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

Automatics, AI, Flight Mechanics, Cognitive Sciences

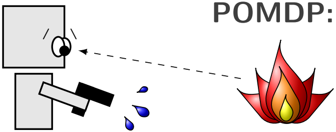
Onera, DCSD

Automatics, AI, Flight Mechanics, Cognitive Sciences

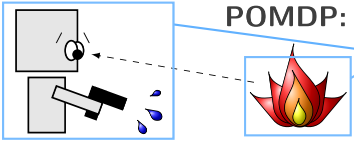
among many other works:

- autonomy, steering architectures and human factors
- decision making, planning
- experimental/industrial applications: UAVs,
  orbital systems,
  exploration robots

POMDP: model for autonomous robotics



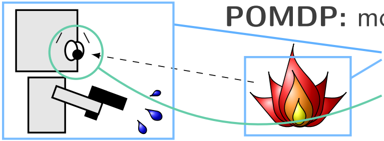
POMDP: model for autonomous robotics



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

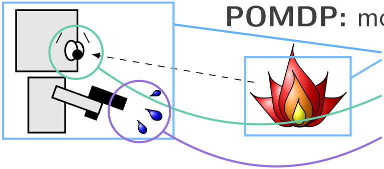
Context

Partially Observable Markov Decision Processes (POMDPs)



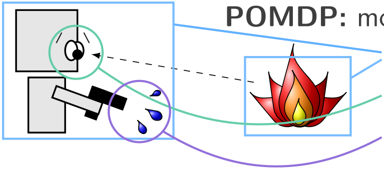
POMDP: model for autonomous robotics

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



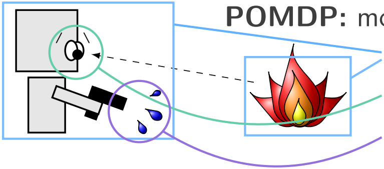
POMDP: model for autonomous robotics

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



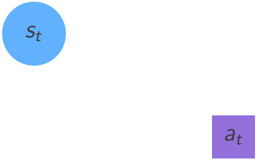
POMDP: model for autonomous robotics

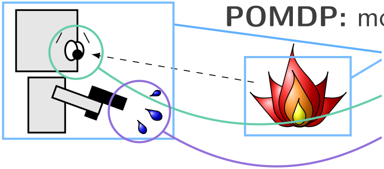
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**



POMDP: model for autonomous robotics

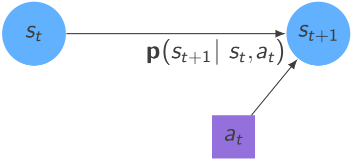
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**

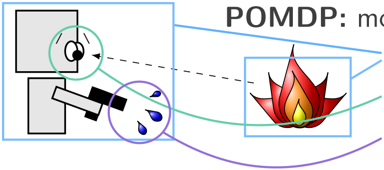




POMDP: model for autonomous robotics

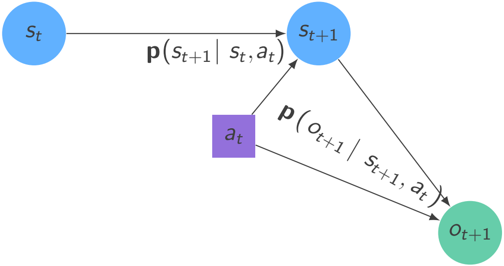
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**

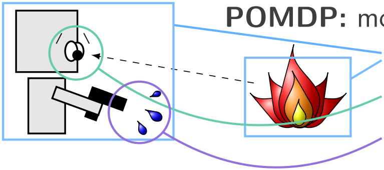




POMDP: model for autonomous robotics

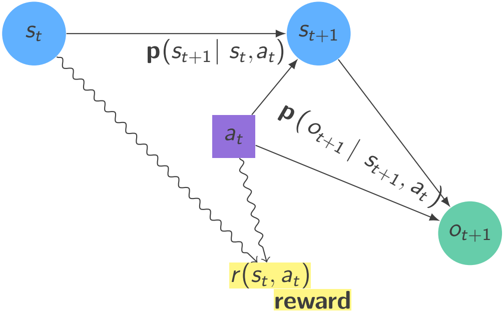
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**

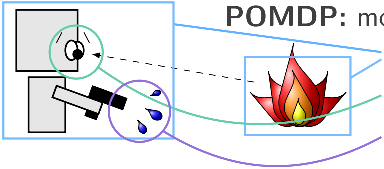




POMDP: model for autonomous robotics

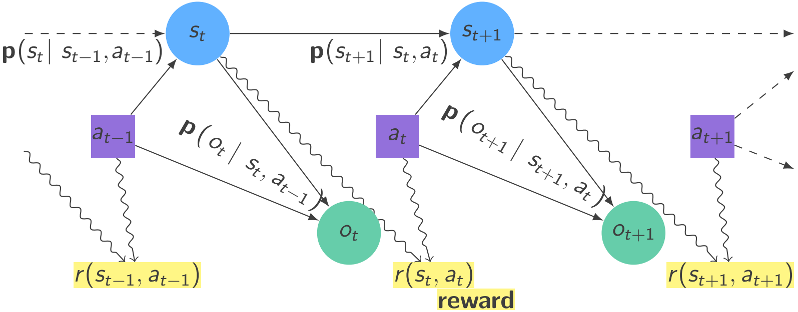
- $s \in \mathcal{S}$: **system states**
- $o \in \mathcal{O}$: **system observations**
- $a \in \mathcal{A}$: **agent's actions**





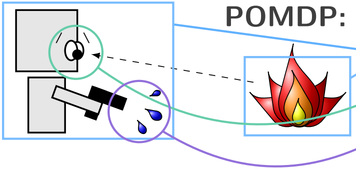
POMDP: model for autonomous robotics

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



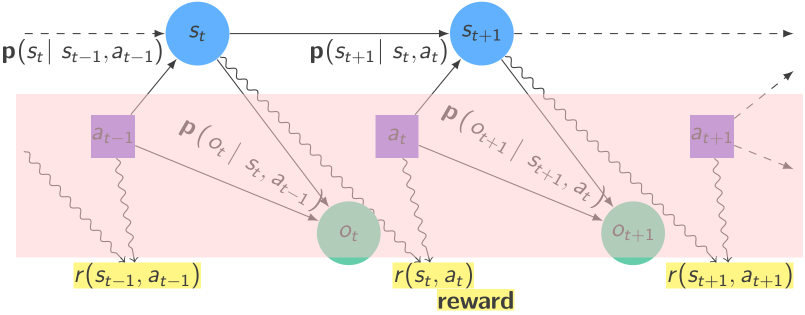
Context

Partially Observable Markov Decision Processes (POMDPs)

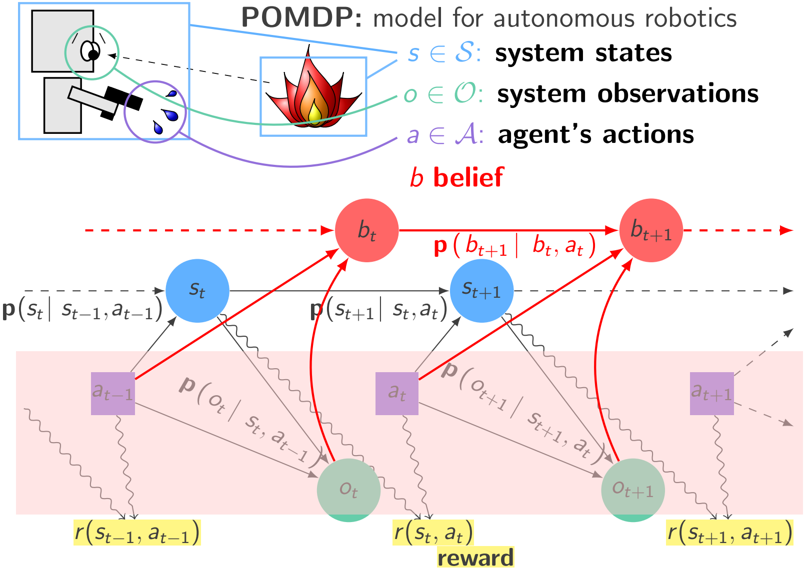


POMDP: model for autonomous robotics

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



POMDP: model for autonomous robotics



POMDP: $\langle \mathcal{S}, \mathcal{A}, \mathcal{O}, T, O, r, \gamma \rangle$ *(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
- agent's **ignorance** not taken into account

- **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)*

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

- **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

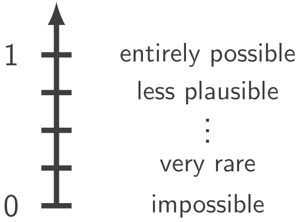
→ unknown environments: image variability of the datasets?

