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

Context

Autonomous robotics

Onera, Flight Dynamics & System control

Control Engineering, Artificial intelligence, Cognitive Sciences

Context
Autonomous robotics

Onera, Flight Dynamics & System control

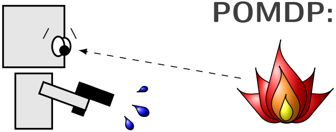
Control Engineering, Artificial intelligence, Cognitive Sciences

among many other works:

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

Context

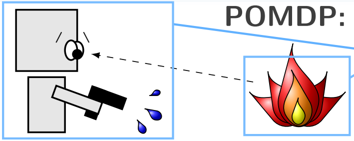
Partially Observable Markov Decision Processes (POMDPs)



POMDP: model for autonomous robotics

Context

Partially Observable Markov Decision Processes (POMDPs)

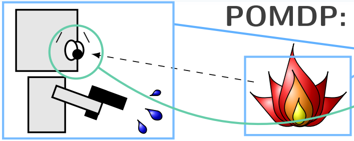


POMDP: model for autonomous robotics

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

Context

Partially Observable Markov Decision Processes (POMDPs)

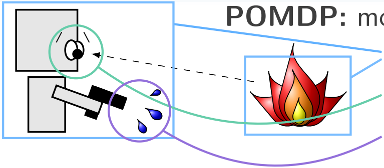


POMDP: model for autonomous robotics

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

Context

Partially Observable Markov Decision Processes (POMDPs)

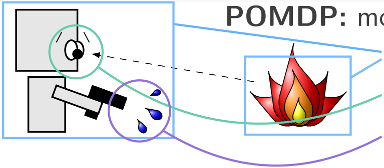


POMDP: model for autonomous robotics

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

Context

Partially Observable Markov Decision Processes (POMDPs)

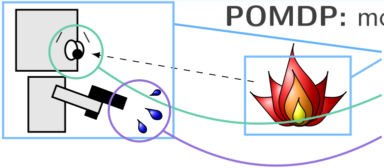


POMDP: model for autonomous robotics

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

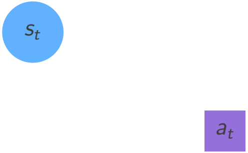
Context

Partially Observable Markov Decision Processes (POMDPs)



POMDP: model for autonomous robotics

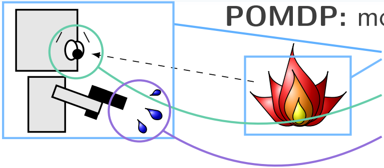
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**



Context

Partially Observable Markov Decision Processes (POMDPs)

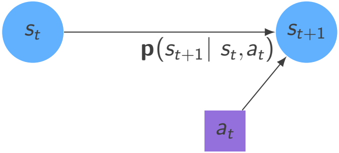
POMDP: model for autonomous robotics



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

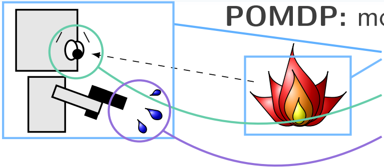
$o \in \mathcal{O}$: **system observations**

$a \in \mathcal{A}$: **agent's actions**



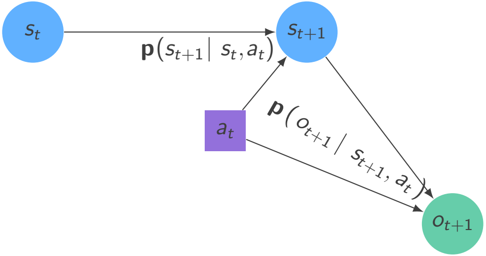
Context

Partially Observable Markov Decision Processes (POMDPs)



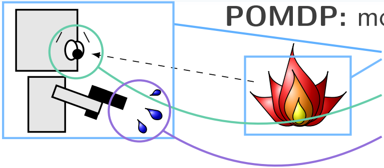
POMDP: model for autonomous robotics

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



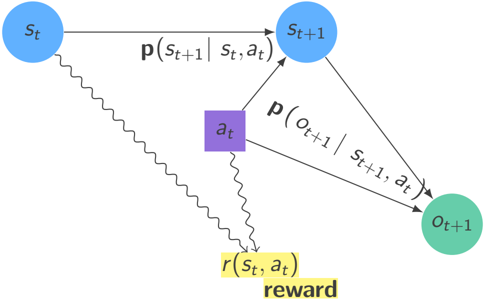
Context

Partially Observable Markov Decision Processes (POMDPs)



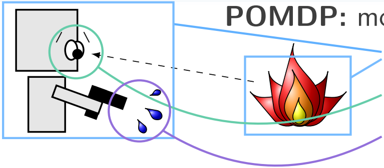
POMDP: model for autonomous robotics

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



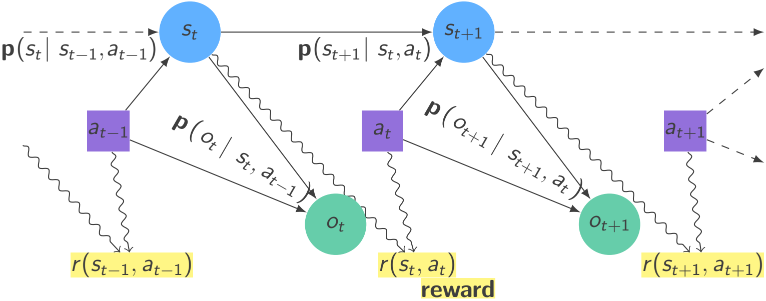
Context

Partially Observable Markov Decision Processes (POMDPs)



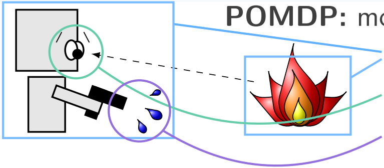
POMDP: model for autonomous robotics

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



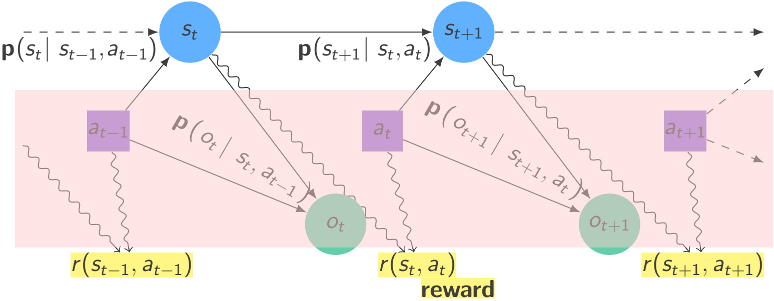
Context

Partially Observable Markov Decision Processes (POMDPs)



