Exploiting Imprecise Information Sources in Sequential Decision Making Problems under Uncertainty

N.Drougard

under **D.Dubois**, **J-L.Farges** and **F.Teichteil-Königsbuch** supervision

doctoral school: EDSYS **institution:** ISAE–SUPAERO

laboratory: ONERA–The French Aerospace Lab

retour sur innovation

Overview

Overview

Onera, DCSD

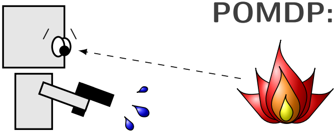
Control Engeenering, AI, Flight Dynamics, Cognitive Sciences

Onera, DCSD

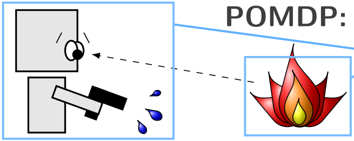
Control Engeenering, AI, Flight Dynamics, Cognitive Sciences

among many other works:

- autonomy and human factors
- decision making, planning
- experimental/industrial applications: UAVs, human-machine interaction, exploration robots

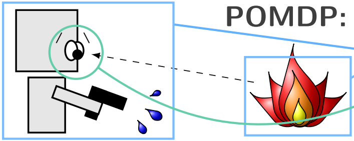


POMDP: model for autonomous robotics



POMDP: model for autonomous robotics

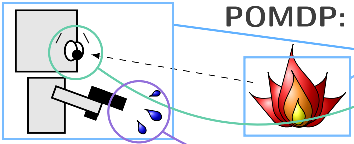
$s \in \mathcal{S}$: **system states**



POMDP: model for autonomous robotics

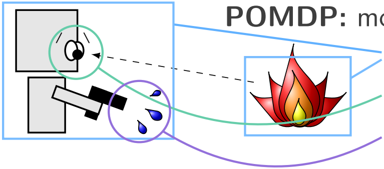
$s \in \mathcal{S}$: **system states**

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



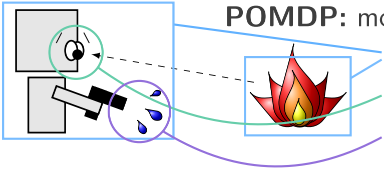
POMDP: model for autonomous robotics

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



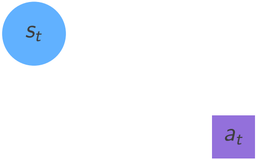
POMDP: model for autonomous robotics

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



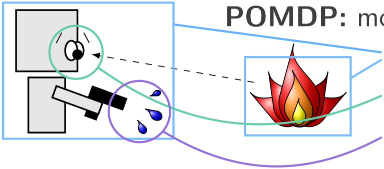
POMDP: model for autonomous robotics

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



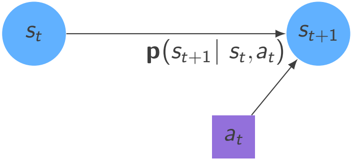
Context

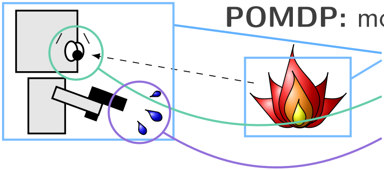
Partially Observable Markov Decision Processes (POMDPs)



POMDP: model for autonomous robotics

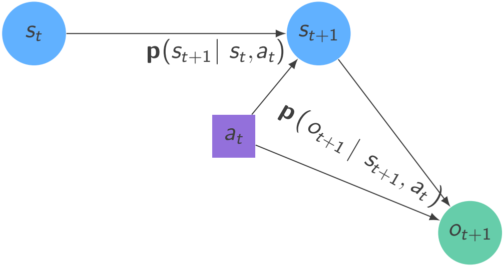
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**





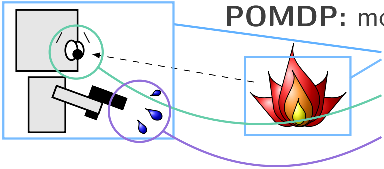
POMDP: model for autonomous robotics

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



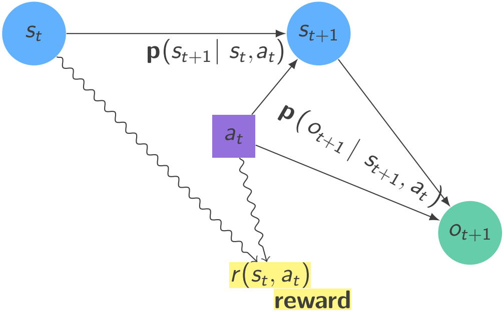
Context

Partially Observable Markov Decision Processes (POMDPs)

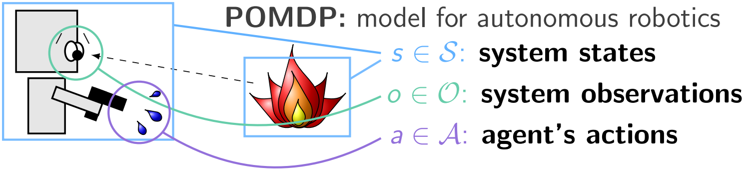


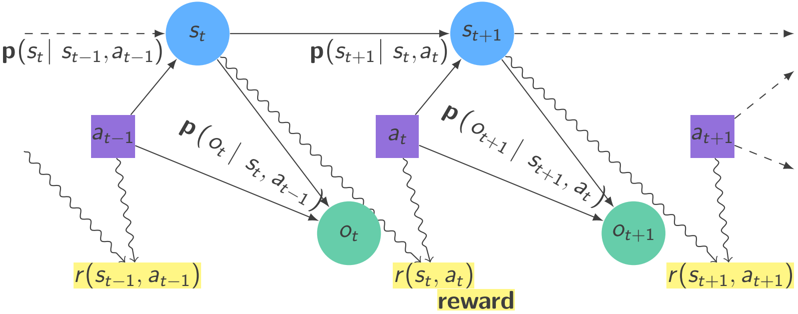
POMDP: model for autonomous robotics

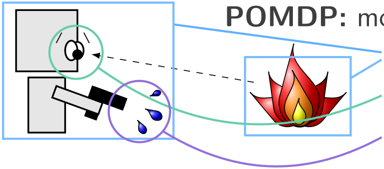
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**