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

→uniform probability distribution ≠ ignorance!

finite scale $\mathcal{L}$

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



events $e \subset \Omega$ (universe)

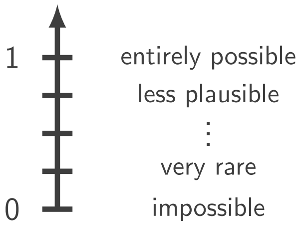
sorted using possibility **degrees** $\pi(e) \in \mathcal{L}$

$\neq$

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

finite scale $\mathcal{L}$

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



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** $\mathcal{S}$
  π-POMDP $\rightarrow$ π-MDP with **finite** $\mathcal{S}$
- $\forall s \in \mathcal{S}$, $\pi(s) = 1 \Leftrightarrow$ 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

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

$\rightarrow$ *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

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

$\rightarrow$ *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}: $\pi\left(o', s' \;\middle|\; b_t^{\pi}, a \right)$

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**

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

possibilistic belief update – a selected, o' received

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

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

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

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**

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

possibilistic belief update – a selected, o' received

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

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

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

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

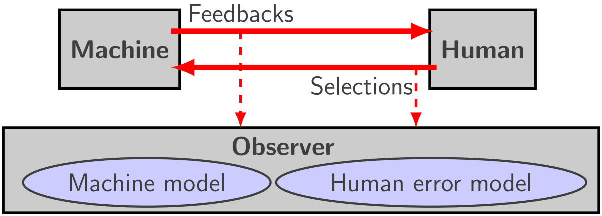
Qualitative Possibility Theory:
$\rightarrow$ simplification, ignorance and imprecision modeling

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

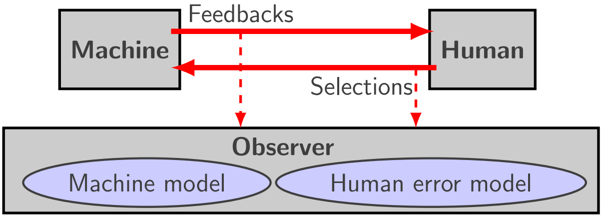
1. introduction
2. natural use of a qualitative possibilistic model
3. updates and first use of the π-POMDP model
4. simplify computation: ADDs and factorization
5. probabilistic-possibilistic (hybrid) approach
6. conclusion

Overview

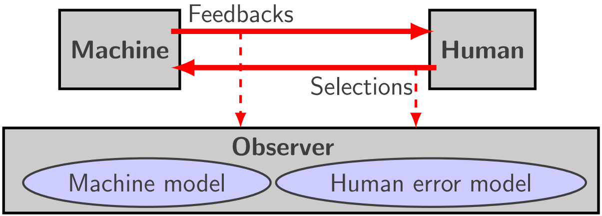
Example: Human-Machine Interaction
joint work with Sergio Pizziol – Context



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



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



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 selections

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 with states A, B, C, ...

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

Machine state $A \rightarrow B$

- **machine feedback** observation $o_f \in \mathcal{O}$:

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

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

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

Machine state $A \rightarrow B$

- **machine feedback** observation $o_f \in \mathcal{O}$:

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

- **human selection** observation $o_s \in \mathcal{O}$:

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

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 $A \rightarrow B$

- **machine feedback** observation $o_f \in \mathcal{O}$:

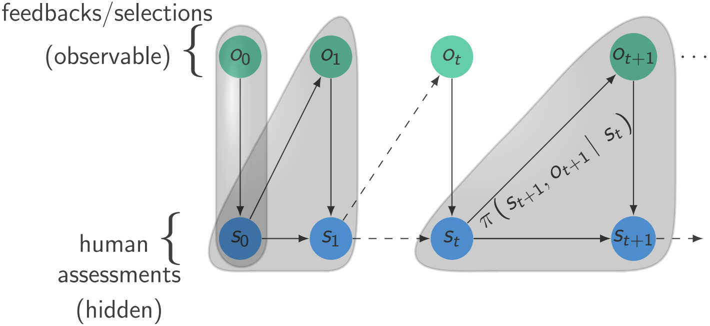
human usually aware of feedbacks $\rightarrow \pi\left(s'_B, o'_f \mid s_A\right) = 1$
but may lost a feedback $\rightarrow \pi\left(s'_A, o'_f \mid s_A\right) = \frac{2}{3}$

- **human selection** observation $o_s \in \mathcal{O}$:

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

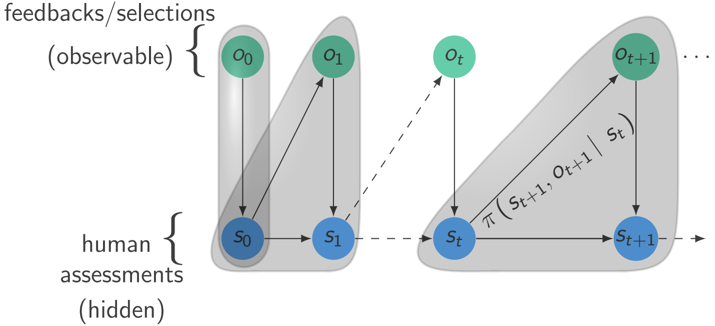
- impossible cases: possibility degree 0

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



Qualitative Possibilistic Hidden Markov Process:

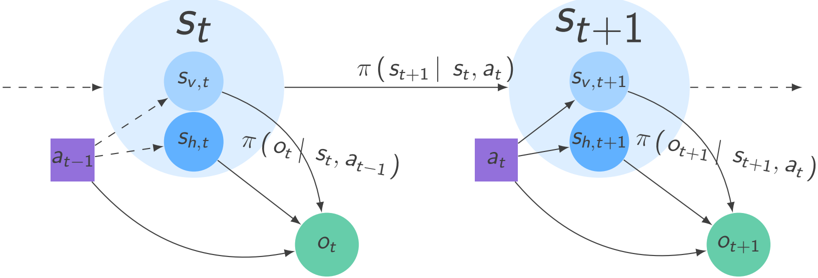
diagnosis tool for Human-Machine interaction (with Sergio Pizziol)



- **estimation** of the human assessment
  $\Leftrightarrow$ **possibilistic belief state**
- **detection** of human assessment errors
- **diagnosis** using *leximin* operator
- results on flight simulator missions with pilots

Overview

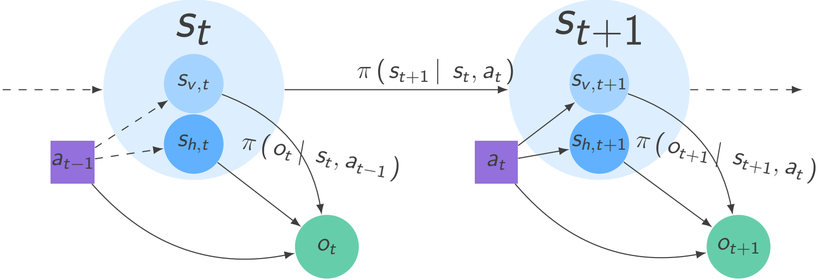
contribution (UAI-13):



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):