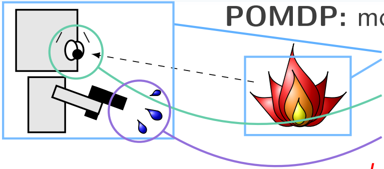
POMDP: model for autonomous robotics

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



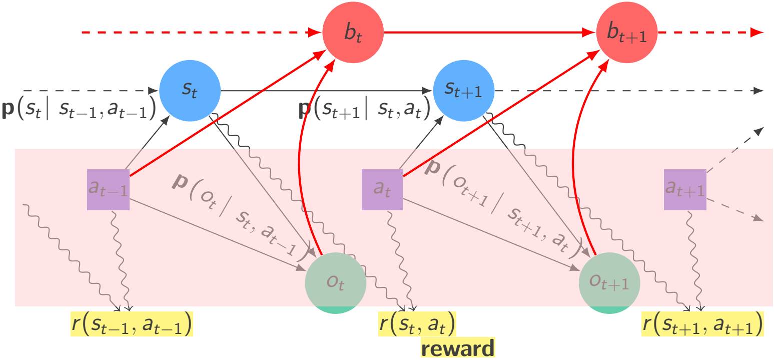
Context

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**
- b **belief**: **probability distribution**



Context

belief state, strategy, criterion

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 $O(s',a,o') = \mathbf{p}\left(o' \mid s', a \right)$
- **reward** function $r(s,a) \in \mathbb{R}$

Context

belief state, strategy, criterion

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 $O(s',a,o') = \mathbf{p}\left(o' \mid s', a\right)$
- **reward** function $r(s,a) \in \mathbb{R}$

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

Context

belief state, strategy, criterion

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 $O(s',a,o') = \mathbf{p}\left(o' \mid s', a \right)$
- **reward** function $r(s,a) \in \mathbb{R}$

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

T, O

Context

belief state, strategy, criterion

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 $O(s', a, o') = \mathbf{p}\left(o' \mid s', a \right)$
- **reward** function $r(s, a) \in \mathbb{R}$

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

$$b_t,\ T, O \longrightarrow \mathbf{p}\left(s', o' \mid b_t, a_t \right)$$

Context

belief state, strategy, criterion

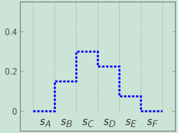
POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$ (*Smallwood et al. 1973*)

- **transition** function $T(s,a,s') = \mathbf{p}\left(s' \mid s,a\right)$
- **observation** function $O(s',a,o') = \mathbf{p}\left(o' \mid s',a\right)$
- **reward** function $r(s,a) \in \mathbb{R}$

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, T,O $\longrightarrow$ $\mathbf{p}\left(s', o' \mid b_t, a_t\right)$ $\xrightarrow[o_{t+1} \text{ fixed}]{\text{received observation}}$



Context

belief state, strategy, criterion

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

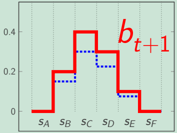
- **transition** function $T(s, a, s') = \mathbf{p}\left(s' \mid s, a \right)$
- **observation** function $O(s', a, o') = \mathbf{p}\left(o' \mid s', a \right)$
- **reward** function $r(s, a) \in \mathbb{R}$

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

$$b_t, \; T, O \longrightarrow \mathbf{p}\left(s', o' \mid b_t, a_t \right) \xrightarrow[o_{t+1} \text{ fixed}]{\text{received observation}}$$

probabilistic normalization



Context

belief state, strategy, criterion

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

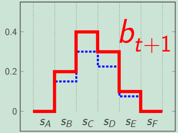
- **transition** function $T(s,a,s') = \mathbf{p}\left(s' \mid s, a\right)$
- **observation** function $O(s',a,o') = \mathbf{p}\left(o' \mid s', a\right)$
- **reward** function $r(s,a) \in \mathbb{R}$

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

probabilistic belief update

b_t, T,O → $\mathbf{p}\left(s', o' \mid b_t, a_t\right)$ —received observation, o_{t+1} fixed→

probabilistic normalization



strategy $d_t : b_t \mapsto a_t \in \mathcal{A}$

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

Flaws of the POMDP model

POMDPs in practice

- optimal strategy computation **PSPACE-hard**
  (*Papadimitriou et al., 1987*)

- probabilities are **imprecisely known** in practice

- **prior ignorance** semantic/management?

Context

practical issues: Complexity, Vision and Initial Belief

- **POMDP optimal strategy computation undecidable** in infinite horizon (*Madani et al. 1999*)

Context

practical issues: Complexity, Vision and Initial Belief

- **POMDP optimal strategy computation undecidable** in infinite horizon (*Madani et al. 1999*)

$\rightarrow$ optimality for “small” or “structured” POMDPs

$\rightarrow$ approximate computations

Context

practical issues: Complexity, Vision and Initial Belief

- **POMDP optimal strategy computation undecidable** in infinite horizon (*Madani et al. 1999*)

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

$\rightarrow$ approximate computations

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

Context

practical issues: Complexity, Vision and Initial Belief

- **POMDP optimal strategy computation undecidable** in infinite horizon (*Madani et al. 1999*)

→ optimality for “small” or “structured” POMDPs
→ approximate computations

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

→ unknown environments: image variability of the datasets?

Context

practical issues: Complexity, Vision and Initial Belief

- **POMDP optimal strategy computation undecidable** in infinite horizon (*Madani et al. 1999*)

→ optimality for "small" or "structured" POMDPs

→ approximate computations

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

→ 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. 1999*)

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

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

→ unknown environments: image variability of the datasets?

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

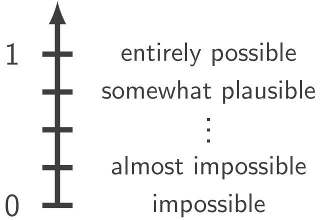
→ uniform probability distribution $\neq$ ignorance!

