Exploiting Imprecise Information Sources in Sequential Decision Making Problems under Uncertainty

N.Drougard

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

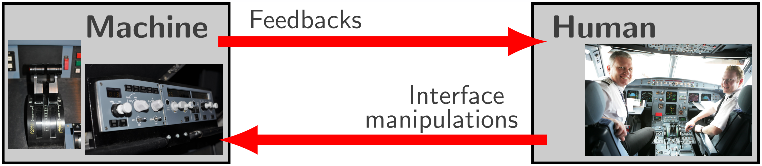
doctoral school: EDSYS **institution:** ISAE–SUPAERO

laboratory: ONERA–The French Aerospace Lab

retour sur innovation

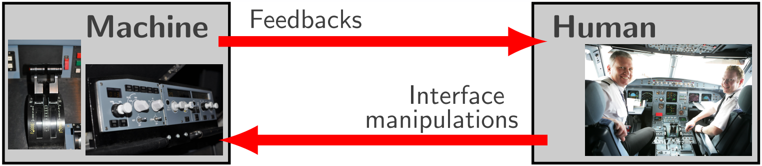
Example: Human-Machine Interaction (HMI)

joint work with **Sergio Pizziol** – Context



Example: Human-Machine Interaction (HMI)

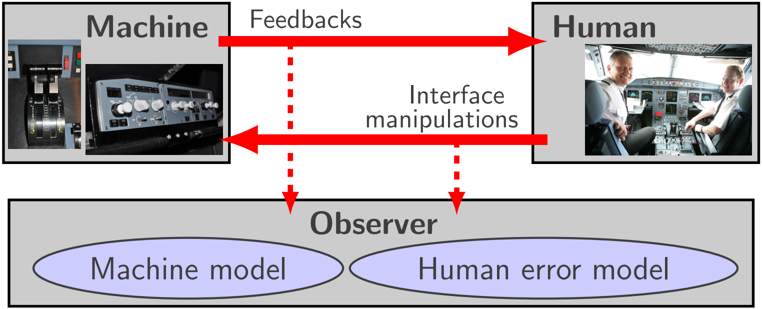
joint work with **Sergio Pizziol** – Context



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

Example: Human-Machine Interaction (HMI)

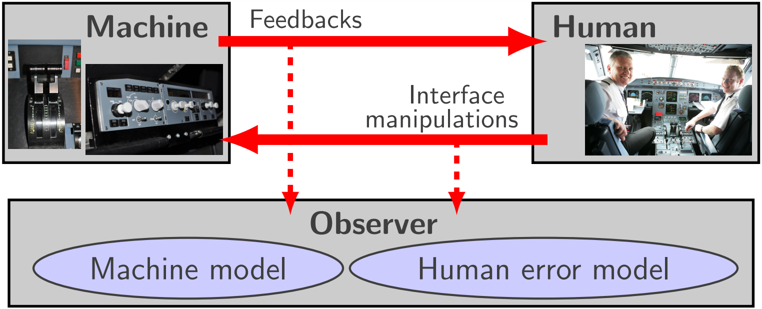
joint work with **Sergio Pizziol** – Context



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 $A \rightarrow B$

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

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

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

Example: Human-Machine Interaction (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" $\rightarrow$ $= \frac{1}{3}$

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

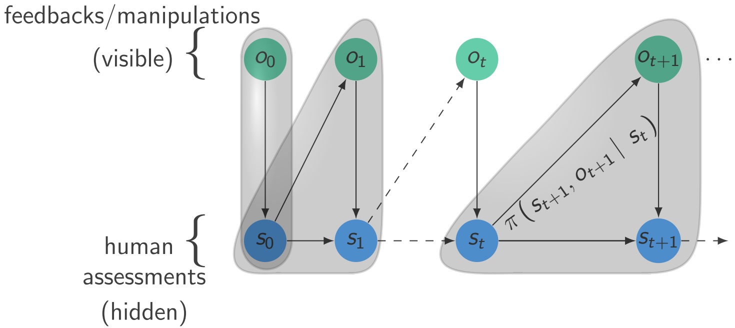
"manipulation o'_m is normal under s_A" $\rightarrow \pi\left(s'_B, o'_m \mid s_A\right) = 1$