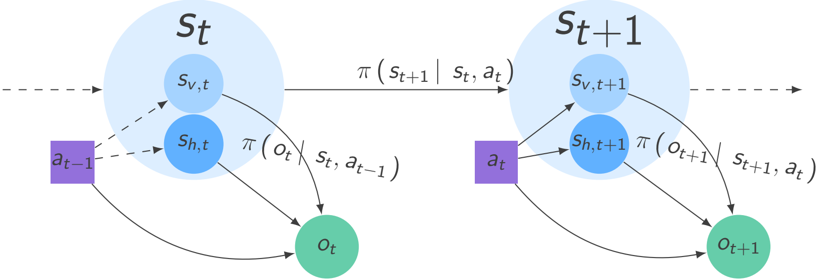


Mixed-Observability: system state $s \in \mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$

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

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

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

contribution (UAI-13): Undefined Horizon

contribution (UAI-13): Undefined Horizon

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

- assumption: $\exists$ artificial "stay" action
  as in classical planning/ γ counterpart
- value function $=$ criterion: non decreasing with horizon

contribution (UAI-13): Undefined Horizon

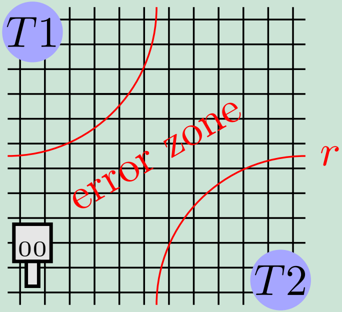
Dynamic Programming scheme: $\#$ iterations $< \#\mathcal{X}$

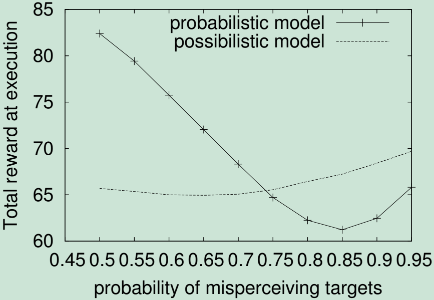
- assumption: $\exists$ artificial "stay" action
  as in classical planning/ γ counterpart
- value function $=$ criterion: non decreasing with horizon
- action update for states increasing the value function
- proof of optimality

Use of the π-MOMDP in practice
simulations

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

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





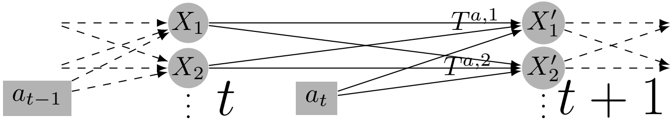
in reality, misperception probability in the error zone: $P_{bad} > \frac{1}{2}$

Overview

Factored π-MOMDP and computations with ADDs

qualitative possibilistic models to reduce complexity

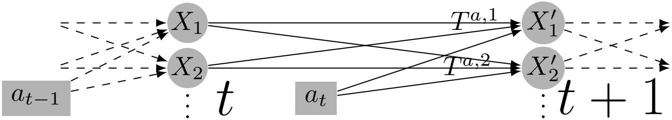
contribution (AAAI-14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$ = Boolean variables $(X_1, \dots, X_n)$
$+$ independence assumptions $\Leftarrow$ graphical model



Factored π-MOMDP and computations with ADDs

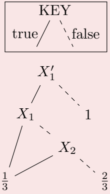
qualitative possibilistic models to reduce complexity

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

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**.

- 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 $\Rightarrow$ DP steps divided into n stages
  $\rightarrow$ used ADDs smaller $\Rightarrow$ **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

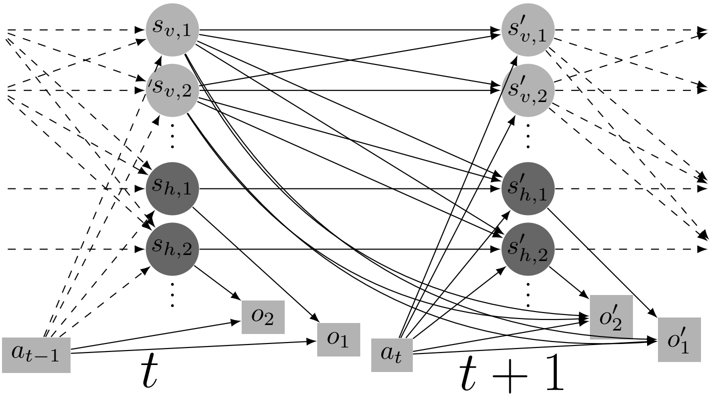
- factorization $\Rightarrow$ DP steps divided into n stages
  $\rightarrow$ used ADDs smaller $\Rightarrow$ **faster computations**
- computations on trees: *CU Decision Diagram Package*.

Simplify computations with π-MOMDPs

Natural factorization: belief independence

contribution (AAAI-14):

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

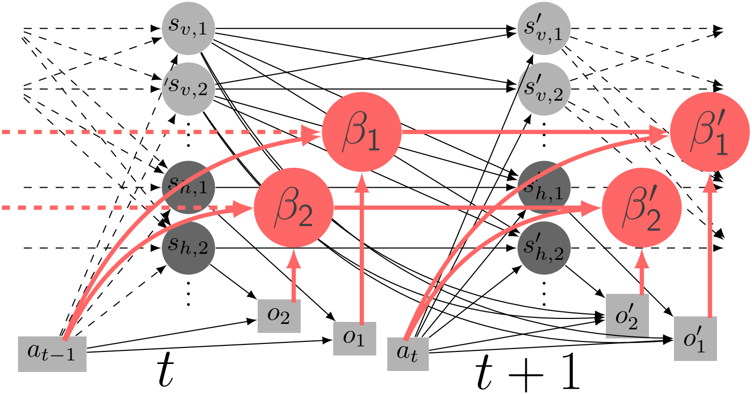


contribution (AAAI-14):

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

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

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

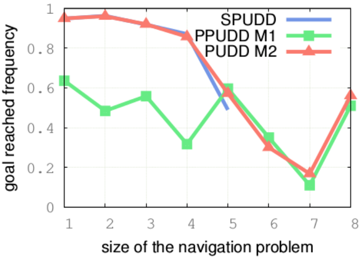
PPUDD vs SPUDD *Hoey et al.*

Navigation benchmark: reach a goal – spots with accident risk

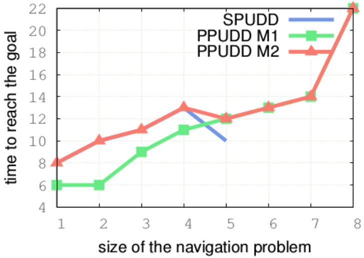
M1 (resp. M2) optimistic (resp. cautious) criterion

Performances, function of the instance size

reached goal frequency

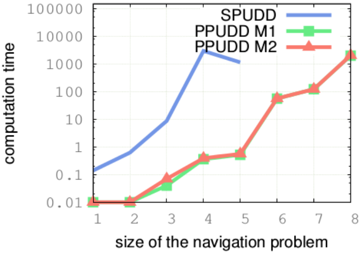


steps to reach the goal

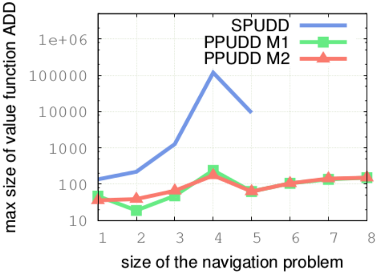


ONERA
THE FRENCH AEROSPACE LAB

computation time



max size of ADDs

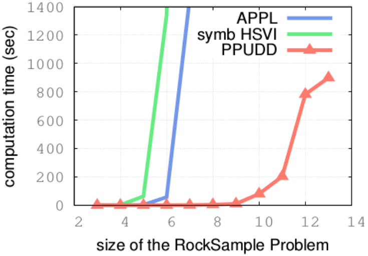


- PPUDD + M2 (pessimistic criterion) **faster with same performances** as SPUDD
- SPUDD only solves the first 5 instances
- verified intuition: ADDs are smaller

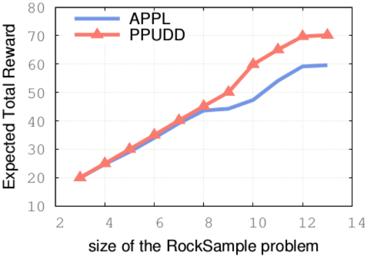
PPUDD vs APPL *Kurniawati et al.*, solver MOMDP
symbolic HSVI *Sim et al.*, solver POMDP
RockSample benchmark: recognize and sample "good" rocks

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

PPUDD + BDD mask over reachable states.