Qualitative Possibility Theory

presentation – (max,min) "tropical" algebra

finite scale $\mathcal{L}$

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



events $e \subset \Omega$ (universe)

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

$\neq$

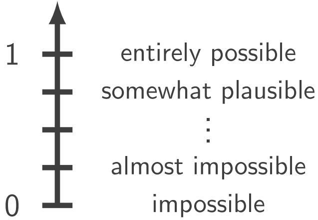
quantified with **frequencies** $\mathbf{p}(e) \in [0, 1]$ (probabilities)

Qualitative Possibility Theory

presentation – (max,min) "tropical" algebra

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)

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

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

Qualitative Possibility Theory

Criteria from Sugeno integral

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:** $\mathbb{S}_{\Pi}[X] = \max_{x \in X} \min \{ x, \pi(x) \}$ **pessimistic:** $\mathbb{S}_{\mathcal{N}}[X] = \min_{x \in X} \max \{ x, 1 - \pi(x) \}$

Qualitative Possibility Theory

qualitative possibilistic POMDP (π-POMDP)

Sabbadin (UAI-98) introduces

the qualitative possibilistic POMDP

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

Qualitative Possibility Theory

qualitative possibilistic POMDP (π-POMDP)

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 $\rho : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{L}$

Qualitative Possibility Theory

qualitative possibilistic POMDP (π-POMDP)

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
- problem becomes **decidable**
- $\forall s \in \mathcal{S}$, $\pi(s) = 1 \Leftrightarrow$ 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**

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

A possibilistic belief state

finite belief space

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

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

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

possibilistic belief update

β_t

T^π, O^π

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**

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

possibilistic belief update

$$\left.\begin{array}{r} \beta_t \longrightarrow \\ T^{\pi}, O^{\pi} \longrightarrow \end{array}\right. \pi \left(s', o' \mid \beta_t, a_t \right)$$

A possibilistic belief state
finite belief space

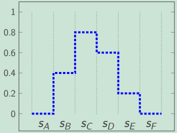
$$\Pi_{\mathcal{L}}^{\mathcal{S}} = \Big\{ \text{ possibility distributions } \Big\} \colon \#\Pi_{\mathcal{L}}^{\mathcal{S}} \sim \#\mathcal{L}^{\#\mathcal{S}} < +\infty$$

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

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

possibilistic belief update

β_t, T^π, O^π $\longrightarrow$ $\pi\left(s', o' \,\middle|\, \beta_t, a_t\right)$ $\xrightarrow[o_{t+1} \text{ fixed}]{\text{received observation}}$



A possibilistic belief state

finite belief space

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

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

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

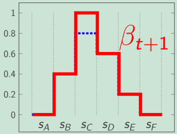
possibilistic belief update

possibilistic normalization

β_t

T^π, O^π

$\pi\left(s', o' \,\middle|\, \beta_t, a_t\right) \xrightarrow[o_{t+1} \text{ fixed}]{\text{received observation}}$



A possibilistic belief state

finite belief space

$$\Pi_{\mathcal{L}}^{\mathcal{S}} = \Big\{ \text{possibility distributions} \Big\} \colon \#\Pi_{\mathcal{L}}^{\mathcal{S}} \sim \#\mathcal{L}^{\#\mathcal{S}} < +\infty$$

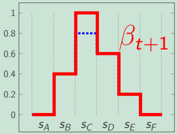
$\rightarrow$ *i.e.* **finite belief space**

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

possibilistic belief update

β_t, T^π, O^π $\rightarrow$ $\pi\left(s', o' \mid \beta_t, a_t\right)$ $\xrightarrow[o_{t+1}\ \text{fixed}]{\text{received observation}}$

possibilistic normalization



- **Markovian update:** only depends on o_{t+1}, a_t and b_t^π

Overview

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

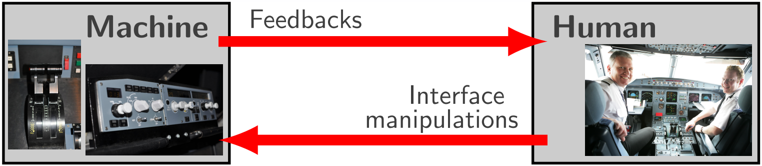
Overview

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

1. introductory example: qualitative **possibilistic modeling**
2. **advancements** in π-POMDP:
   mixed-observability & indefinite horizon
3. **simplifying computations**:
   ADD-based solver & factorization
4. **probabilistic-possibilistic** (*hybrid*) approach

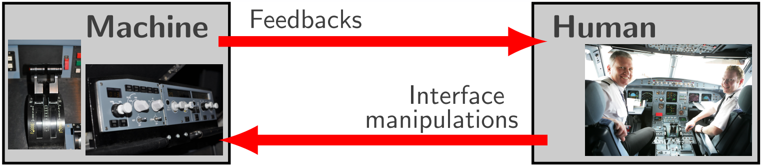
Example: Human-Machine Interaction (HMI)

joint work with **Sergio Pizziol** – Context



Example: Human-Machine Interaction (HMI)

joint work with **Sergio Pizziol** – Context

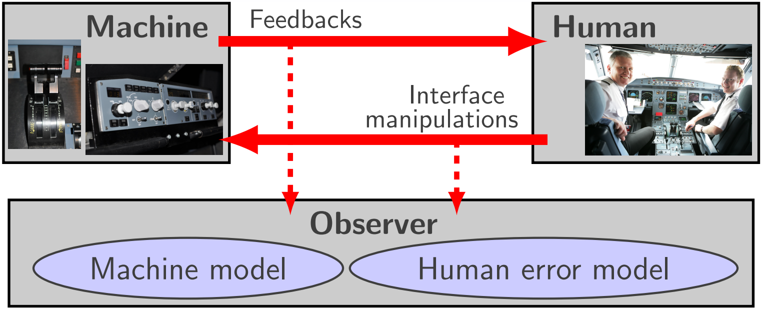


Issue: incorrect human assessment of the machine state

$\rightarrow$ **accident risk**

Example: Human-Machine Interaction (HMI)