POMDP: model for autonomous robotics

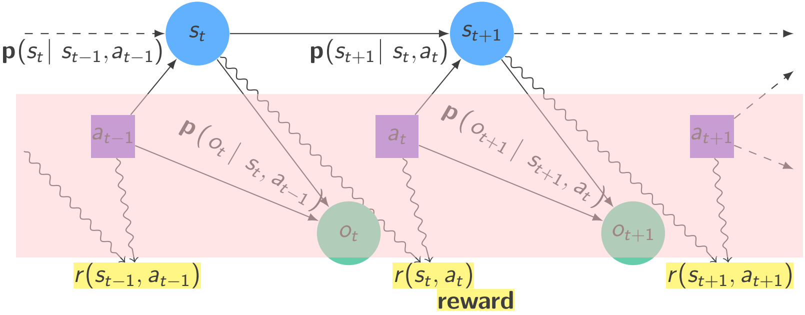




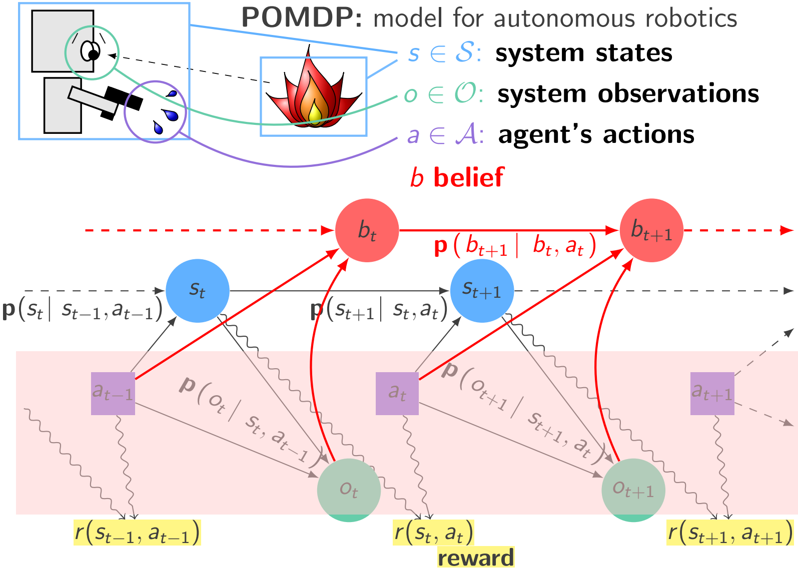


POMDP: model for autonomous robotics

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



POMDP: model for autonomous robotics



POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$ *(Smallwood et al. 1973)*

- **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: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$ *(Smallwood et al. 1973)*

- **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: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$ *(Smallwood et al. 1973)*

- **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 – a selected, o' received

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

POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$ *(Smallwood et al. 1973)*

- **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 – a selected, o' received

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

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** (*Papadimitriou et al. 1987*)
- probabilities are **imprecisely known** in practice
- **prior ignorance** management?

- **POMDP optimal strategy computation undecidable** in infinite horizon – *Madani et al. (AAAI-99)*

- **POMDP optimal strategy computation undecidable** in infinite horizon – *Madani et al. (AAAI-99)*

$\rightarrow$ optimality for "small" or "structured" POMDPs

$\rightarrow$ approximate computations

- **POMDP optimal strategy computation undecidable** in infinite horizon – *Madani et al. (AAAI-99)*

$\rightarrow$ optimality for "small" or "structured" POMDPs

$\rightarrow$ approximate computations

- **Imprecise model**, e.g. vision from statistical learning

- **POMDP optimal strategy computation undecidable** in infinite horizon – *Madani et al. (AAAI-99)*

→ optimality for "small" or "structured" POMDPs
→ approximate computations

- **Imprecise model**, e.g. vision from statistical learning

→ unknown environments: image variability of the datasets?

- **POMDP optimal strategy computation undecidable** in infinite horizon – *Madani et al. (AAAI-99)*

$\rightarrow$ optimality for "small" or "structured" POMDPs

$\rightarrow$ approximate computations

- **Imprecise model**, e.g. vision from statistical learning

$\rightarrow$ unknown environments: image variability of the datasets?

- **Lack of prior information** on the system state: initial belief state b_0

Context
practical issues: Complexity, Vision and Initial Belief

- **POMDP optimal strategy computation undecidable** in infinite horizon – *Madani et al. (AAAI-99)*

$\rightarrow$ optimality for "small" or "structured" POMDPs
$\rightarrow$ approximate computations

- **Imprecise model**, e.g. vision from statistical learning

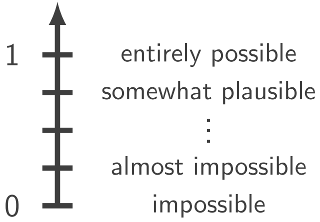
$\rightarrow$ unknown environments: image variability of the datasets?

- **Lack of prior information** on the system state: initial belief state b_0

$\rightarrow$ uniform probability distribution $\neq$ ignorance!

finite scale $\mathcal{L}$

usually $\{0, \frac{1}{k}, \frac{2}{k}, \dots, 1\}$



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)

finite scale $\mathcal{L}$

usually $\{0, \frac{1}{k}, \frac{2}{k}, \ldots, 1\}$

1	entirely possible
	somewhat plausible
	⋮
	almost impossible
0	impossible

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 / Possibility:

$+$	max
$\times$	min
$X \in \mathbb{R}$	$X \in \mathcal{L}$
$\mathbb{E}[X] = \sum_{x \in X} x \cdot \mathbf{p}(x)$	**optimistic:** $\mathbb{S}_{\Pi}[X] = \max_{x \in X} \min \{ x, \pi(x) \}$ **cautious:** $\mathbb{S}_{\mathcal{N}}[X] = \min_{x \in X} \max \{ x, 1 - \pi(x) \}$

Sabbadin (UAI-98) introduces

the qualitative possibilistic POMDP

π-**POMDP:** $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T^{\pi}, O^{\pi}, \rho \rangle$

Sabbadin (UAI-98) introduces

the qualitative possibilistic POMDP

π-**POMDP:** $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T^{\pi}, O^{\pi}, \rho \rangle$

- **transition** function $T^{\pi}(s, a, s') = \pi\left(s' \mid s, a\right) \in \mathcal{L}$
- **observation** function $O^{\pi}(s', a, o') = \pi\left(o' \mid s', a\right) \in \mathcal{L}$
- **preference** function $\rho : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{L}$

Sabbadin (UAI-98) introduces

the qualitative possibilistic POMDP

π-**POMDP:** $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T^{\pi}, O^{\pi}, \rho \rangle$