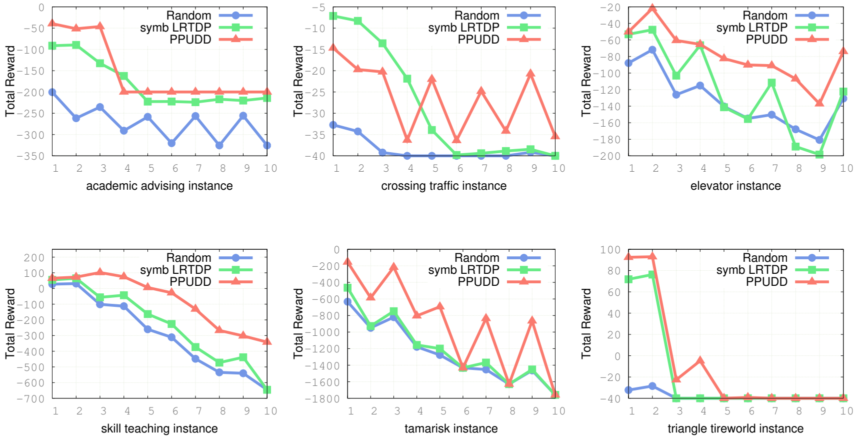


Figure : average of rewards over simulations

Overview

Qualitative possibilistic approach: benefits/drawbacks

towards a hybrid POMDP

Qualitative Possibilistic models:

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

Qualitative Possibilistic models:

- **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: winners of IPPC 2014, $2\times$ better than PPUDD
- choice of the qualitative criterion (optimistic/pessimistic)
- non additive utility degrees
  same scale as possibility degrees (commensurability)
- frequentist information lost

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

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

results in a standard (finite state space) MDP

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

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

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

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

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

- 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 summary **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 summary **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 summary **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 $\tilde{T}\left(\beta, a, \beta'\right) = \mathbf{p}\left(\beta' \middle| \beta, a\right)$;

translation summary **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** $\tilde{r}(a, \beta) = \underline{Ch}\Big(r(a, .)\Big)$,

Resulting MDP

translation summary **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** $\tilde{r}(a, \beta) = \underline{Ch}\Big(r(a, .)\Big)$,

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

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

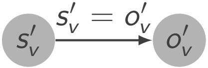
s'_h

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



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

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

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

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

$\Leftrightarrow$ deterministic belief variable

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

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

---

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

$$s'_h$$

---

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

$$s'_f$$

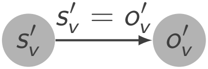
General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

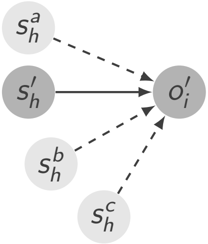
<ins>variable:</ins> visible $s'_v \in \mathbb{S}_v$

$\Leftrightarrow$ 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 $s'_f \in \mathbb{S}_f$

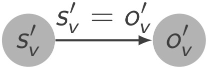
General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

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

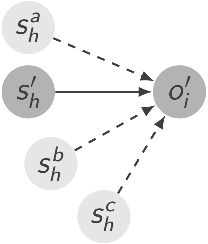
$\Leftrightarrow$ deterministic belief variable

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



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

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



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

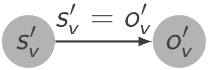
General variable classification **contribution (SUM-15):**

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable

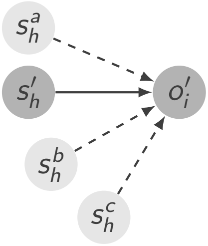
$$\beta_{t+1}(s'_v) = \mathbb{1}_{\{s'_v = o'_v\}}(s'_v)$$



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

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

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

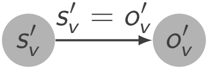


fully hidden $s'_f \in \mathbb{S}_f$

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

$\Leftrightarrow$ deterministic belief variable

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

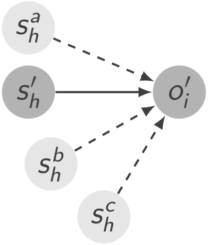


inferred hidden $s'_h \in \mathbb{S}_h$

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

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

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

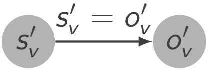


fully hidden $s'_f \in \mathbb{S}_f$

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

$\Leftrightarrow$ deterministic belief variable

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



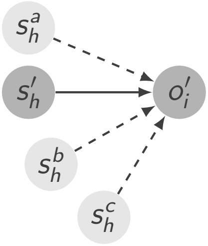
---

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.



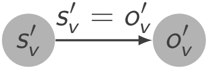
---

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

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

$\Leftrightarrow$ deterministic belief variable

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

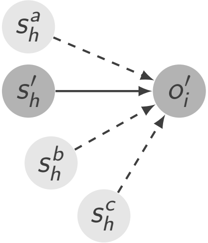


inferred hidden $s'_h \in \mathbb{S}_h$

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

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

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



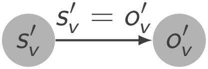
fully hidden $s'_f \in \mathbb{S}_f$

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

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

$\Leftrightarrow$ deterministic belief variable

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

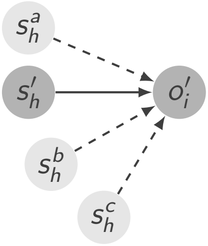


inferred hidden $s'_h \in \mathbb{S}_h$

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

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

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



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

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

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

Possibilistic belief variables

global belief state

bound over the global belief state

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

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

bound over the global belief state

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

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

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

Overview

Conclusion
contributions

- udates: → mixed-observability modeling
  → undefined horizon
- modeling: → human-machine interaction
  → robust recognition mission with possibilistic beliefs
- computations: factorization work & PPUDD algorithm (competitive solver, IPPC 2014)

- udates: $\rightarrow$ mixed-observability modeling
  $\rightarrow$ undefined horizon
- modeling: $\rightarrow$ human-machine interaction
  $\rightarrow$ robust recognition mission with possibilistic beliefs
- computations: factorization work & PPUDD algorithm (competitive solver, IPPC 2014)

new refined criteria $\rightarrow$ finer π-POMDP

Conclusion
contributions

- udates: $\rightarrow$ mixed-observability modeling
  $\rightarrow$ undefined horizon
- modeling: $\rightarrow$ human-machine interaction
  $\rightarrow$ robust recognition mission with possibilistic beliefs
- computations: factorization work & PPUDD algorithm (competitive solver, IPPC 2014)

new refined criteria $\rightarrow$ finer π-POMDP

quantitative information may be available: hybrid work

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

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

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

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

Conclusion
perspectives

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

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

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

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

perspectives:

- IPPC problems (factored POMDPs);

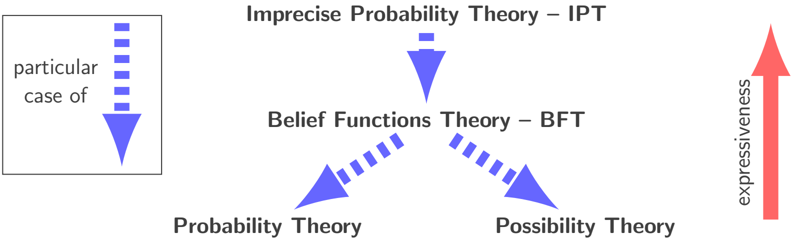
POMDP $\xrightarrow{\textbf{translation}}$ MDP with finite state space

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

perspectives:

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

Thank you!