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

- impossible cases: possibility degree 0

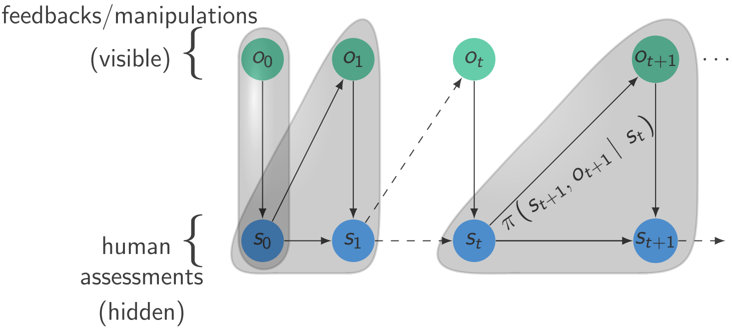
Qualitative Possibilistic Hidden Markov Process:

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:

maximizing

$$\min_{t=0}^{\#\delta} \min \left\{ \pi\left(s' \middle| s, \delta_t(s) \right), \Psi(s) \right\}$$

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

Indefinite Horizon

criterion, DP scheme, optimal strategy

indefinite horizon criterion:

maximizing

$$\min_{t=0}^{\#\delta} \min \left\{ \pi\left(s' \middle| s, \delta_t(s) \right), \Psi(s) \right\}$$

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:

maximizing

$$\min_{t=0}^{\#\delta} \min \left\{ \pi\left(s' \middle| s, \delta_t(s) \right), \Psi(s) \right\}$$

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

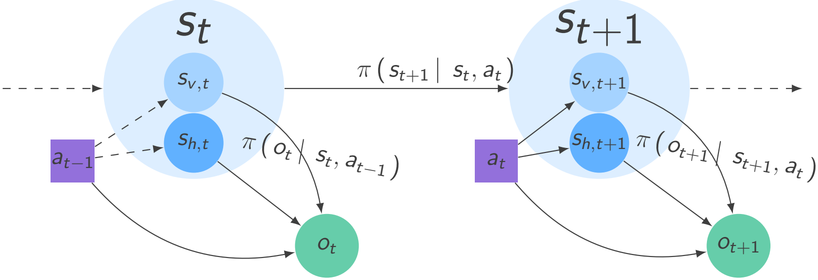
Dynamic Programming scheme: $\#$ iterations $< \#\mathcal{S}$

- 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:



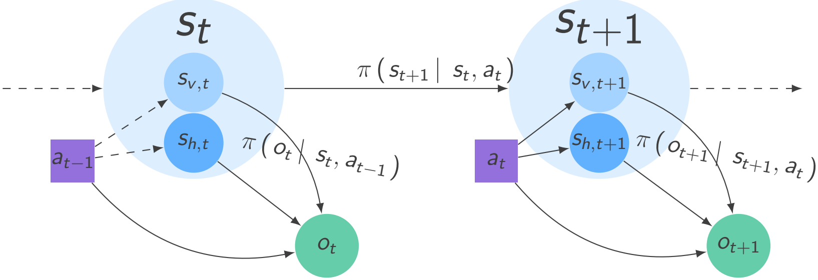
Mixed-Observability: system state $s \in \mathcal{S} = \mathcal{S}_v \times \mathcal{S}_h$

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

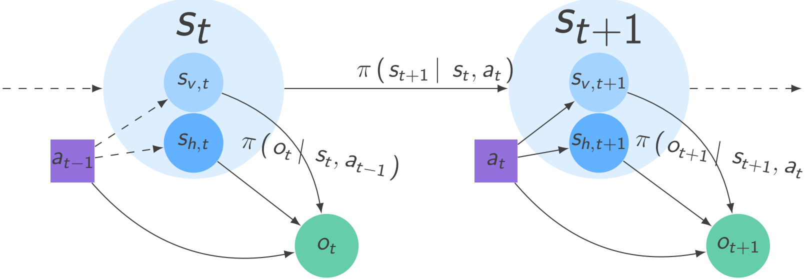
i.e. state s = visible component s_v & hidden component s_h

- 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 $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$

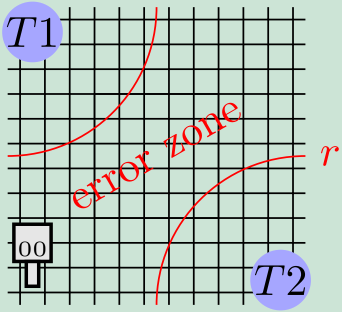
  $\#\mathcal{X} \sim \#\mathcal{S}_v \cdot \#\mathcal{L}^{\#\mathcal{S}_h} \ll \#\Pi_{\mathcal{L}}^{\mathcal{S}}$

