

Contraintes et observabilité dans les systèmes de Markov décentralisés

Camille Besse

Département d'Informatique et de Génie Logiciel
<http://damas.ift.ulaval.ca>

Septembre 2010

Prologue

Cette présentation ...

Prologue

Cette présentation ...

- ne s'adresse *pas* à un public averti

Prologue

Cette présentation ...

- ne s'adresse *pas* à un public averti
- contient seulement **8.5%** de formules mathématiques

Prologue

Cette présentation ...

- ne s'adresse *pas* à un public averti
- contient seulement **8.5%** de formules mathématiques
- est composée à **82.4%** d'images et exemples

Prologue

Cette présentation ...

- ne s'adresse *pas* à un public averti
- contient seulement **8.5%** de formules mathématiques
- est composée à **82.4%** d'images et exemples
- se veut compréhensible par tous et toutes

Prologue

Cette présentation ...

- ne s'adresse *pas* à un public averti
 - contient seulement **8.5%** de formules mathématiques
 - est composée à **82.4%** d'images et exemples
 - se veut compréhensible par tous et toutes
- ➡ Donc si vous avez des questions ...

Prologue

Cette présentation ...

- ne s'adresse *pas* à un public averti
 - contient seulement **8.5%** de formules mathématiques
 - est composée à **82.4%** d'images et exemples
 - se veut compréhensible par tous et toutes
- ➡ Donc si vous avez des questions ...

N'hésitez pas et posez les !!!

Exemple introductif

Imaginons une ferme automatisée ...

Exemple introductif

Imaginons une ferme automatisée ...

- Où travaille un robot ...

Exemple introductif

Imaginons une ferme automatisée ...

- Où travaille un robot ...
- devant réaliser plusieurs tâches journalières :

Exemple introductif

Imaginons une ferme automatisée ...

- Où travaille un robot ...
- devant réaliser plusieurs tâches journalières :
 - Nourrir les poules
 - Traire les vaches
 - Arroser le potager
 - (etc ...)

Exemple introductif

Imaginons une ferme automatisée ...

- Où travaille un robot ...
- devant réaliser plusieurs tâches journalières :
 - Nourrir les poules
 - Traire les vaches
 - Arroser le potager
 - (etc ...)
- Ce robot n'est malheureusement pas parfait

Exemple introductif

Imaginons une ferme automatisée ...

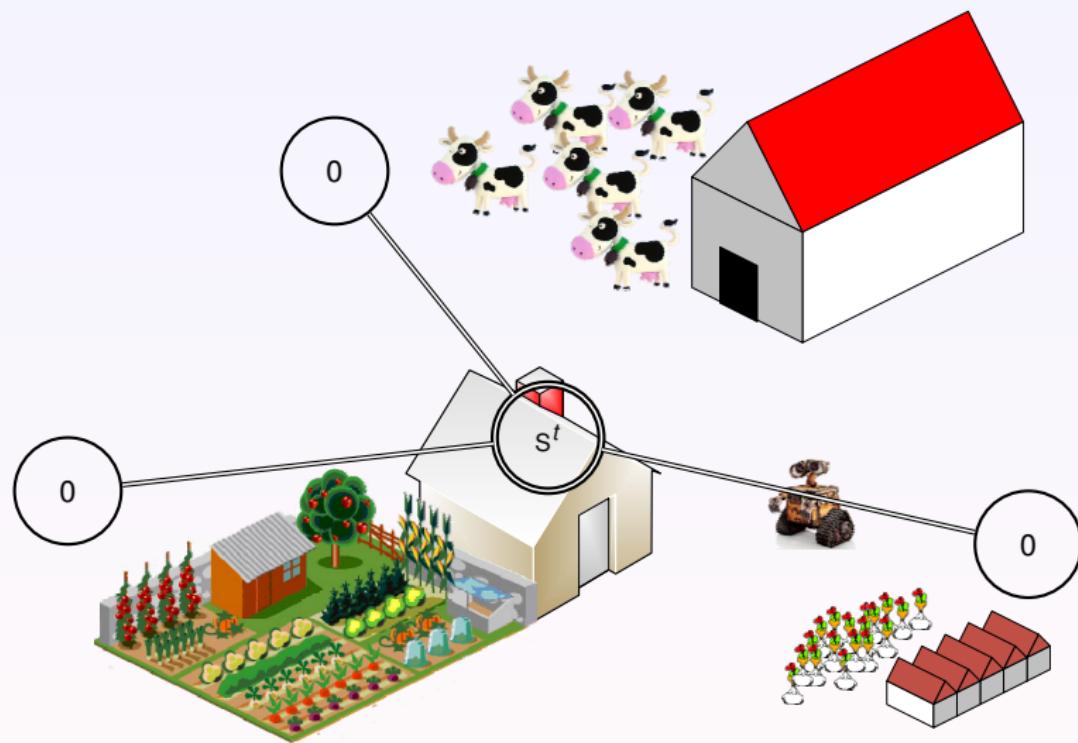
- Où travaille un robot ...
- devant réaliser plusieurs tâches journalières :
 - Nourrir les poules
 - Traire les vaches
 - Arroser le potager
 - (etc ...)
- Ce robot n'est malheureusement pas parfait
 - ➔ Les tâches qu'il réalise peuvent parfois prendre plus de temps
 - ➔ Il a en plus des chances d'échouer (ou de mal réaliser ces tâches)