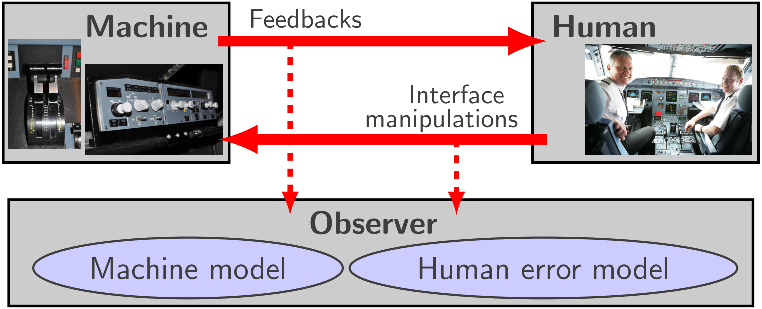
joint work with **Sergio Pizziol** – Context



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

Example: Human-Machine Interaction (HMI)

joint work with **Sergio Pizziol** – Context



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

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

Example: Human-Machine Interaction (HMI)

Human error model from expert knowledge

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

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

Example: Human-Machine Interaction (HMI)

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** $o'_f \in \mathcal{O}$:

$$\text{"human usually aware of feedbacks"} \to \pi\left(s'_B, o'_f \mid s_A\right) = 1$$

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

Example: Human-Machine Interaction (HMI)

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** $o'_f \in \mathcal{O}$:

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

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

$$\text{“manipulation } o'_m \text{ is normal under } s_A\text{”} \rightarrow \pi\left(s'_B, o'_m \mid s_A\right) = 1$$
$$\text{“is abnormal”} \rightarrow \quad = \frac{1}{3}$$

Example: Human-Machine Interaction (HMI)

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" $\rightarrow \pi \left(s'_A, o'_f \mid s_A \right) = \frac{2}{3}$

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

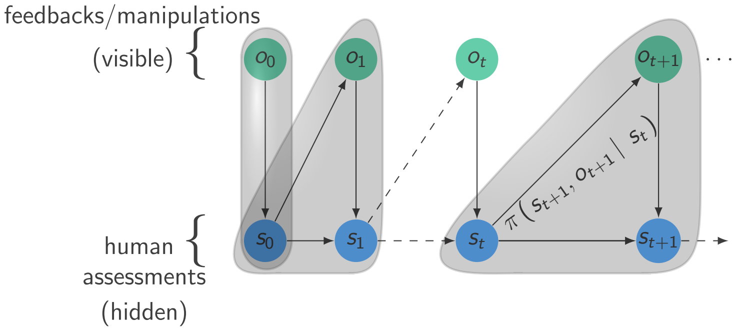
"manipulation o'_m is normal under s_A" $\rightarrow \pi \left(s'_B, o'_m \mid s_A \right) = 1$

"is abnormal" $\rightarrow \quad = \frac{1}{3}$

- impossible cases: possibility degree 0

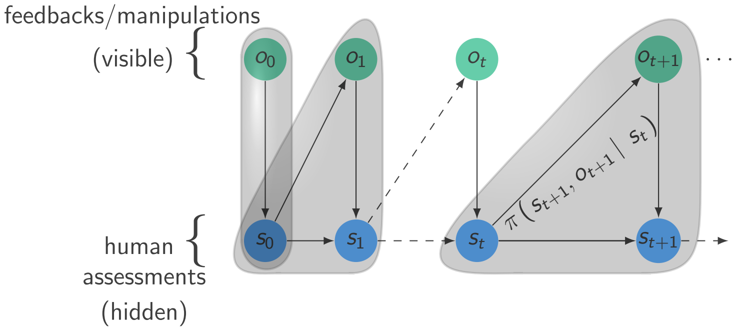
Qualitative Possibilistic Hidden Markov Process:

π-HMP, detection & diagnosis tool for HMI (with **Sergio Pizziol**)



Qualitative Possibilistic Hidden Markov Process:

π-HMP, detection & diagnosis tool for HMI (with **Sergio Pizziol**)



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

Applicability of the π-POMDPs

three advancements

- lack of proof of optimality in indefinite horizon settings
- criterion/algorithm/proof

- curse of dimensionality:
  $\rightarrow$ belief space size of a π-POMDP: exponential in $\#\mathcal{S}$
- in practice, part of $s \in \mathcal{S}$ is visible
  $\Rightarrow$ complexity reduction

- lack of possibilistic strategy evaluation
- demonstration of usefulness
  when probabilities are imprecise

Applicability of the π-POMDPs

three advancements

- lack of proof of optimality in indefinite horizon settings
- criterion/algorithm/proof

- curse of dimensionality:
  $\rightarrow$ belief space size of a π-POMDP: exponential in $\#\mathcal{S}$
- in practice, part of $s \in \mathcal{S}$ is visible
  $\Rightarrow$ complexity reduction

- lack of possibilistic strategy evaluation
- demonstration of usefulness
  when probabilities are imprecise

Applicability of the π-POMDPs

three advancements

- lack of proof of optimality in indefinite horizon settings
- criterion/algorithm/proof

- curse of dimensionality:
  $\rightarrow$ belief space size of a π-POMDP: exponential in $\#\mathcal{S}$
- in practice, part of $s \in \mathcal{S}$ is visible
  $\Rightarrow$ complexity reduction

- lack of possibilistic strategy evaluation
- demonstration of usefulness
  when probabilities are imprecise

Applicability of the π-POMDPs

three advancements

- lack of proof of optimality in indefinite horizon settings
- criterion/algorithm/proof

- curse of dimensionality:
  $\rightarrow$ belief space size of a π-POMDP: exponential in $\#\mathcal{S}$
- in practice, part of $s \in \mathcal{S}$ is visible
  $\Rightarrow$ complexity reduction

- lack of possibilistic strategy evaluation
- demonstration of usefulness
  when probabilities are imprecise

Indefinite Horizon, Mixed-Observability, Simulations

contribution UAI 2013

Indefinite Horizon

criterion, DP scheme, optimal strategy

indefinite horizon criterion $\Psi : \mathcal{S} \to \mathcal{L}$ terminal pref. func.

$$\forall s \in \mathcal{S}, \text{ maximizing } \mathbb{S}_{\Pi}\left[\Psi(S_{\#\delta}) \middle| S_0 = s \right]$$

with respect to the strategy $\delta : (t, s) \mapsto a_t \in \mathcal{A}$.