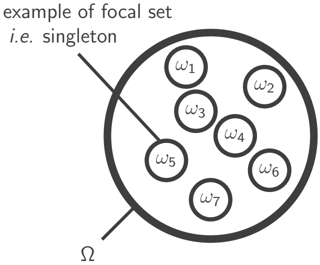


- IPT: most general uncertainty theory.
  Use of sets of probability measures over Ω.
- BFT: use of a mass function $m : 2^{\Omega} \to [0, 1]$, with $\sum_{A \subset \Omega} m(A) = 1$.
  1. plausibility measure: $\forall A \subset \Omega,\ Pl(A) = \sum_{B \cap A \neq \emptyset} m(B)$.
  2. belief function: $\forall A \subset \Omega,\ bel(A) = \sum_{B \subseteq A} m(B)$.

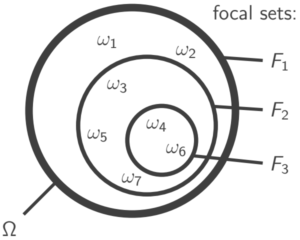
Focal sets of a mass function $m : 2^{\Omega} \to [0,1]$: subsets A of $\Omega = \{\omega_1, \ldots, \omega_7\}$ such that $m(A) > 0$.

- if focal sets are all singletons
  $\to$ probability distribution ($bel = Pl = \mathbb{P}$)
- if focal sets are nested, e.g. $F_3 \subset F_2 \subset F_1 = \Omega$,
  $\to$ possibility distribution:
  bel=**necessity measure**, *Pl*=**possibility measure**.

probabilistic case



possibilistic case



Probabilistic belief update

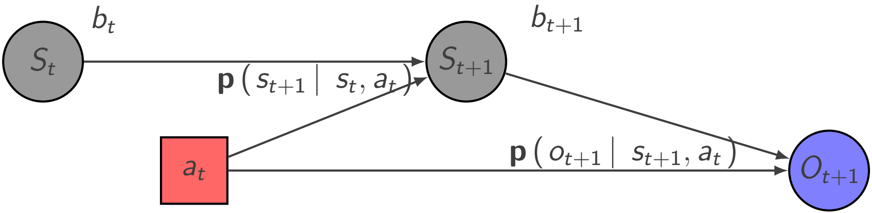


Figure : Bayesian Network illustrating the belief update

- the **system states** are the gray circular nodes,
- the **action** is the red square node ,
- and the **observation** is the blue circular node.

The belief state b_t (resp. b_{t+1}) is the probabilistic estimation of the current (resp. next) system state s_t (resp. s_{t+1})

probabilistic belief update

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

Rewritings of parameters
PROBABILISTIC parameters

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

Rewritings of parameters
PROBABILISTIC parameters

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

consequences on the joint distribution

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

Rewritings of parameters
PROBABILISTIC parameters

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

consequences on the joint distribution

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

observation probabilities

epistemic state

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

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

Parameters rewritings
POSSIBILISTIC parameters

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

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

marginal possibilistic belief states

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

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

Parameters rewritings
POSSIBILISTIC parameters

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

marginal possibilistic belief states

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

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

Parameters rewritings
POSSIBILISTIC parameters

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

marginal possibilistic belief states

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

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

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

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

A possibilistic belief state
finite belief space

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

$\rightarrow$ **finite belief space**

A possibilistic belief state
finite belief space

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

$\rightarrow$ **finite belief space**

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

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

$\rightarrow$ **finite belief space**

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

update – possibilistic belief state

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

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

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

$\rightarrow$ **finite belief space**

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

update – possibilistic belief state

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

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

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

A possibilistic belief state
finite belief space

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

$\rightarrow$ **finite belief space**

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

update – possibilistic belief state

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

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

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

ONERA THE FRENCH AEROSPACE LAB

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

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

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

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

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

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

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

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

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

Resulting π-MOMDP solver: PPUDD

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

Resulting π-MOMDP solver: PPUDD

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

PPUDD: Possibilistic Planning Using Decision Diagrams

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

computations on trees: *CU Decision Diagram Package*.

Resulting π-MOMDP solver: PPUDD

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

PPUDD: Possibilistic Planning Using Decision Diagrams

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

factorization

$\Rightarrow$ dynamic programming

computations on trees: *CU Decision Diagram Package*.

Resulting π-MOMDP solver: PPUDD

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

PPUDD: Possibilistic Planning Using Decision Diagrams

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

factorization
$\Rightarrow$ dynamic programming
divided into n stages

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

computations on trees: *CU Decision Diagram Package*.

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

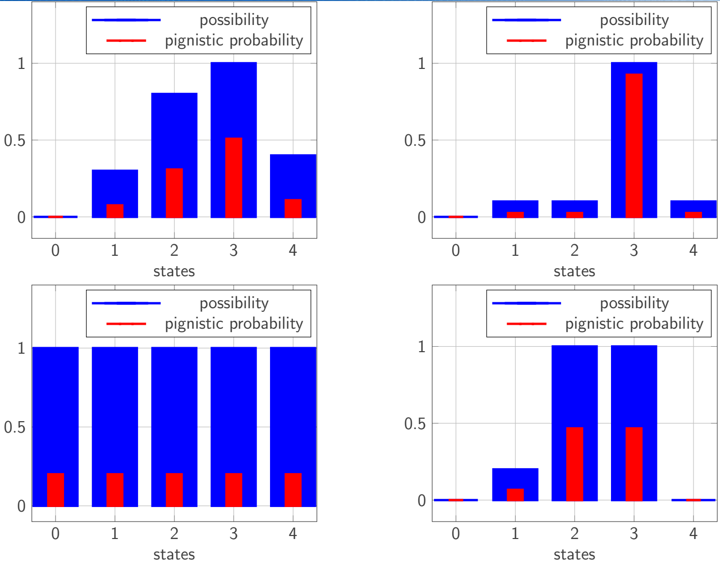
pignistic transformation – $P : \Pi_\mathcal{S} \rightarrow \mathbb{P}_\mathcal{S}$

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

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

Pignistic transformation

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



hybrid POMDP and π-POMDP

differences with possibilistic models

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

hybrid POMDP and π-POMDP

differences with possibilistic models

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

hybrid model:

- only belief states are possibilistic:

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

- probabilistic dynamics:

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

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

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

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

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

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

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

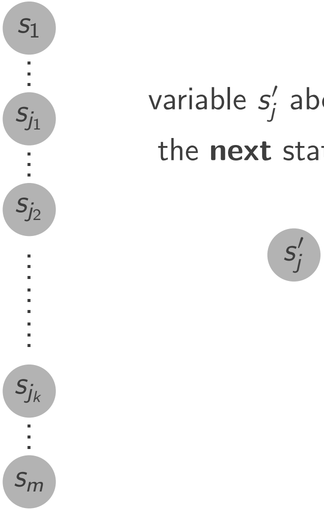
independences:

