_f)\right) \right\}.$$

$\mathbb{O}_h = \mathbb{O} \setminus \mathbb{S}_v.$

bound over the global belief state

$$b^{\pi}_{t+1}(\mathbb{S}') = \pi\left(\mathbb{S}' \,\middle|\, a_0, o_1, \ldots, a_t, o_{t+1}\right)$$

$$\leqslant \beta_{t+1}(\mathbb{S}')$$
$$= \min\left\{ \min_{s'_j \in \mathbb{S}_v} \left[\mathbb{1}_{\left\{ s'_j = o'_j \right\}} \right], \min_{s'_j \in \mathbb{S}_f} \left[b^{\pi}_{t+1}(s'_j) \right], \min_{o'_i \in \mathbb{O}_h} \left[b^{\pi}_{t+1}\left(\mathcal{P}(o'_i)\right) \right] \right\}$$

$$\mathbb{O}_h = \mathbb{O} \setminus \mathbb{S}_v.$$

bound over the global belief state

$$b^{\pi}_{t+1}(\mathbb{S}') = \pi\left(\mathbb{S}' \mid a_0, o_1, \ldots, a_t, o_{t+1}\right)$$

$$\leqslant \beta_{t+1}(\mathbb{S}')$$
$$= \min\left\{ \min_{s'_j \in \mathbb{S}_v} \left[\mathbb{1}_{\{s'_j = o'_j\}} \right], \min_{s'_j \in \mathbb{S}_f} \left[b^{\pi}_{t+1}(s'_j) \right], \min_{o'_i \in \mathbb{O}_h} \left[b^{\pi}_{t+1}\left(\mathcal{P}(o'_i)\right) \right] \right\}$$

- β_t = **less informative** version of the belief state: $b^{\pi}_t \leqslant \beta_t$;
- computed using **marginal belief states** $\leftrightarrow$ factorization.

Variables de croyance
different according to the class of the variable

$$\lambda = \#\mathcal{L}$$

Variables de croyance

different according to the class of the variable

$\lambda = \#\mathcal{L}$

- $\forall s'_v \in \mathbb{S}_v$, 1 variable β'_v is enough.

Variables de croyance
different according to the class of the variable

$\lambda = \#\mathcal{L}$

- $\forall s'_v \in \mathbb{S}_v$, 1 variable β'_v is enough.

- $p_i = \#\mathcal{P}(o'_i)$.

  $\forall o_i \in \mathbb{O} \setminus \mathbb{S}_v$, $\lambda^{2^{p_i}} - (\lambda - 1)^{2^{p_i}}$ belief states,

  $\Rightarrow \lceil \log_2(\lambda^{2^{p_i}} - (\lambda - 1)^{2^{p_i}}) \rceil$ boolean variables β'_h .

Variables de croyance

different according to the class of the variable

$\lambda = \#\mathcal{L}$

- $\forall s'_v \in \mathbb{S}_v$, 1 variable β'_v is enough.
- $p_i = \#\mathcal{P}(o'_i)$.

  $\forall o_i \in \mathbb{O} \setminus \mathbb{S}_v$, $\lambda^{2^{p_i}} - (\lambda - 1)^{2^{p_i}}$ belief states,

  $\Rightarrow \lceil \log_2(\lambda^{2^{p_i}} - (\lambda - 1)^{2^{p_i}}) \rceil$ boolean variables β'_h .
- $\forall s'_f \in \mathbb{S}_f$, $\lambda^2 - (\lambda - 1)^2 = 2\lambda - 1$ belief states,

  $\Rightarrow \lceil \log_2(2\lambda - 1) \rceil$ boolean variables β'_f.

resulting MDP in practice
trick: "flipflop" variable

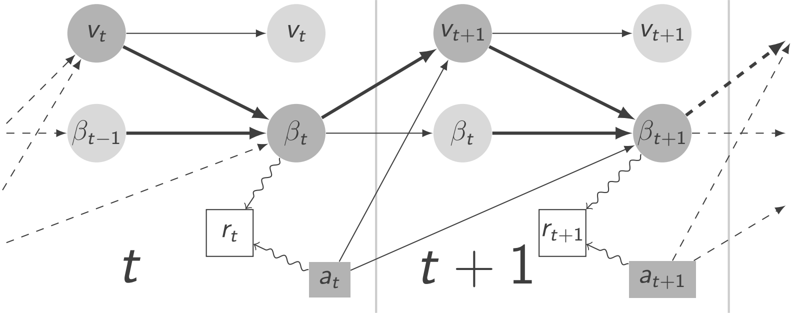
boolean variable "*flipflop*" f changes state at each time step
$\rightarrow$ defines 2 phases:

1. *observation generation*,
2. *belief update* (deterministic knowing the observation).