Indefinite Horizon

criterion, DP scheme, optimal strategy

indefinite horizon criterion $\Psi : \mathcal{S} \rightarrow \mathcal{L}$ terminal pref. func.

$\forall s \in \mathcal{S}$, maximizing $\mathbb{S}_{\Pi} \left[\Psi(S_{\#\delta}) \middle| S_0 = s \right]$

$$= \max_{(s_1,\dots,s_{\#\delta})} \min \left\{ \min_{t=0}^{\#\delta-1} \pi\left(s_{t+1} \middle| s_t, \delta_t(s_t) \right), \Psi(s_{\#\delta}) \right\}$$

with respect to the strategy $\delta : (t,s) \mapsto a_t \in \mathcal{A}$.

Indefinite Horizon

criterion, DP scheme, optimal strategy

indefinite horizon criterion $\Psi : \mathcal{S} \to \mathcal{L}$ terminal pref. func.

$$\forall s \in \mathcal{S}, \text{ maximizing } \mathbb{S}_\Pi \left[\Psi(S_{\#\delta}) \middle| S_0 = s \right]$$

$$= \max_{(s_1,\ldots,s_{\#\delta})} \min \left\{ \min_{t=0}^{\#\delta-1} \pi\left(s_{t+1} \middle| s_t, \delta_t(s_t) \right), \Psi(s_{\#\delta}) \right\}$$

with respect to the strategy $\delta : (t, s) \mapsto a_t \in \mathcal{A}$.

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

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

Indefinite Horizon

criterion, DP scheme, optimal strategy

indefinite horizon criterion $\Psi : \mathcal{S} \to \mathcal{L}$ terminal pref. func.

$\forall s \in \mathcal{S}$, maximizing $\mathbb{S}_{\Pi}\left[\Psi(S_{\#\delta}) \middle| S_0 = s \right]$

$$= \max_{(s_1,\dots,s_{\#\delta})} \min \left\{ \min_{t=0}^{\#\delta-1} \pi\left(s_{t+1} \middle| s_t, \delta_t(s_t) \right), \Psi(s_{\#\delta}) \right\}$$

with respect to the strategy $\delta : (t,s) \mapsto a_t \in \mathcal{A}$.

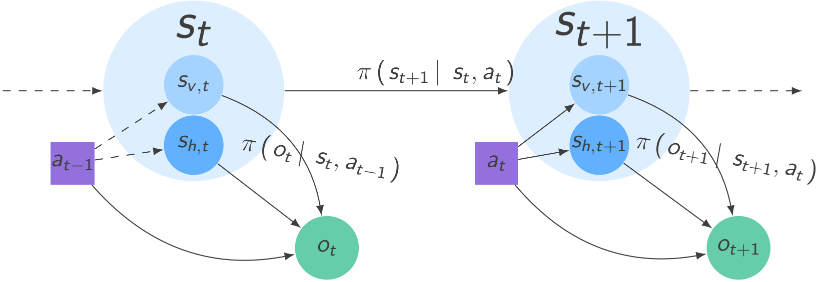
Dynamic Programming scheme: $\#$ iterations $< \#\mathcal{S}$

- 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** of the resulting **stationary** strategy

Mixed-Observability (MOMDP, *Ong et al., 2005*)

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

graphical model of a π-MOMDP:



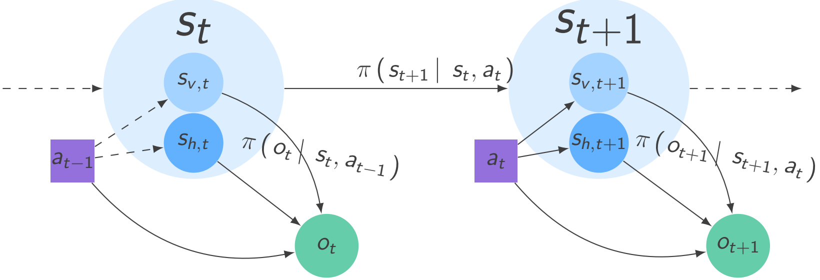
Mixed-Observability: system state $s \in \mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$

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

Mixed-Observability (MOMDP, *Ong et al., 2005*)

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

graphical model of a π-MOMDP:



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

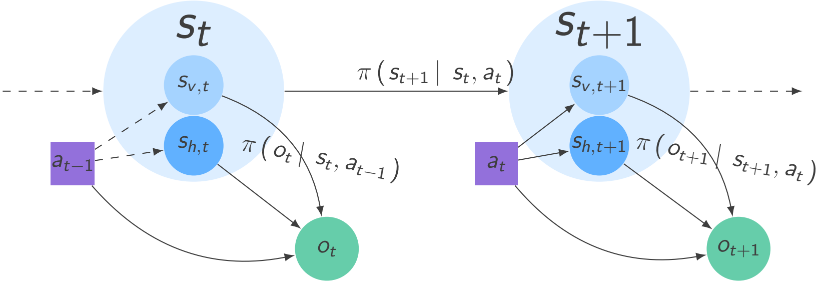
i.e. state s = visible component s_v & hidden component s_h

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

Mixed-Observability (MOMDP, *Ong et al., 2005*)

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

graphical model of a π-MOMDP:



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

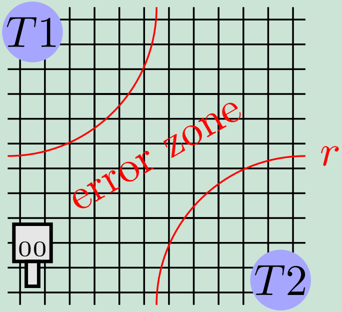
- belief states only over $\mathcal{S}_h$ (component s_v observed)
- $\rightarrow$ π-POMDP: belief space $\Pi_{\mathcal{L}}^{\mathcal{S}}$ $\quad \#\Pi_{\mathcal{L}}^{\mathcal{S}} \sim \#\mathcal{L}^{\#\mathcal{S}}$
  $\rightarrow$ π-MOMDP: computations on $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$
  $\#\mathcal{X} \sim \#\mathcal{S}_v \cdot \#\mathcal{L}^{\#\mathcal{S}_h} \ll \#\Pi_{\mathcal{L}}^{\mathcal{S}}$