- **transition** function $T^{\pi}(s,a,s') = \pi\left(s' \mid s,a\right) \in \mathcal{L}$
- **observation** function $O^{\pi}(s',a,o') = \pi\left(o' \mid s',a\right) \in \mathcal{L}$
- **preference** function $\rho : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{L}$

- belief space trick: POMDP $\rightarrow$ MDP with **infinite** space
π-POMDP $\rightarrow$ π-MDP with **finite** space
- $\forall s \in \mathcal{S}$, $\pi(s) = 1 \Leftrightarrow$ total ignorance about s

A possibilistic belief state
finite belief space

$\Pi_{\mathcal{L}}^{\mathcal{S}} = \left\{ \text{ possibility distributions } \right\}$: $\#\Pi_{\mathcal{L}}^{\mathcal{S}} \sim \#\mathcal{L}^{\#\mathcal{S}} < +\infty$

$\rightarrow$ *i.e.* **finite belief space**

A possibilistic belief state

finite belief space

$\Pi_{\mathcal{L}}^{\mathcal{S}} = \left\{ \text{ possibility distributions } \right\}$: $\#\Pi_{\mathcal{L}}^{\mathcal{S}} \sim \#\mathcal{L}^{\#\mathcal{S}} < +\infty$

$\rightarrow$ *i.e.* **finite belief space**

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

A possibilistic belief state
finite belief space

$\Pi^{\mathcal{S}}_{\mathcal{L}} = \left\{ \text{possibility distributions} \right\}$: $\#\Pi^{\mathcal{S}}_{\mathcal{L}} \sim \#\mathcal{L}^{\#\mathcal{S}} < +\infty$

$\rightarrow$ *i.e.* **finite belief space**

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

possibilistic belief update – a selected, o' received

joint distribution on $\mathcal{S} \times \mathcal{O}$ from b^{π}_t: $\pi\left(o', s' \,\middle|\, b^{\pi}_t, a \right)$

ONERA THE FRENCH AEROSPACE LAB

A possibilistic belief state

finite belief space

$\Pi^{\mathcal{S}}_{\mathcal{L}} = \left\{ \text{ possibility distributions } \right\}$: $\#\Pi^{\mathcal{S}}_{\mathcal{L}} \sim \#\mathcal{L}^{\#\mathcal{S}} < +\infty$

$\rightarrow$ *i.e.* **finite belief space**

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

possibilistic belief update – a selected, o' received

joint distribution on $\mathcal{S} \times \mathcal{O}$ from b^{π}_t: $\pi \left(o', s' \mid b^{\pi}_t, a \right)$

$\rightarrow$ **next belief state:** $b^{\pi}_{t+1}(s') = \pi \left(o', s' \mid b^{\pi}_t, a \right)$
unless s' maximizes $\pi \left(o', s' \mid b^{\pi}_t, a \right)$, then $b^{\pi}_{t+1}(s') = 1$

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

$$\Pi_{\mathcal{L}}^{\mathcal{S}} = \left\{ \text{ possibility distributions } \right\}: \; \#\Pi_{\mathcal{L}}^{\mathcal{S}} \sim \#\mathcal{L}^{\#\mathcal{S}} < +\infty$$

$\rightarrow$ *i.e.* **finite belief space**

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

possibilistic belief update – a selected, o' received

joint distribution on $\mathcal{S} \times \mathcal{O}$ from b_t^{π}: $\pi\left(o', s' \,\middle|\, b_t^{\pi}, a \right)$

$\rightarrow$ **next belief state:** $b_{t+1}^{\pi}(s') = \pi\left(o', s' \,\middle|\, b_t^{\pi}, a \right)$
unless s' maximizes $\pi\left(o', s' \,\middle|\, b_t^{\pi}, a \right)$, then $b_{t+1}^{\pi}(s') = 1$

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

- the update **only depends on** o'**,** a **and** b_t^{π}

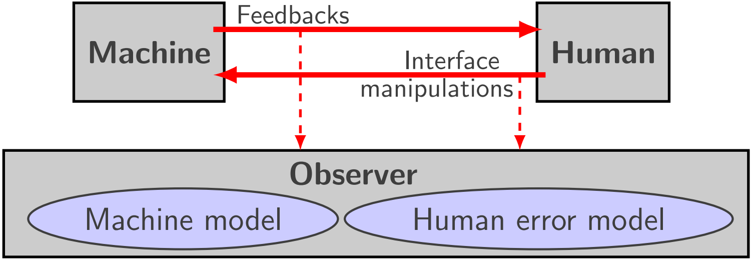
Qualitative Possibility Theory:
$\longrightarrow$ simplification, imprecision/prior ignorance modeling

Qualitative Possibility Theory:
$\rightarrow$ simplification, imprecision/prior ignorance modeling

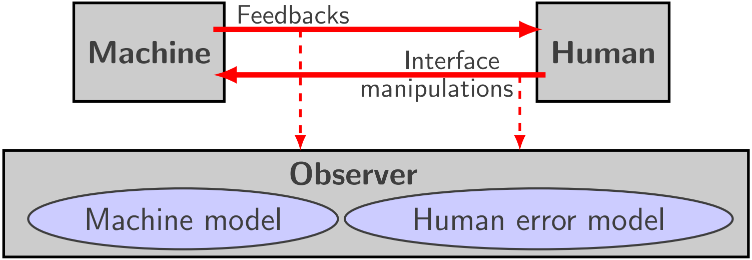
1. example of a qualitative possibilistic model
2. advancements and first use of the π-POMDP model
3. simplify computation: ADDs and factorization
4. probabilistic-possibilistic (hybrid) approach
5. conclusion/perspectives

Overview

Example: Human-Machine Interaction
joint work with Sergio Pizziol – Context

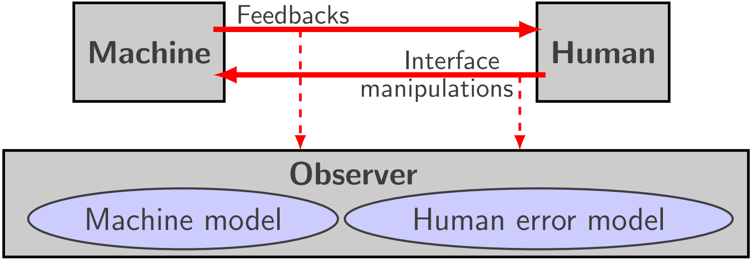


Example: Human-Machine Interaction
joint work with Sergio Pizziol – Context



Issue: incorrect human assessment of the machine state
$\rightarrow$ **accident**