MDP variables:

$$\tilde{\mathbb{S}} =$$
$$\textbf{beliefs}: \ \beta = \beta_v^1 \times \ldots \times \beta_v^{m_v} \times \beta_h^1 \times \ldots \times \beta_h^{m_h} \times \beta_f^1 \times \ldots \times \beta_f^{m_f}$$
$$\times$$
$$\textbf{visible variables}: \ v = f \times s_v^1 \times \ldots \times s_v^{m_v} \times o_1 \times \ldots \times o_k .$$

resulting MDP in practice
final structured MDP

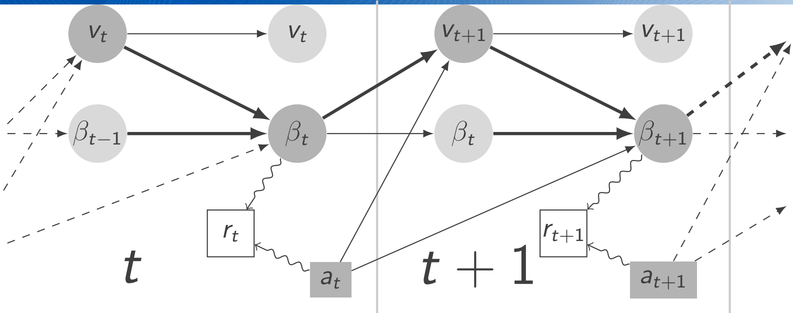


$\tilde{\mathbb{S}} =$

beliefs: $\beta = \beta_v^1 \times \ldots \times \beta_v^{m_v} \times \beta_h^1 \times \ldots \times \beta_h^{m_h} \times \beta_f^1 \times \ldots \times \beta_f^{m_f}$

$\times$

visible variables : $v = f \times s_v^1 \times \ldots \times s_v^{m_v} \times o_1 \times \ldots \times o_k$.

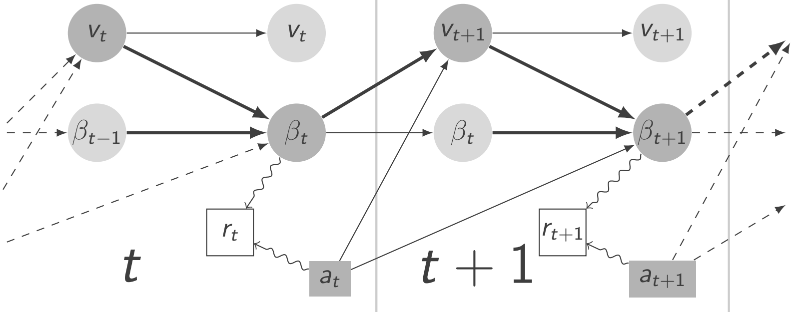


$\#$ **factorized model**'s variables: $\#\mathbb{O} + \#\mathbb{S}_v +$

$$+\sum_{i=1}^{\#\mathbb{O}_h} \left\lceil \log_2\left(\lambda^{2^{p_i}} - (\lambda-1)^{2^{p_i}}\right) \right\rceil + \#\mathbb{S}_f \cdot \left\lceil \log_2\left(2\lambda - 1\right) \right\rceil$$

$\ll$ $\#$ **initial hybrid model**'s variables:

$$\left\lceil \log_2\left(\lambda^{2^{\#\mathbb{S}}} - (\lambda-1)^{2^{\#\mathbb{S}}}\right) \right\rceil$$



\# **factorized model**'s variables:

$$\leqslant \#\mathbb{O} + \#\mathbb{S}_v + \sum_{i=1}^{\#\mathbb{O}_h} \log_2(\lambda) \cdot 2^{p_i} + \#\mathbb{S}_f \cdot (1 + \log_2(\lambda))$$

$\ll$ \# **initial hybrid model**'s variables:

$$\geqslant \log_2(\lambda) \cdot (2^{\#\mathbb{S}} - 1).$$

Plan

Conclusion

POMDP $\xrightarrow{\textbf{translation}}$ MDP with finite state space

- transition probabilities on the **possibilistic belief states**;

Conclusion

POMDP $\xrightarrow{\textbf{translation}}$ MDP with finite state space

- transition probabilities on the **possibilistic belief states**;
- **pessimistic** evaluation of the **rewards** (Choquet integral);

Conclusion

POMDP $\xrightarrow{\textbf{translation}}$ MDP with finite state space

- transition probabilities on the **possibilistic belief states**;
- **pessimistic** evaluation of the **rewards** (Choquet integral);
- factorized POMDP $\xrightarrow{\textbf{translation}}$ **factorized** MPD.

Conclusion

POMDP $\xrightarrow{\textbf{translation}}$ MDP with finite state space

- transition probabilities on the **possibilistic belief states**;
- **pessimistic** evaluation of the **rewards** (Choquet integral);
- factorized POMDP $\xrightarrow{\textbf{translation}}$ **factorized** MPD.

perspectives:

- IPPC problems (factorized POMDPs);

Conclusion

POMDP $\xrightarrow{\textbf{translation}}$ MDP with finite state space

- transition probabilities on the **possibilistic belief states**;
- **pessimistic** evaluation of the **rewards** (Choquet integral);
- factorized POMDP $\xrightarrow{\textbf{translation}}$ **factorized** MPD.

perspectives:

- IPPC problems (factorized POMDPs);
- tests of this approach:
  1. **simplification:** π distributions definition (π-normalization, pignistic transformation, most specific, . . .);
  2. **imprecision:** robust in practice?

Thank you!