π-MOMDP for robotics with imprecise probabilities

simulations with machine vision behavior imprecisely known

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

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

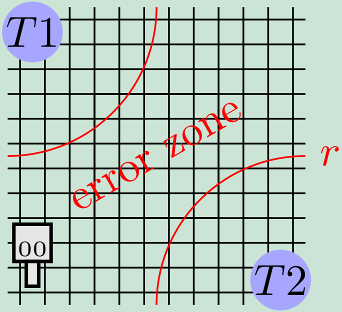


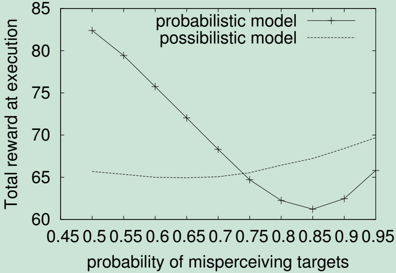
π-MOMDP for robotics with imprecise probabilities

simulations with machine vision behavior imprecisely known

- **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: $P_{bad} > \frac{1}{2}$

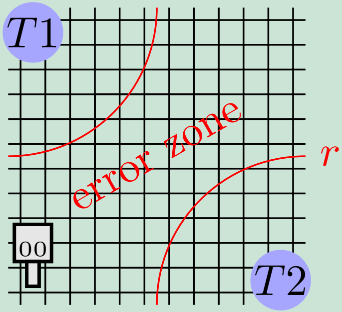
π-MOMDP for robotics with imprecise probabilities

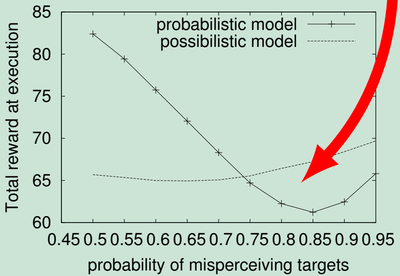
simulations with machine vision behavior imprecisely known

- **goal:** reach the object A
- noisy observations of the

probabilistic model inappropriate with too imprecise probabilities

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



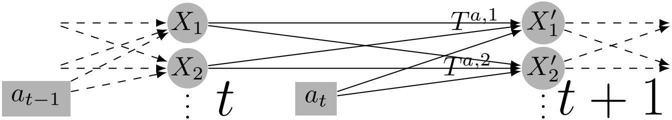


in reality, misperception probability in the error zone: $P_{bad} > \frac{1}{2}$

Factored π-MOMDP and computations with ADDs

qualitative possibilistic models to reduce complexity

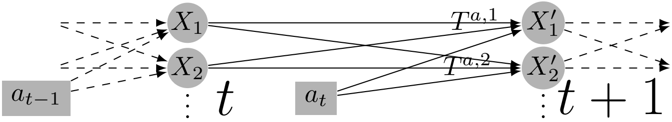
contribution (AAAI-14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$ = Boolean variables $(X_1, \dots, X_n)$
+ independence assumptions $\Leftarrow$ graphical model



Factored π-MOMDP and computations with ADDs

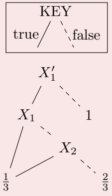
qualitative possibilistic models to reduce complexity

contribution (AAAI-14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$ = Boolean variables $(X_1, \dots, X_n)$
+ 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., 1999*)



example of ADD

Simplify computations with π-MOMDPs

Resulting π-MOMDP solver: PPUDD

- 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**.

Simplify computations with π-MOMDPs

Resulting π-MOMDP solver: PPUDD

- 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**

Simplify computations with π-MOMDPs

Resulting π-MOMDP solver: PPUDD

- 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

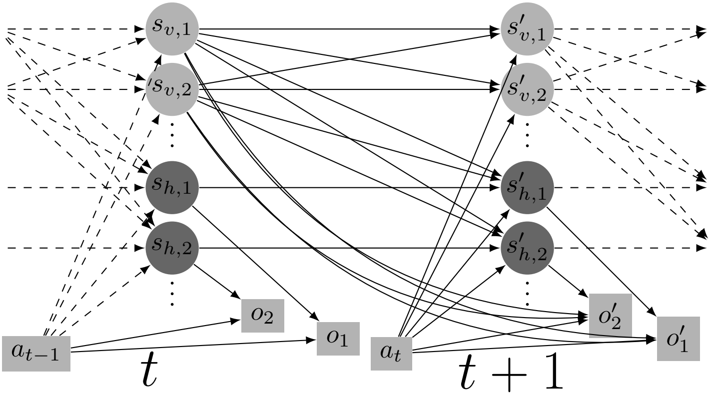
- factorization $\Rightarrow$ each DP steps divided into n stages
  $\rightarrow$ smaller ADDs $\Rightarrow$ **faster computations**
- computations on trees: *CU Decision Diagram Package*.

Simplifying computations with π-MOMDPs

Natural factorization: belief independence

contribution (AAAI-14):

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



Simplifying computations with π-MOMDPs

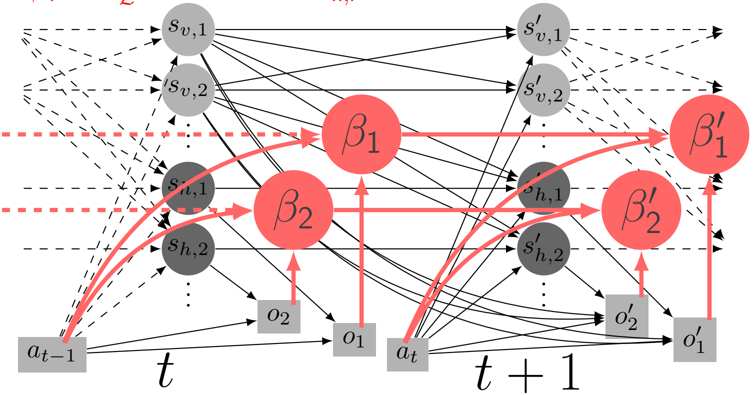
Natural factorization: belief independence