Example: Human-Machine Interaction
joint work with Sergio Pizziol – Context



Issue: incorrect human assessment of the machine state
$\rightarrow$ **accident**

π-POMDP without actions: π-Hidden Markov Process

- **system space** $\mathcal{S}$: set of human assessments $\rightarrow$ **hidden**
- **observation space** $\mathcal{O}$: feedbacks/human manipulations

Machine with states A, B, C, ...

state $s_A \in \mathcal{S}$: "human thinks machine state is A"

Example: Human-Machine Interaction
Human error model from expert knowledge

Machine with states A, B, C, ...

state $s_A \in \mathcal{S}$: "human thinks machine state is A"

Machine state transition $A \rightarrow B$

- observation: **machine feedback** $o'_f \in \mathcal{O}$:

human usually aware of feedbacks $\rightarrow \pi\left(s'_B, o'_f \mid s_A\right) = 1$

but may lose a feedback $\rightarrow \pi\left(s'_A, o'_f \mid s_A\right) = \frac{2}{3}$

Example: Human-Machine Interaction

Human error model from expert knowledge

Machine with states A, B, C, ...

state $s_A \in \mathcal{S}$: "human thinks machine state is A"

Machine state transition $A \to B$

- observation: **machine feedback** $o'_f \in \mathcal{O}$:

human usually aware of feedbacks $\to \pi\left(s'_B, o'_f \mid s_A\right) = 1$
but may lose a feedback $\to \pi\left(s'_A, o'_f \mid s_A\right) = \frac{2}{3}$

- observation: **human manipulation** $o'_m \in \mathcal{O}$:

manipulation o'_m normal under $s_A \to \pi\left(s'_B, o'_m \mid s_A\right) = 1$
anormal manipulation $= \frac{1}{3}$

Machine with states A, B, C, ...

state $s_A \in \mathcal{S}$: "human thinks machine state is A"

Machine state transition $A \rightarrow B$

- observation: **machine feedback** $o'_f \in \mathcal{O}$:

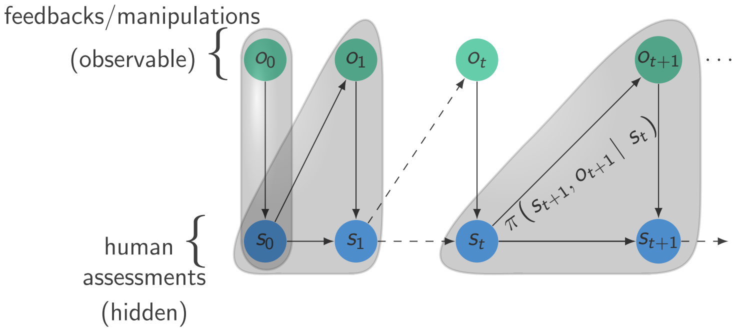
human usually aware of feedbacks $\rightarrow \pi\left(s'_B, o'_f \mid s_A\right) = 1$
but may lose a feedback $\rightarrow \pi\left(s'_A, o'_f \mid s_A\right) = \frac{2}{3}$

- observation: **human manipulation** $o'_m \in \mathcal{O}$:

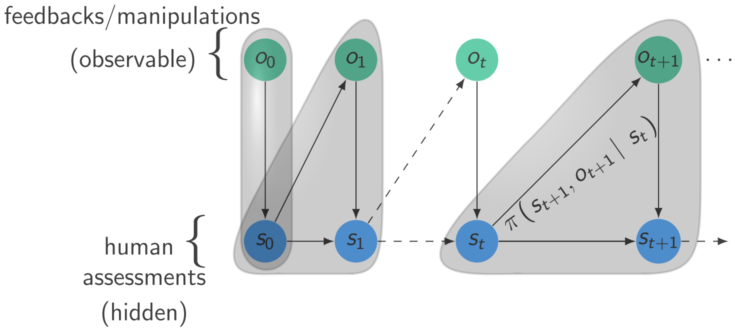
manipulation o'_m normal under $s_A \rightarrow \pi\left(s'_B, o'_m \mid s_A\right) = 1$
anormal manipulation $= \frac{1}{3}$

- impossible cases: possibility degree 0

Qualitative Possibilistic Hidden Markov Process:
diagnosis tool for Human-Machine interaction (with Sergio Pizziol)



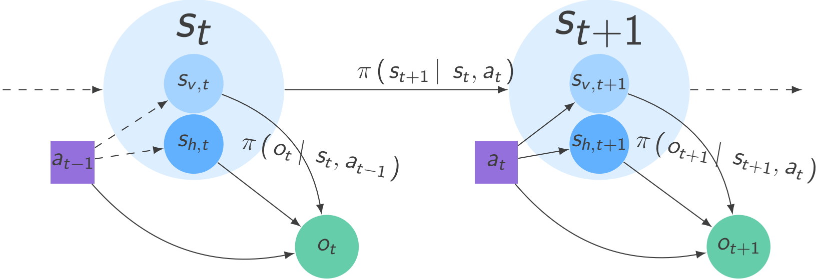
Qualitative Possibilistic Hidden Markov Process:
diagnosis tool for Human-Machine interaction (with Sergio Pizziol)



- **estimation** of the human assessment
  $\Leftrightarrow$ **possibilistic belief state**
- **detection** of human assessment errors + **diagnosis**
- validated with pilots on a flight simulator missions

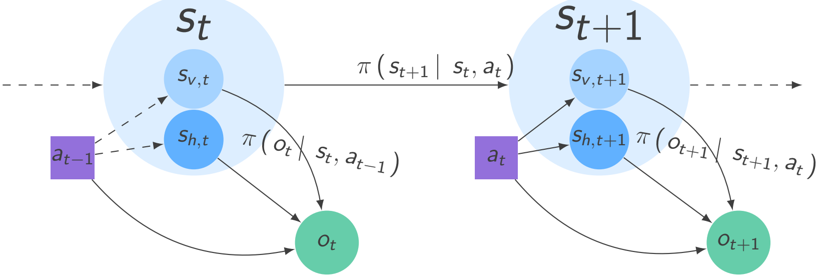
Overview

contribution (UAI-13):



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

contribution (UAI-13):