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 $T1$ & $T2$

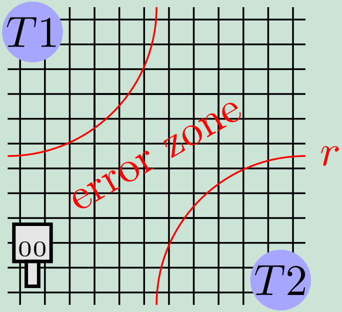


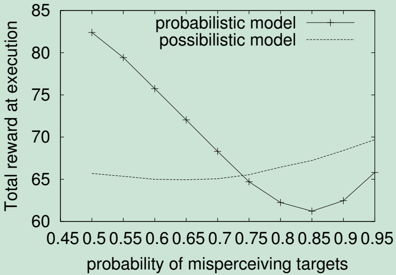
π-MOMDP for robotics with imprecise probabilities

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

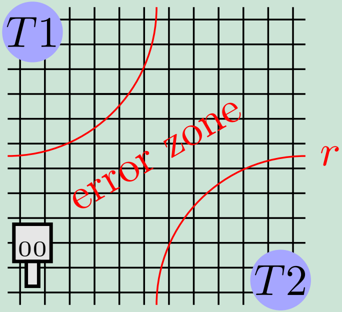
π-MOMDP for robotics with imprecise probabilities

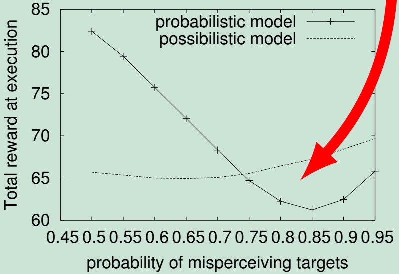
simulations with machine vision behavior imprecisely known

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

probabilistic model inappropriate with too imprecise probabilities

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



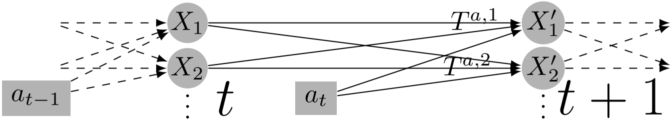


in reality, misperception probability in the error zone: $P_{bad} > \frac{1}{2}$

Factored π-MOMDP and computations with ADDs

qualitative possibilistic models to reduce complexity

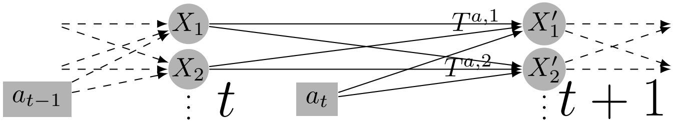
contribution (AAAI-14): factored π-MOMDP
$\Leftrightarrow$ state space $\mathcal{X} = \mathcal{S}_v \times \Pi_{\mathcal{L}}^{\mathcal{S}_h}$ = Boolean variables $(X_1, \dots, X_n)$
+ independence assumptions $\Leftarrow$ graphical model



Factored π-MOMDP and computations with ADDs

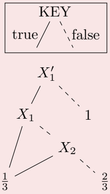
qualitative possibilistic models to reduce complexity

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

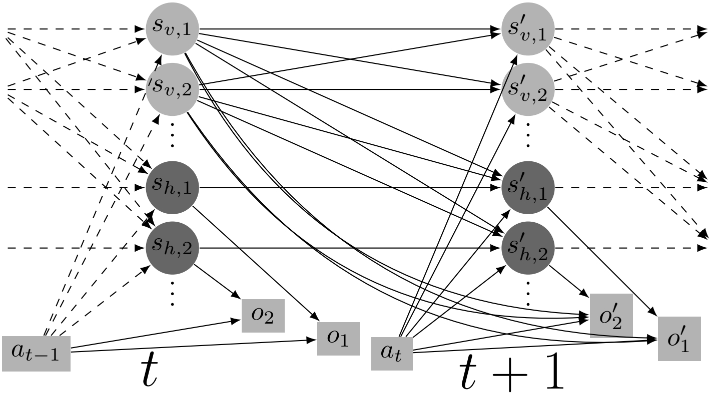
- factorization $\Rightarrow$ each DP steps divided into n stages
  $\rightarrow$ smaller ADDs $\Rightarrow$ **faster computations**
- computations on trees: *CU Decision Diagram Package*.