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 $s_{h,i}$

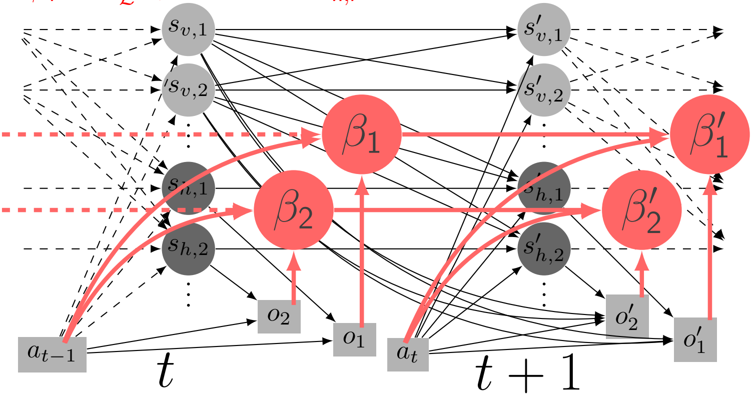


Simplifying computations with π-MOMDPs

Natural factorization: belief independence

$\perp\!\!\!\perp$ assumptions on state & observation variables
$\rightarrow$ belief variable factorization
$\rightarrow$ **additional** computation savings

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



Simplify computations with π-MOMDPs

Experiments – perfect sensing: Navigation problem

PPUDD vs SPUDD (*Hoey et al.*, 1999)

Navigation benchmark: reach a goal – spots with accident risk

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

Simplify computations with π-MOMDPs

Experiments – perfect sensing: Navigation problem

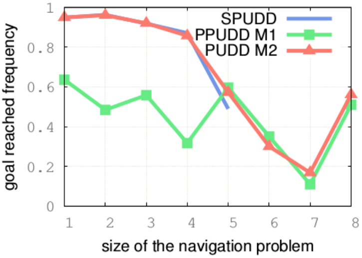
PPUDD vs SPUDD (*Hoey et al.*, 1999)

Navigation benchmark: reach a goal – spots with accident risk

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

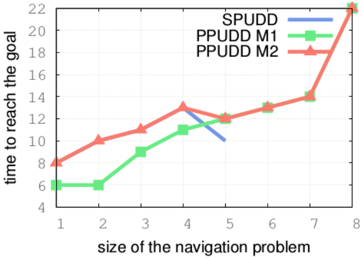
Performances, function of the problem index

reached goal frequency



the higher the better

steps to reach the goal

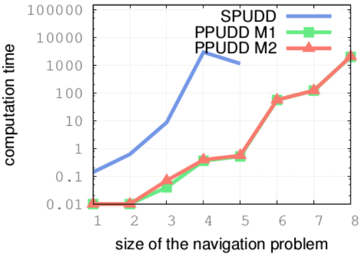


the lower the better

Simplify computations with π-MOMDPs

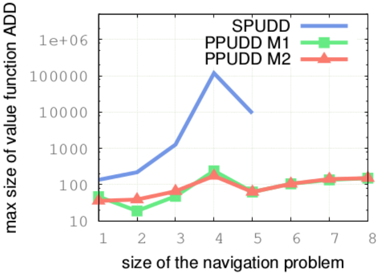
Experiments – perfect sensing: Navigation problem

computation time



the lower the better

max size of ADDs



the lower the better

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

Simplify computations with π-MOMDPs

Experiments – imperfect sensing: RockSample problem

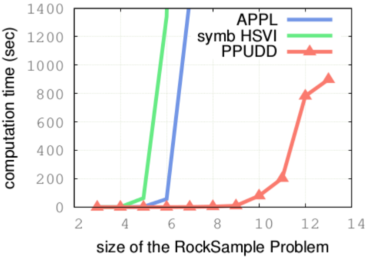
PPUDD vs APPL (*Kurniawati et al.*, 2008, solver MOMDP)
symbolic HSVI (*Sim et al.*, 2008, solver POMDP)
RockSample benchmark: recognize and sample “good” rocks

Simplify computations with π-MOMDPs

Experiments – imperfect sensing: RockSample problem

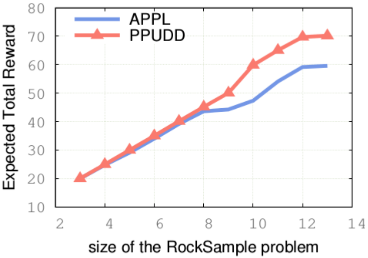
PPUDD vs APPL (*Kurniawati et al.*, 2008, solver MOMDP)
symbolic HSVI (*Sim et al.*, 2008, solver POMDP)
RockSample benchmark: recognize and sample "good" rocks

computation time:



the lower the better

average of rewards

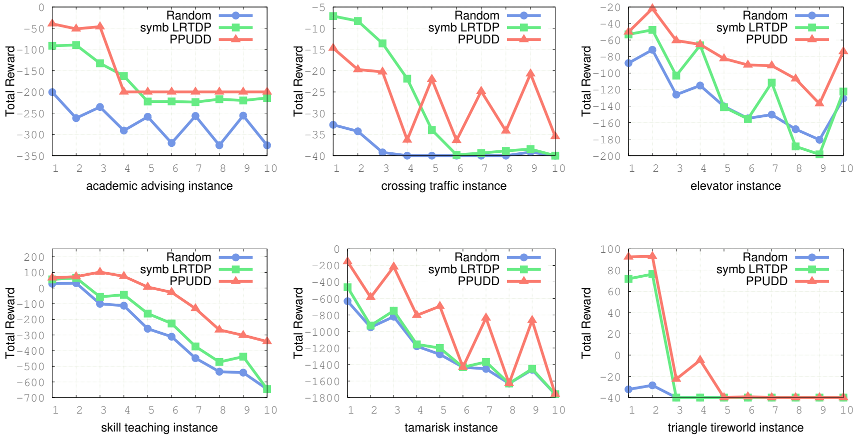


the higher the better

- approximate model + exact resolution solver can be **better than** exact model + approximate resolution solver