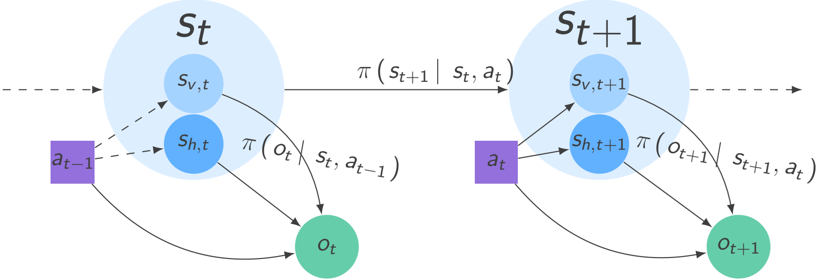
Mixed-Observability: system state $s \in \mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$
i.e. state s = visible component s_v & hidden component s_h

- belief states only over $\mathcal{S}_h$ (component s_v observed)

Mixed-Observability (MOMDP) – *Ong et al. (RSS-05)*

π-Mixed-Observable Markov Decision Process (π-MOMDP)

contribution (UAI-13):



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

- belief states only over $\mathcal{S}_h$ (component s_v observed)
- $\rightarrow$ π-POMDP: belief space $\Pi_{\mathcal{L}}^{\mathcal{S}}$ $\quad \#\Pi_{\mathcal{L}}^{\mathcal{S}} \sim \#\mathcal{L}^{\#\mathcal{S}}$

  $\rightarrow$ π-MOMDP: computations on $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$

  $\#\mathcal{X} \sim \#\mathcal{S}_v \cdot \#\mathcal{L}^{\#\mathcal{S}_h} \ll \#\Pi_{\mathcal{L}}^{\mathcal{S}}$

ONERA THE FRENCH AEROSPACE LAB

contribution (UAI-13): undeterminate Horizon

contribution (UAI-13): undeterminate Horizon

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

- assumption: $\exists$ artificial **"stay" action**
  as in classical planning/ γ counterpart
- criterion **non decreasing** with iterations

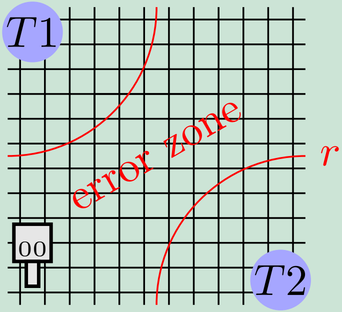
contribution (UAI-13): undeterminate Horizon

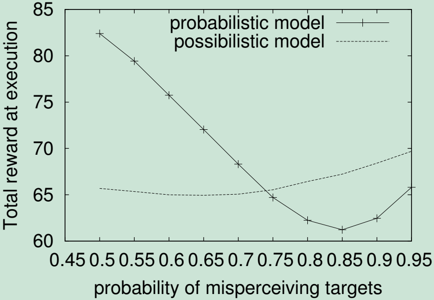
Dynamic Programming scheme: # iterations $< \#\mathcal{X}$

- assumption: $\exists$ artificial **"stay" action**
  as in classical planning/ γ counterpart
- criterion **non decreasing** with iterations
- action update for states increasing the criterion
- proof of optimality

- **goal:** reach the object $A = T1$ or $T2$
- noisy observations of the location of the object A

Recognition mission: robot on a grid, targets $T1$ & $T2$



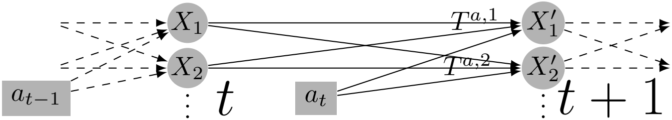


in reality, misperception probability in the error zone: $P_{bad} > \frac{1}{2}$

Overview

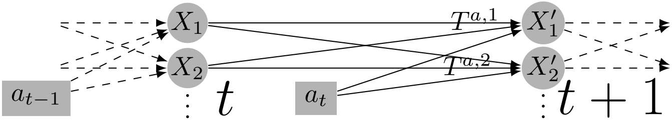
ONERA THE FRENCH AEROSPACE LAB

contribution (AAAI-14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$ = Boolean variables $(X_1, \ldots, X_n)$
$+$ independence assumptions $\Leftarrow$ graphical model



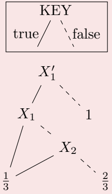
Factored π-MOMDP and computations with ADDs
qualitative possibilistic models to reduce complexity

contribution (AAAI-14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$ = Boolean variables $(X_1, \ldots, X_n)$
$+$ independence assumptions $\Leftarrow$ graphical model



- **factorization:** transition functions $T_i^a = \pi\left(X'_i \mid parents(X'_i), a\right)$ stored as **Algebraic Decision Diagrams (ADD)**

  probabilistic case:
  SPUDD, *Hoey et al., UAI-99*



example of ADD

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

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

PPUDD: Possibilistic Planning Using Decision Diagrams

- factorization $\Rightarrow$ each DP steps divided into n stages
  $\rightarrow$ smaller ADDs $\Rightarrow$ **faster computations**

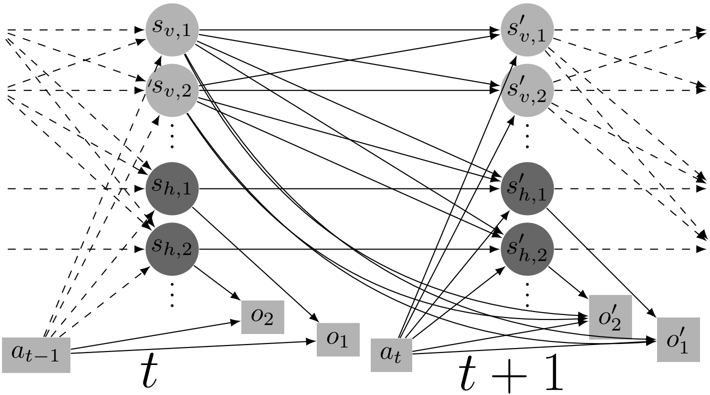
- probabilistic model: $+$ and $\times$ $\Rightarrow$ new values created
  $\Rightarrow$ number of ADDs leaves **potentially huge**
- possibilistic model: min and max $\Rightarrow$ values $\in \mathcal{L}$ finite
  $\Rightarrow$ number of leaves bounded, **ADDs smaller**.