Simplifying computations with π-MOMDPs

Natural factorization: belief independence

contribution (AAAI-14):

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



Simplifying computations with π-MOMDPs

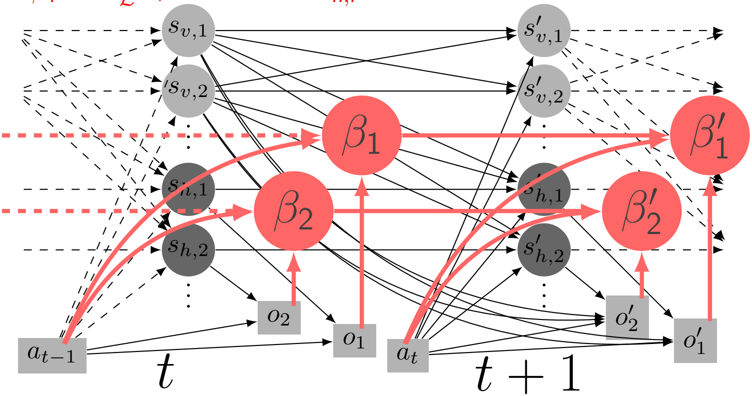
Natural factorization: belief independence

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 $s_{h,i}$

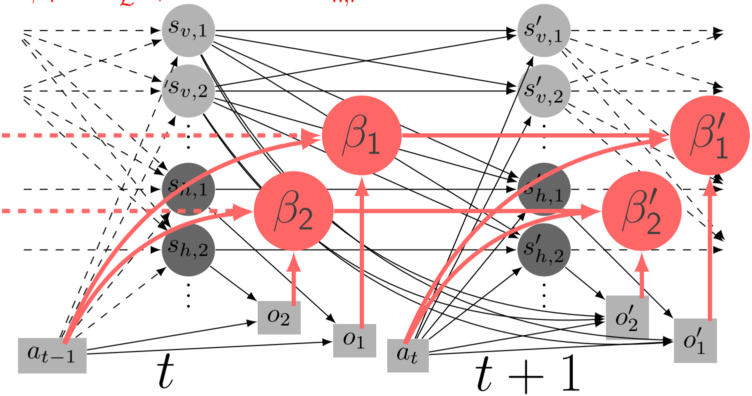


Simplifying computations with π-MOMDPs

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

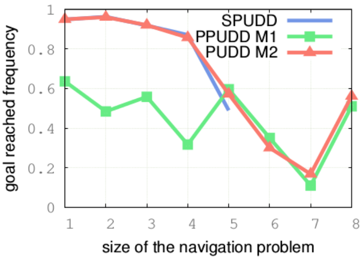
PPUDD vs SPUDD (*Hoey et al.*, 1999)

Navigation benchmark: reach a goal – spots with accident risk

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

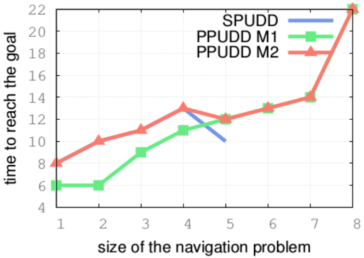
Performances, function of the problem index

reached goal frequency



the higher the better

\# steps to reach the goal

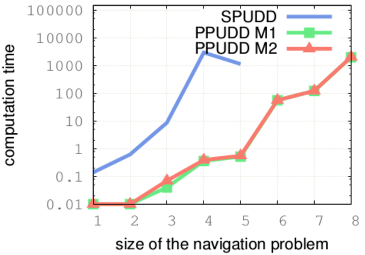


the lower the better

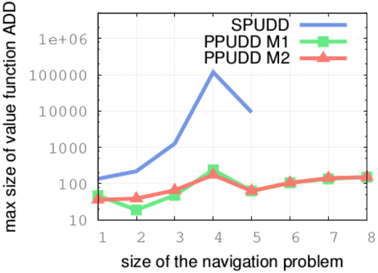
Simplify computations with π-MOMDPs

Experiments – perfect sensing: Navigation problem

computation time



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

Simplify computations with π-MOMDPs

Experiments – imperfect sensing: RockSample problem

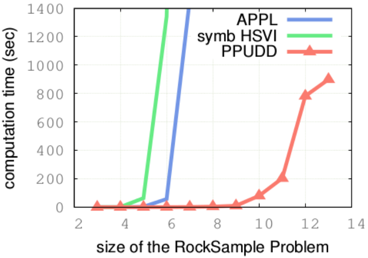
PPUDD vs APPL (*Kurniawati et al.*, 2008, solver MOMDP)
symbolic HSVI (*Sim et al.*, 2008, solver POMDP)
RockSample benchmark: recognize and sample “good” rocks

Simplify computations with π-MOMDPs

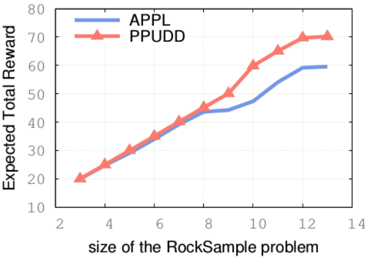
Experiments – imperfect sensing: RockSample problem

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:
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

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