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

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

some variables does not interact with each other

variables about the **current** system state,



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

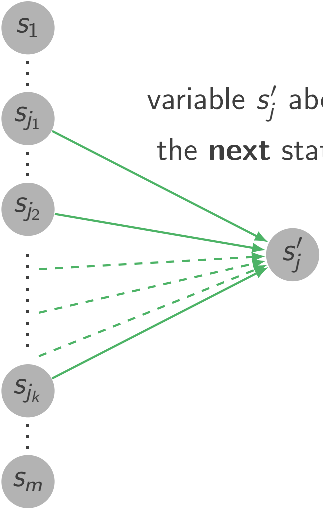
Notations

some variables does not interact with each other

variables about the **current** system state,

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

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



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

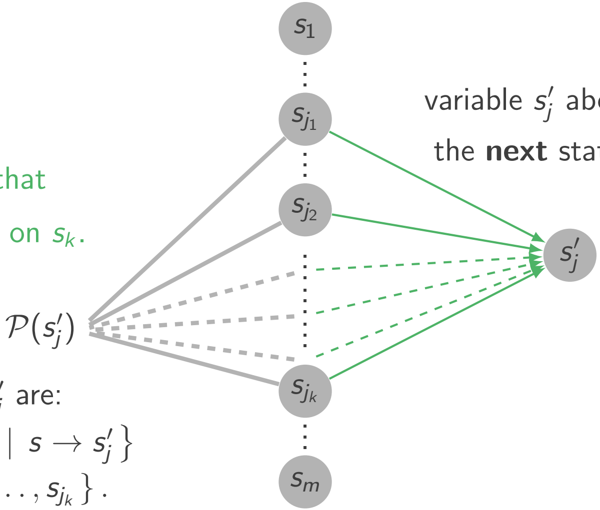
variables about the **current** system state,

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

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

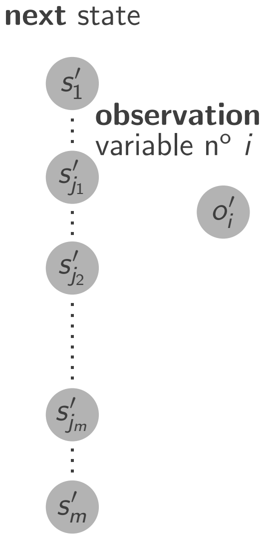
variable s'_j about

the **next** state.



the parents of s'_j are:

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



next state
s'_1
observation
variable n° i
s'_{j1}
o'_i
s'_{j2}
s'_{jm}
s'_m

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

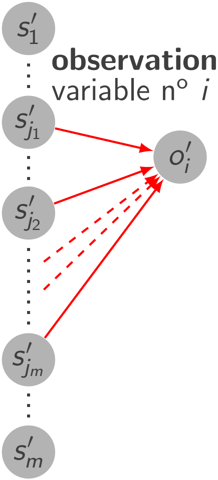
$$\Updownarrow$$

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

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

depends on s'_j.

next state



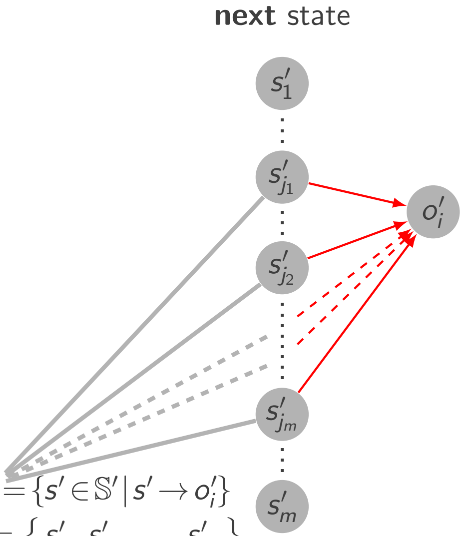
$$s'_j \to o'_i$$

$$\Updownarrow$$

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

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

depends on s'_j.



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

Notations

concerning observation variables

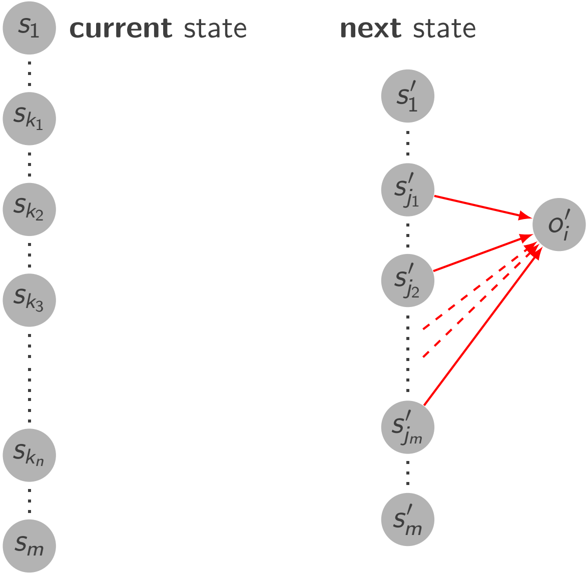
$s'_j \to o'_i$

$\Updownarrow$

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

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

depends on s'_j.

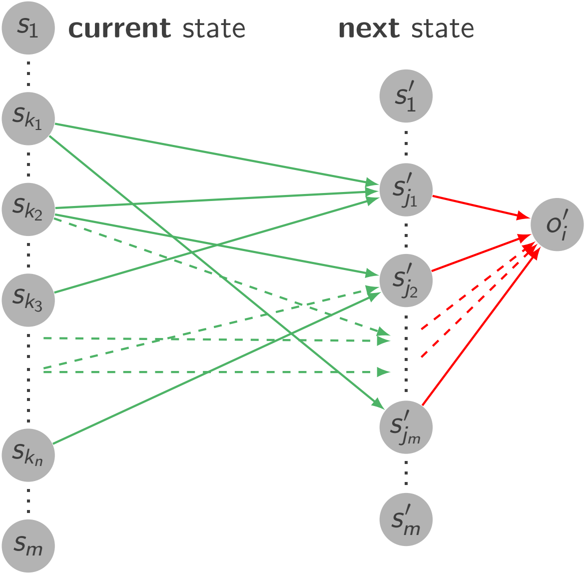


Notations
concerning observation variables

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

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

depends on s'_j.



Notations

concerning observation variables

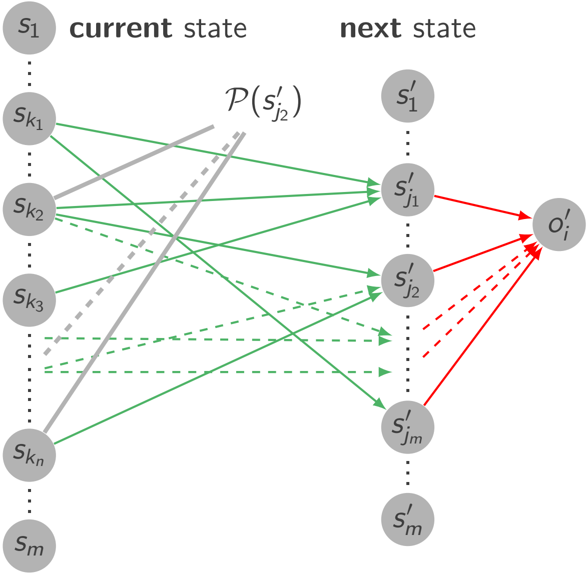
$$s'_j \rightarrow o'_i$$

$$\Updownarrow$$

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

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

depends on s'_j.



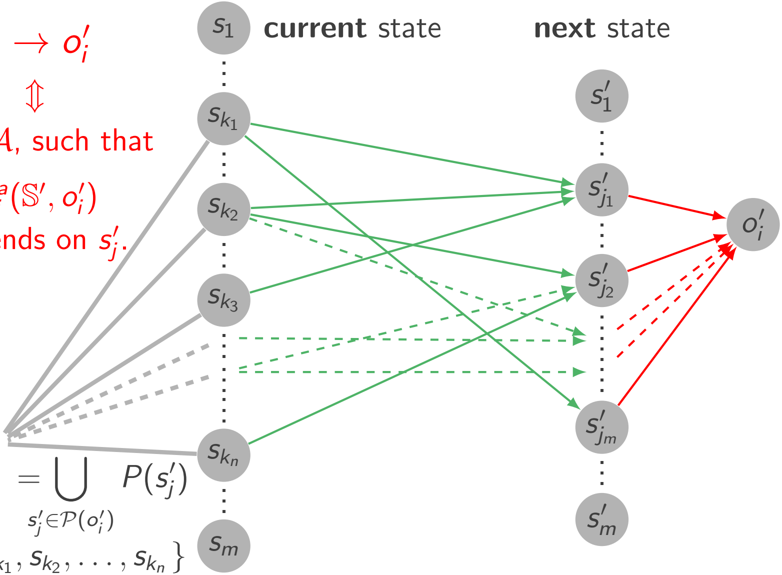
Notations

concerning observation variables

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

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

depends on s'_j.