PPUDD: Possibilistic Planning Using Decision Diagrams

- factorization $\Rightarrow$ each DP steps divided into n stages
  $\rightarrow$ smaller ADDs $\Rightarrow$ **faster computations**
- computations on trees: *CU Decision Diagram Package*.

contribution (AAAI-14):

independent sensors, hidden states, ... $\Rightarrow$ graphical model

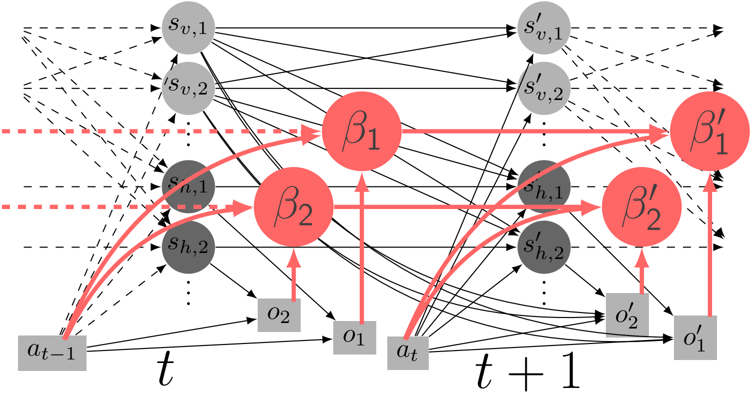


contribution (AAAI-14):

independent sensors, hidden states, ... $\Rightarrow$ graphical model

d-Separation $\Rightarrow$ $(s_v, \beta) = (s_{v,1}, \dots, s_{v,m}, \beta_1, \dots, \beta_l)$

$\beta_i \in \Pi_{\mathcal{L}}^{\mathcal{S}_{h,i}}$, belief over $s_{h,i}$



PPUDD vs SPUDD *Hoey et al.*

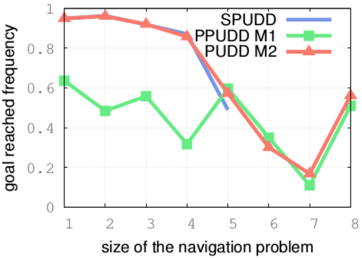
Navigation benchmark: reach a goal – spots with accident risk

M1 (resp. M2) optimistic (resp. cautious) criterion

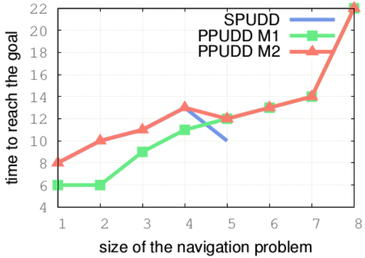
PPUDD vs SPUDD *Hoey et al.*
Navigation benchmark: reach a goal – spots with accident risk
M1 (resp. M2) optimistic (resp. cautious) criterion

Performances, function of the instance size

reached goal frequency



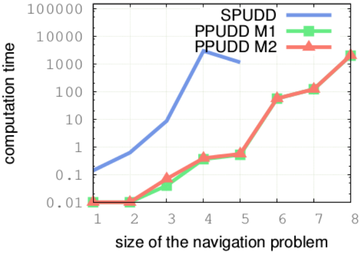
steps to reach the goal



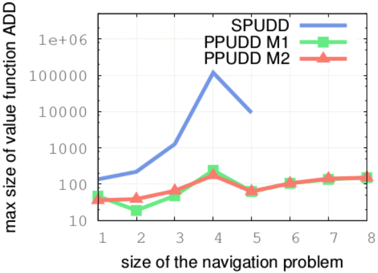
Simplify computations with π-MOMDPs

Experiments – perfect sensing: Navigation problem

computation time



max size of ADDs



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

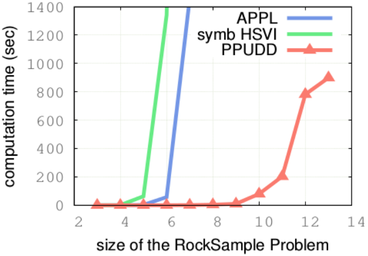
ONERA
THE FRENCH AEROSPACE LAB

PPUDD vs APPL *Kurniawati et al.*, solver MOMDP
symbolic HSVI *Sim et al.*, solver POMDP

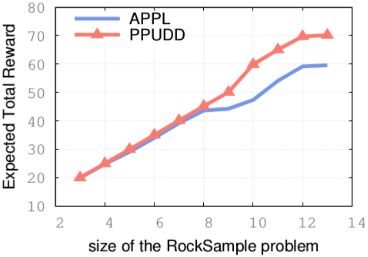
RockSample benchmark: recognize and sample “good” rocks

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 end



- **approximate model + exact resolution solver**
  $\rightarrow$ improvement of computation time and performances

IPPC 2014 – ADD-based approaches: PPUDD vs symbolic LRTDP

PPUDD + BDD mask over reachable states.

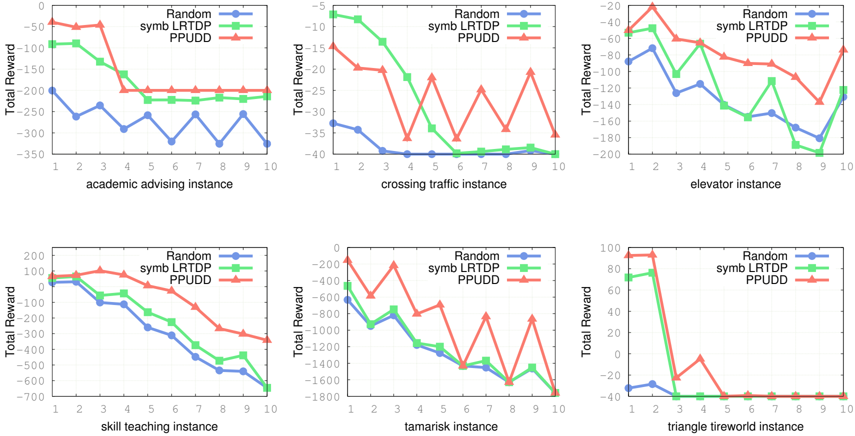


Figure : average of rewards over simulations

Overview

Qualitative possibilistic approach: benefits/drawbacks
towards a hybrid POMDP

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

Qualitative possibilistic approach: benefits/drawbacks

towards a hybrid POMDP

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

- ADD methods $\prec$ state space search methods
  $\rightarrow$ winners of IPPC 2014: $2\times$ better than PPUDD
- choice of the qualitative criterion (optimistic/pessimistic)
- preference $\rightarrow$ non additive degrees
  $\rightarrow$ same scale as possibility degrees (commensurability)
- coarse approximation of probabilistic model
  $\rightarrow$ no frequentist information

A hybrid model

a probabilistic POMDP with possibilistic belief states

hybrid approach

- agent knowledge = possibilistic belief states
- probabilistic dynamics & quantitative rewards