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

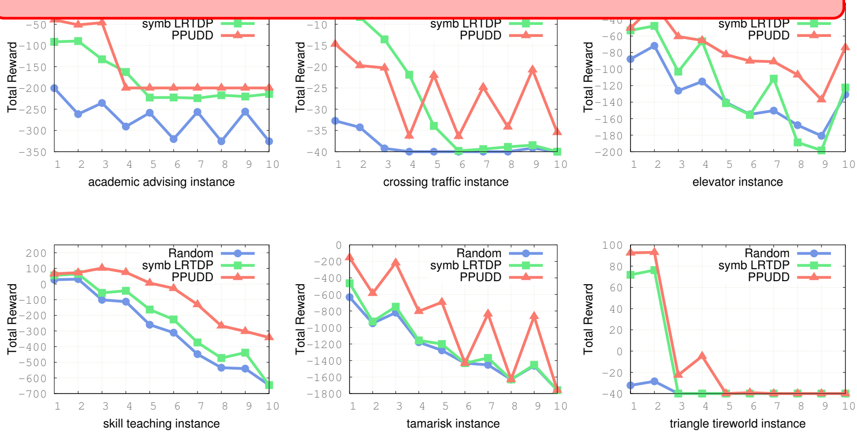
PPUDD + BDD mask over reachable states.



average of rewards over simulations – the higher the better

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

PPUDD outperforms its probabilistic counterpart



average of rewards over simulations – the higher the better

Qualitative possibilistic approach: benefits/drawbacks

towards a hybrid POMDP

- **granulated** belief space (discrete)
- efficient problem **simplification** (PPUDD 2× 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

→ **heuristic** for solving POMDPs:
results in a standard (finite state space) MDP
→ 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

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

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** $\tilde{\mathcal{S}} = \Pi_{\mathcal{L}}^{\mathcal{S}}$,
  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 $\tilde{T}\left(\beta, a, \beta'\right) = \mathbf{p}\left(\beta' \middle| \beta, a\right)$

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 $\tilde{T}\Big(\beta, a, \beta'\Big) = \mathbf{p}\Big(\beta' \Big| \beta, a\Big)$
- **reward** $\tilde{r}(a, \beta) = \underline{Ch}\Big(r(a, .)\Big)$,

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 $\tilde{T}\Big(\beta, a, \beta'\Big) = \mathbf{p}\Big(\beta' \Big| \beta, a\Big)$
- **reward** $\tilde{r}(a, \beta) = \underline{Ch}\Big(r(a, .)\Big)$,

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

Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

s'_v

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

s'_h

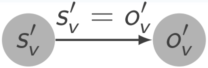
fully hidden $s'_f \in \mathbb{S}_f$

s'_f

Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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



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

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

Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

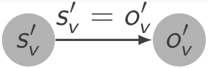
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

---

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

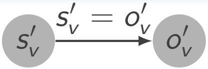
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

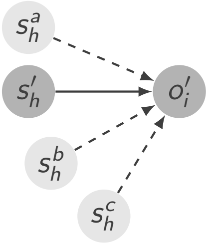
variable: **visible** $s'_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable

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



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



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

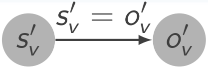
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

$\Leftrightarrow$ deterministic belief variable

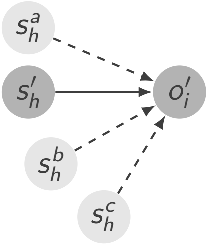
$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



---

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

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



---

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

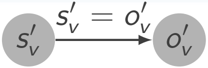
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

$\Leftrightarrow$ deterministic belief variable

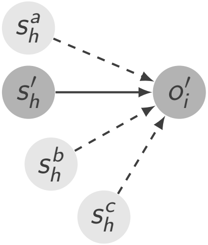
$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



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

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

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



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

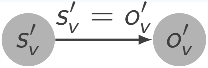
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

$\Leftrightarrow$ deterministic belief variable

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

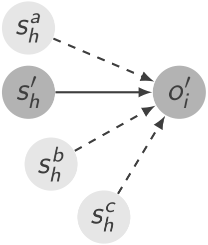


inferred hidden $s'_h \in \mathbb{S}_h$

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

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

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



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

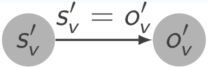
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

$\Leftrightarrow$ deterministic belief variable

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

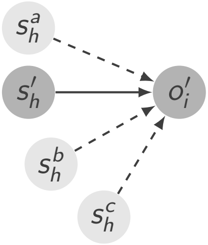


inferred hidden $s'_h \in \mathbb{S}_h$

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

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

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



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

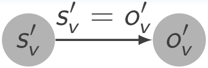
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

$\Leftrightarrow$ deterministic belief variable

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

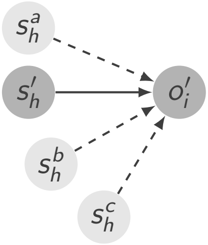


inferred hidden $s'_h \in \mathbb{S}_h$

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

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

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



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

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

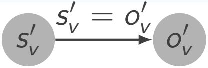
Belief variable factorization

3 classes of state variables – **contribution (SUM-15)**

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

$\Leftrightarrow$ deterministic belief variable

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

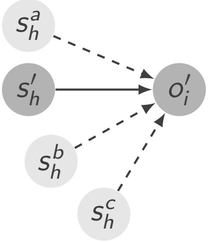


inferred hidden $s'_h \in \mathbb{S}_h$

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

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

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



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

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

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

Belief variable factorization

global belief state from marginal belief variables

bound over the global belief state

$$\beta_{t+1}(s'_1, \dots, s'_n) = \pi\left(s'_1, \dots, s'_n \;\middle|\; a_0, o_1, \dots, a_t, o_{t+1} \right)$$

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

Belief variable factorization

global belief state from marginal belief variables

bound over the global belief state

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

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

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

Conclusion
contributions

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

Conclusion
contributions

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

Conclusion

contributions

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

Conclusion

contributions

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

Conclusion

contributions

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

Conclusion

perspectives

- refined criteria *(Weng 2005, Dubois et al. 2005)*
  $\Rightarrow$ finer π-POMDP
- combination with reinforcement learning

Conclusion

perspectives

- refined criteria *(Weng 2005, Dubois et al. 2005)*
  $\Rightarrow$ finer π-POMDP
- combination with reinforcement learning

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!

produced work:

- *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 french journal**)
- *A Possibilistic Estimation of Human Attentional Errors*, submitted to **IEEE-TFS journal**