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

Variables de croyance

different according to the class of the variable

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

Variables de croyance
different according to the class of the variable

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

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

Variables de croyance
different according to the class of the variable

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

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

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

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

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

Variables de croyance

different according to the class of the variable

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

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

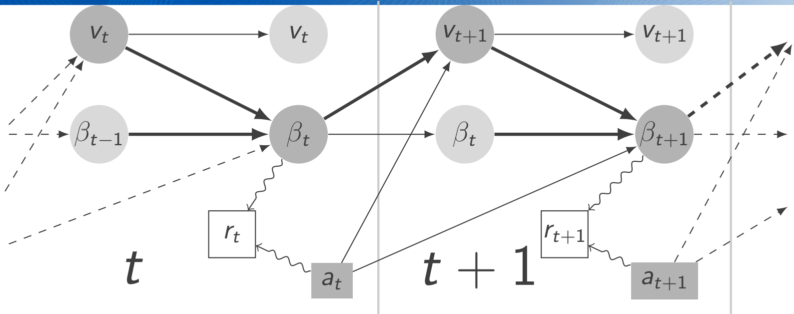
- $p_i = \#\mathcal{P}(o'_i)$.

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

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

- $\forall s'_f \in \mathbb{S}_f$, $\lambda^2 - (\lambda - 1)^2 = 2\lambda - 1$ belief states,

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

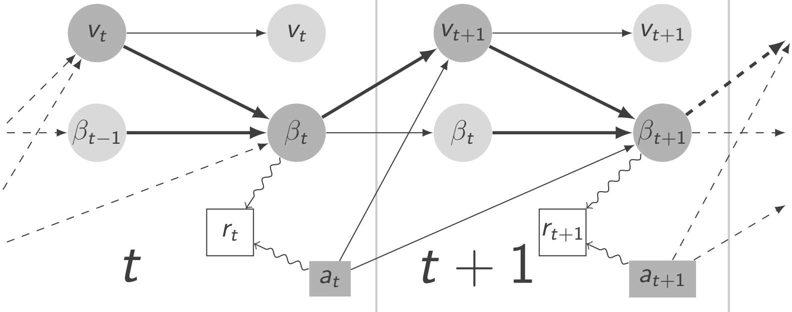


factorized model's variables: $\#\mathbb{O} + \#\mathbb{S}_v +$

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

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

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



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

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

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

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

Variable classification

3 classes of state variables – state space factorization

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

s'_v

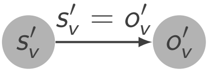
inferred hidden $s'_h \in \mathbb{S}_h$

s'_h

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

s'_f

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



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

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

Variable classification

3 classes of state variables – state space factorization

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

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

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

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

s'_h

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

s'_f

Variable classification

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable.

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

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

---

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

$$s'_h$$

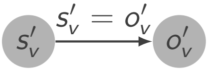
---

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

$$s'_f$$

Variable classification

3 classes of state variables – state space factorization



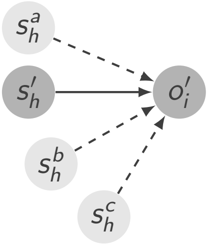
<u>variable:</u> visible $s'_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable.

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

---

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



---

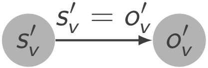
fully hidden $s'_f \in \mathbb{S}_f$

Variable classification

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable.

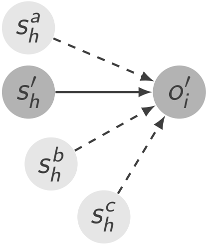


$$\mathbf{p}\left(s'_v \,\middle|\, b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\Big(\mathcal{P}(s'_v)\Big)$$

---

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

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



---

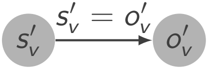
fully hidden $s'_f \in \mathbb{S}_f$

Variable classification

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable.



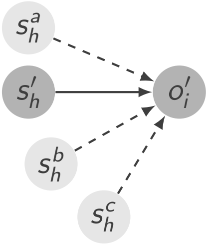
$$\mathbf{p}\left(s'_v \mid b^{\pi}_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^{\pi}_t}\left(\mathcal{P}(s'_v)\right)$$

---

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

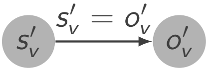
$$b^{\pi}_{t+1}(\mathcal{P}(o'_i)) = b^{\pi}_{t+1}(s_h, s^a_h, s^b_h, s^c_h)$$

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



---

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



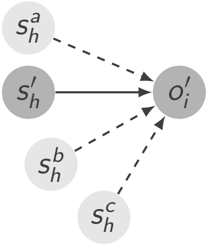
<u>variable:</u> visible $s'_v \in \mathbb{S}_v$

$\Leftrightarrow$ deterministic belief variable.

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

---

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



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

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

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

---

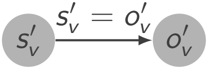
fully hidden $s'_f \in \mathbb{S}_f$

Variable classification

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable.

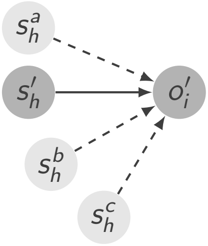


$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\left(\mathcal{P}(s'_v)\right)$$

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

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

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



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

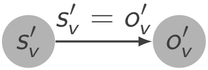
fully hidden $s'_f \in \mathbb{S}_f$

Variable classification

3 classes of state variables – state space factorization

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

$\Leftrightarrow$ deterministic belief variable.



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

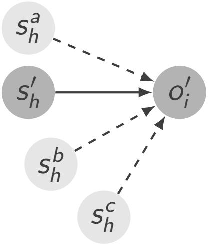
---

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

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

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

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



---

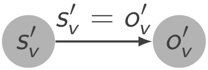
fully hidden $s'_f \in \mathbb{S}_f$

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

ONERA THE FRENCH AEROSPACE LAB

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

$\Leftrightarrow$ deterministic belief variable.



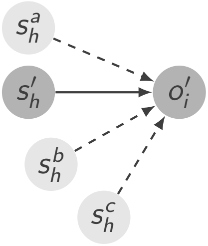
$$\mathbf{p}\left(s'_v \mid b^\pi_t, a\right) = \sum_{2^{\mathcal{P}(s'_v)}} T^a\left(\mathcal{P}(s'_v), s'_v\right) \cdot \overline{b^\pi_t}\left(\mathcal{P}(s'_v)\right)$$

---

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

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

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



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

---

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

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

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

Toy example: 2 machine states, 3 occurrences

columns		1	2	3	4	5
SITUATION						
v'	v_A	1				1
	v_B		1			
	v_C	1			1	
h	s_A	1	1		1	
	s_B			1		1
BEHAVIOUR						
h'	s_A					1
	s_B		1		1	
EFFECT		$\overline{e}$	$\tilde{e}$	$\overline{e}$	$\widehat{e}$	$\underline{e}$
POSSIBILITY		1	ε	1	λ	δ

Probability / Possibility :

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

- $r : \mathcal{S} \times \mathcal{A} \rightarrow \mathbb{R}$ reward function
- $\rho : \mathcal{S} \times \mathcal{A} \rightarrow \mathcal{L}$ preference function

Probability / Possibility:

$R(b_t, d_t) = \sum_s r(s, d_t) \cdot b_t(s)$	**optimistic:** $\overline{\Psi}(\beta_t, \delta_t) = \max_s \min \{ \rho(s, \delta_t), \beta_t(s) \}$ **pessimistic:** $\underline{\Psi}(\beta_t, \delta_t) = \min_s \max \{ \rho(s, \delta_t), 1 - \beta_t(s) \}$
$\mathbb{E}[r(S_t, d_t)] = \mathbb{E}[R(b_t, d_t)]$	$\mathbb{S}_\Pi[\rho(S_t, d_t)] = \mathbb{S}_\Pi[\overline{\Psi}(\beta_t, d_t)]$ $\mathbb{S}_\mathcal{N}[\rho(S_t, d_t)] = \mathbb{S}_\mathcal{N}[\underline{\Psi}(\beta_t, d_t)]$

Note: $\mathbb{S}_\Pi[\underline{\Psi}(\beta_t, d_t)]$; $\mathbb{S}_\mathcal{N}[\overline{\Psi}(\beta_t, d_t)]$?