A hybrid model

a probabilistic POMDP with possibilistic belief states

hybrid approach

- agent knowledge $=$ possibilistic belief states
- probabilistic dynamics & quantitative rewards

Usefullness

$\rightarrow$ **heuristic** for solving POMDPs:

results in a standard (finite state space) MDP

$\rightarrow$ problem with **qualitative** & **quantitative** uncertainty

Transitions and rewards

belief-based transition and reward functions

- possibility distribution $\beta \rightarrow$ probability distribution $\overline{\beta}$ using poss-prob tranformations *(Dubois et al., FSS-92)*

Transition function on belief states

$$\Rightarrow \mathbf{p}\left(\beta' \middle| \overline{\beta}, a\right) = \sum_{\substack{o' \text{ t.q.} \\ update(\beta,a,o')=\beta'}} \mathbf{p}\left(o' \mid \overline{\beta}, a\right)$$

Transitions and rewards

belief-based transition and reward functions

- possibility distribution $\beta \rightarrow$ probability distribution $\overline{\beta}$ using poss-prob tranformations *(Dubois et al., FSS-92)*

Transition function on belief states

$$\Rightarrow \mathbf{p}\left(\beta' \middle| \overline{\beta}, a\right) = \sum_{\substack{o' \text{ t.q.} \\ update(\beta,a,o')=\beta'}} \mathbf{p}\left(o' \mid \overline{\beta}, a\right)$$

- reward cautious according to β

Pessimistic Choquet Integral

$$r(\beta, a) = \sum_{i=1}^{\#\mathcal{L}-1} (l_i - l_{i+1}) \cdot \min_{\substack{s \in \mathcal{S} \text{ s.t.} \\ \beta(s) \geqslant l_i}} r(s, a)$$

ONERA THE FRENCH AEROSPACE LAB

Resulting MDP

translation from hybrid POMDP to MDP – **contribution (SUM-15):**

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

Resulting MDP

translation from hybrid POMDP to MDP – **contribution (SUM-15):**

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** $\tilde{\mathcal{S}} = \Pi^{\mathcal{S}}_{\mathcal{L}}$,
  the set of the possibility distributions over $\mathcal{S}$;

Resulting MDP

translation from hybrid POMDP to MDP – **contribution (SUM-15):**

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** $\tilde{\mathcal{S}} = \Pi_{\mathcal{L}}^{\mathcal{S}}$,
  the set of the possibility distributions over $\mathcal{S}$;
- $\forall \beta, \beta'$ possibilistic belief states $\in \Pi_{\mathcal{L}}^{\mathcal{S}}$, $\forall a \in \mathcal{A}$,
  transitions $\tilde{T}\left(\beta, a, \beta'\right) = \mathbf{p}\left(\beta' \middle| \beta, a\right)$;

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** $\tilde{\mathcal{S}} = \Pi^{\mathcal{S}}_{\mathcal{L}}$,
  the set of the possibility distributions over $\mathcal{S}$;
- $\forall \beta, \beta'$ possibilistic belief states $\in \Pi^{\mathcal{S}}_{\mathcal{L}}$, $\forall a \in \mathcal{A}$,
  transitions $\tilde{T}\left(\beta, a, \beta'\right) = \mathbf{p}\left(\beta' \middle| \beta, a\right)$;
- **reward** $\tilde{r}(a, \beta) = \underline{Ch}\left(r(a, .)\right)$,

Resulting MDP

translation from hybrid POMDP to MDP – **contribution (SUM-15):**

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** $\tilde{\mathcal{S}} = \Pi_{\mathcal{L}}^{\mathcal{S}}$,
  the set of the possibility distributions over $\mathcal{S}$;
- $\forall \beta, \beta'$ possibilistic belief states $\in \Pi_{\mathcal{L}}^{\mathcal{S}}$, $\forall a \in \mathcal{A}$,
  transitions $\tilde{T}\left(\beta, a, \beta'\right) = \mathbf{p}\left(\beta' \middle| \beta, a\right)$;
- **reward** $\tilde{r}(a, \beta) = \underline{Ch}\Big(r(a, .)\Big)$,

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

General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

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

General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

variable: visible $s'_v \in \mathbb{S}_v$

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

inferred hidden $s'_h \in \mathbb{S}_h$

$$s'_h$$

fully hidden $s'_f \in \mathbb{S}_f$

$$s'_f$$

variable: visible $s'_v \in \mathbb{S}_v$

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

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$

---

inferred hidden $s'_h \in \mathbb{S}_h$

$$s'_h$$

---

fully hidden $s'_f \in \mathbb{S}_f$

$$s'_f$$

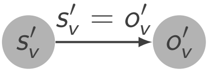
General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



---

inferred hidden $s'_h \in \mathbb{S}_h$

---

fully hidden $s'_f \in \mathbb{S}_f$

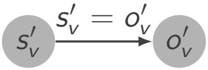
General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

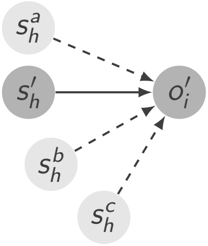
<u>variable:</u> visible $s'_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



inferred hidden $s'_h \in \mathbb{S}_h$

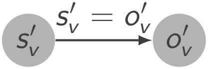


fully hidden $s'_f \in \mathbb{S}_f$

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

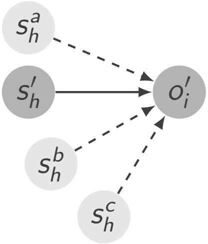
$\Leftrightarrow$ deterministic belief variable

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



inferred hidden $s'_h \in \mathbb{S}_h$

$$\beta_{t+1}\Big(parents(o'_i)\Big) = \beta_{t+1}(s_h, s_h^a, s_h^b, s_h^c)$$

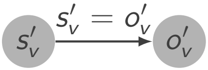


fully hidden $s'_f \in \mathbb{S}_f$

variable: visible $s'_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable

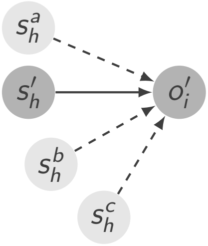
$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



inferred hidden $s'_h \in \mathbb{S}_h$

$$\beta_{t+1}\Big(parents(o'_i)\Big) = \beta_{t+1}(s_h, s_h^a, s_h^b, s_h^c)$$

$$\propto^{\pi} \pi\Big(o'_i, parents(o'_i) \Big| \beta_t, a\Big)$$



fully hidden $s'_f \in \mathbb{S}_f$

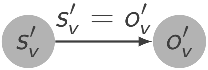
General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{ s'_v = o'_v \}}(s'_v)$$



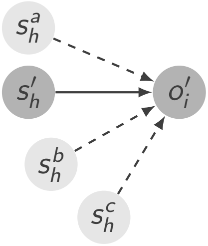
---

inferred hidden $s'_h \in \mathbb{S}_h$

$$\beta_{t+1}\Big(parents(o'_i)\Big) = \beta_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^\pi \pi\Big(o'_i, parents(o'_i) \Big| \beta_t, a\Big)$$

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



---

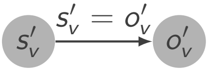
fully hidden $s'_f \in \mathbb{S}_f$

General variable classification **contribution (SUM-15):**

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

$\Leftrightarrow$ deterministic belief variable

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$

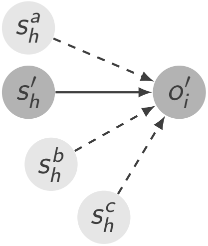


inferred hidden $s'_h \in \mathbb{S}_h$

$$\beta_{t+1}\Big(parents(o'_i)\Big) = \beta_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^{\pi} \pi\Big(o'_i, parents(o'_i) \Big| \beta_t, a\Big)$$

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

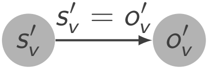


fully hidden $s'_f \in \mathbb{S}_f$

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

$\Leftrightarrow$ deterministic belief variable

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{ s'_v = o'_v \}}(s'_v)$$

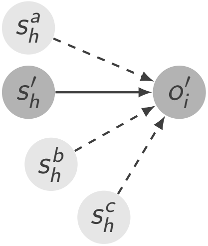


inferred hidden $s'_h \in \mathbb{S}_h$

$$\beta_{t+1}\Big(parents(o'_i)\Big) = \beta_{t+1}(s_h, s_h^a, s_h^b, s_h^c)$$

$$\propto^{\pi} \pi\Big(o'_i, parents(o'_i) \Big| \beta_t, a\Big)$$

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



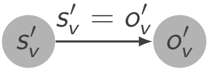
fully hidden $s'_f \in \mathbb{S}_f$

$$\beta_{t+1}(s'_f) = \pi\,(\, s'_f \mid \beta_t, a\,)$$

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

$\Leftrightarrow$ deterministic belief variable

$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



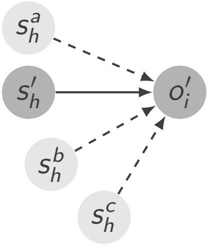
---

inferred hidden $s'_h \in \mathbb{S}_h$

$$\beta_{t+1}\Big(parents(o'_i)\Big) = \beta_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

$$\propto^{\pi} \pi\Big(o'_i, parents(o'_i) \Big| \beta_t, a\Big)$$

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



---

fully hidden $s'_f \in \mathbb{S}_f$

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

$$\beta_{t+1}(s'_f) = \pi\left(s'_f \mid \beta_t, a\right)$$

bound over the global belief state

$$\beta_{t+1}(s'_1, \ldots, s'_n) = \pi\left(s'_1, \ldots, s'_n \mid a_0, o_1, \ldots, a_t, o_{t+1}\right)$$

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

Possibilistic belief variables

global belief state

bound over the global belief state

$$\beta_{t+1}(s'_1, \ldots, s'_n) = \pi\left(s'_1, \ldots, s'_n \mid a_0, o_1, \ldots, a_t, o_{t+1}\right)$$

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

- min of marginals = a **less informative** belief state
- computed using **marginal belief states**
  $\rightarrow$ factorization & smaller state space

Overview

- **modeling efforts**: → human-machine interaction

- **modeling efforts**: $\rightarrow$ human-machine interaction
- **advancements**: $\rightarrow$ mixed-observability modeling
  $\rightarrow$ undeterminate horizon $+$ optimality proof

- **modeling efforts**: $\rightarrow$ human-machine interaction
- **advancements**: $\rightarrow$ mixed-observability modeling
  $\rightarrow$ undeterminate horizon $+$ optimality proof
- **simplifying computations**: factorization work
  & PPUDD algorithm

- **modeling efforts**: $\rightarrow$ human-machine interaction
- **advancements**: $\rightarrow$ mixed-observability modeling
  $\rightarrow$ undeterminate horizon + optimality proof
- **simplifying computations**: factorization work
  & PPUDD algorithm
- **experimentations**: real problems
  $\rightarrow$ robust recognition mission with possibilistic beliefs
  $\rightarrow$ validation of the computation time reduction
  $\rightarrow$ IPPC 2014

- **modeling efforts**: $\rightarrow$ human-machine interaction
- **advancements**: $\rightarrow$ mixed-observability modeling
  $\rightarrow$ undeterminate horizon + optimality proof
- **simplifying computations**: factorization work
  & PPUDD algorithm
- **experimentations**: real problems
  $\rightarrow$ robust recognition mission with possibilistic beliefs
  $\rightarrow$ validation of the computation time reduction
  $\rightarrow$ IPPC 2014
- **hybrid POMDP** $\xrightarrow{\text{translation}}$ MDP
  $\rightarrow$ probabilities on possibilistic belief states
  pessimistic rewards (Choquet integral)
  $\rightarrow$ factored POMDP $\xrightarrow{\text{translation}}$ factored MPD

- *Qualitative Possibilistic Mixed-Observable MDPs*, UAI-2013
- *Structured Possibilistic Planning Using Decision Diagrams*, AAAI-2014
- *Planning in Partially Observable Domains with Fuzzy Epistemic States and Probabilistic Dynamics*, SUM-2015
- *Processus Décisionnels de Markov Possibilistes à Observabilité Mixte*, Revue d'Intelligence Artificielle (RIA)
- *A Possibilistic Estimation of Human Attentional Errors*, submitted to IEEE-TFS

Conclusion
perspectives

- refined criteria *(Weng UAI-05, Dubois et al. EJOR-05)*
  $\Rightarrow$ finer π-POMDP
- link with statistical learning in practice

- refined criteria *(Weng UAI-05, Dubois et al. EJOR-05)*
  $\Rightarrow$ finer π-POMDP
- link with statistical learning in practice

quantitative information may be available: hybrid work

- IPPC problems (factored POMDPs);
- tests of this approach:
  1. **simplification:** π distributions definition?
  2. **imprecision:** robust in practice?

Thank you!