Why model ignorance?

knowledge is not always encouraged with POMDPs

- initial belief deterministic $s_0 = s_A$.

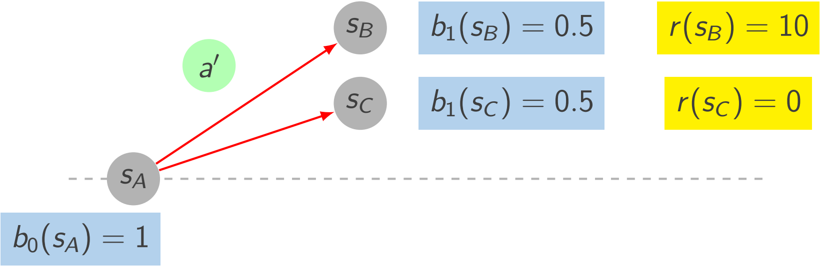
s_A

$b_0(s_A) = 1$

Why model ignorance?

knowledge is not always encouraged with POMDPs

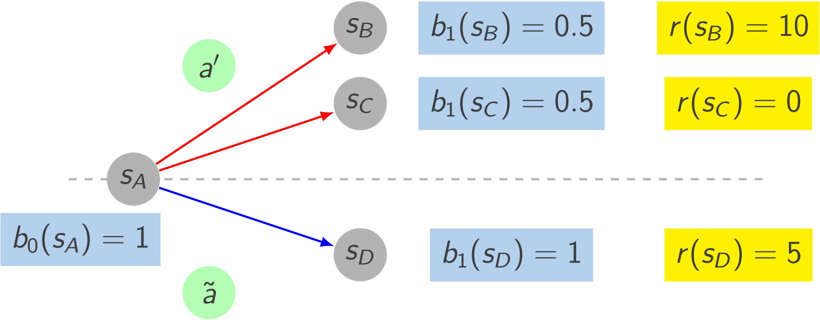
- initial belief deterministic $s_0 = s_A$.



ONERA
THE FRENCH AEROSPACE LAB

Why model ignorance?
knowledge is not always encouraged with POMDPs

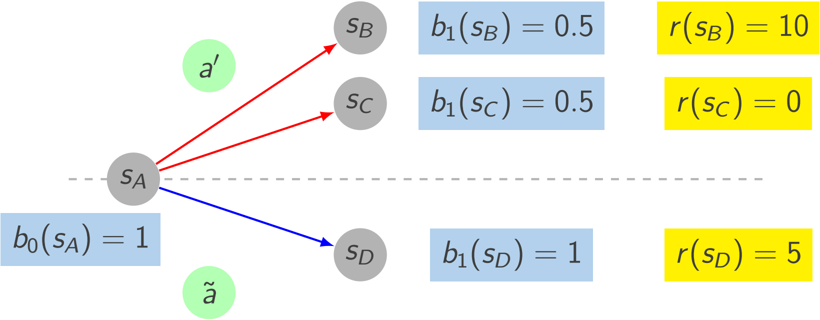
- initial belief deterministic $s_0 = s_A$.



Why model ignorance?

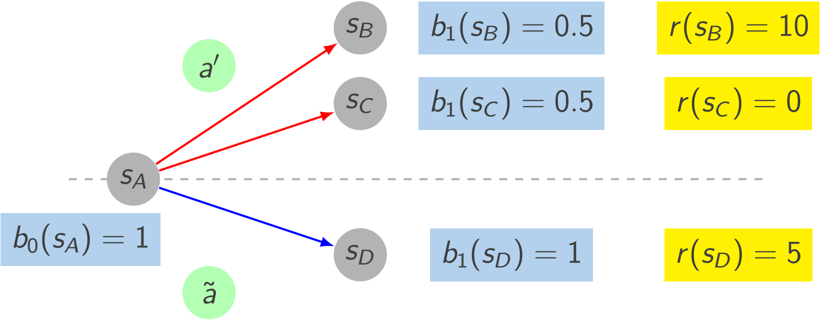
knowledge is not always encouraged with POMDPs

- initial belief deterministic $s_0 = s_A$.



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

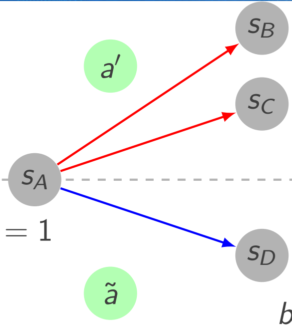
- initial belief deterministic $s_0 = s_A$.



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

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

the safe action is not prefered.



$b_0^\pi(s_A) = 1$

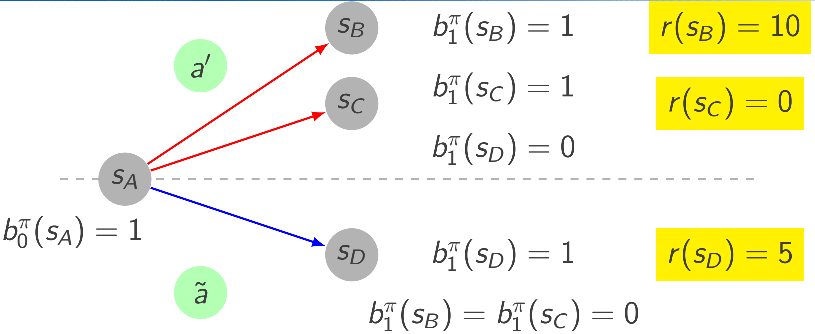
$b_1^\pi(s_B) = 1$ $r(s_B) = 10$

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

$b_1^\pi(s_D) = 0$

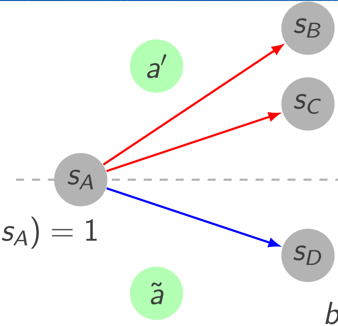
$b_1^\pi(s_D) = 1$ $r(s_D) = 5$

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



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

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



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

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

$b_1^\pi(s_D) = 0$

$b_0^\pi(s_A) = 1$

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

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

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

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

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

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

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

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

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

Choquet integral and rewards

pessimistic evaluation of the rewards – necessity measure

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

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

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

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

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

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

$$(2)$$

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

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

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

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

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

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

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

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

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

MDP variables:

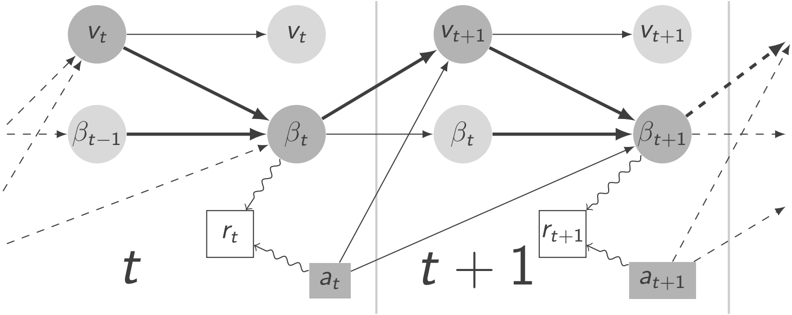
$$\tilde{\mathbb{S}} =$$

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

$$\times$$

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

resulting MDP in practice
final structured MDP



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

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

$\times$

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