Exemple introductif

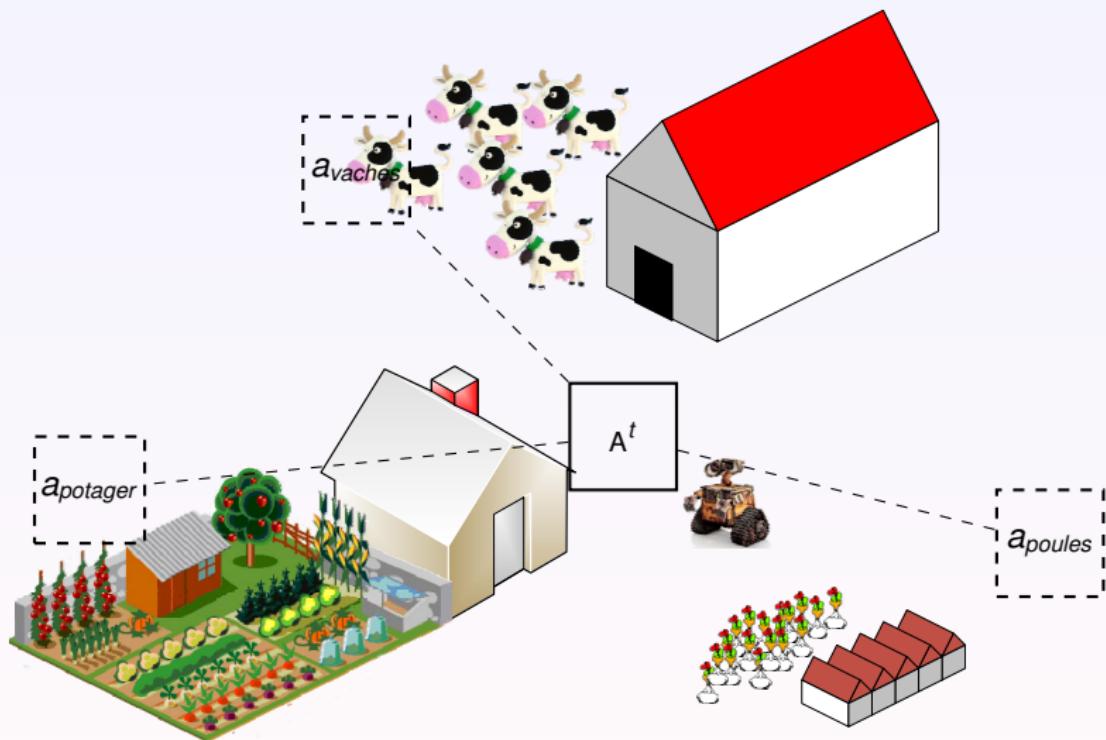
Imaginons une ferme automatisée ...

- Où travaille un robot ...
- devant réaliser plusieurs tâches journalières :
 - Nourrir les poules
 - Traire les vaches
 - Arroser le potager
 - (etc ...)
- Ce robot n'est malheureusement pas parfait
 - Les tâches qu'il réalise peuvent parfois prendre plus de temps
 - Il a en plus des chances d'échouer (ou de mal réaliser ces tâches)
- Comment modéliser ce problème de *planification* ?

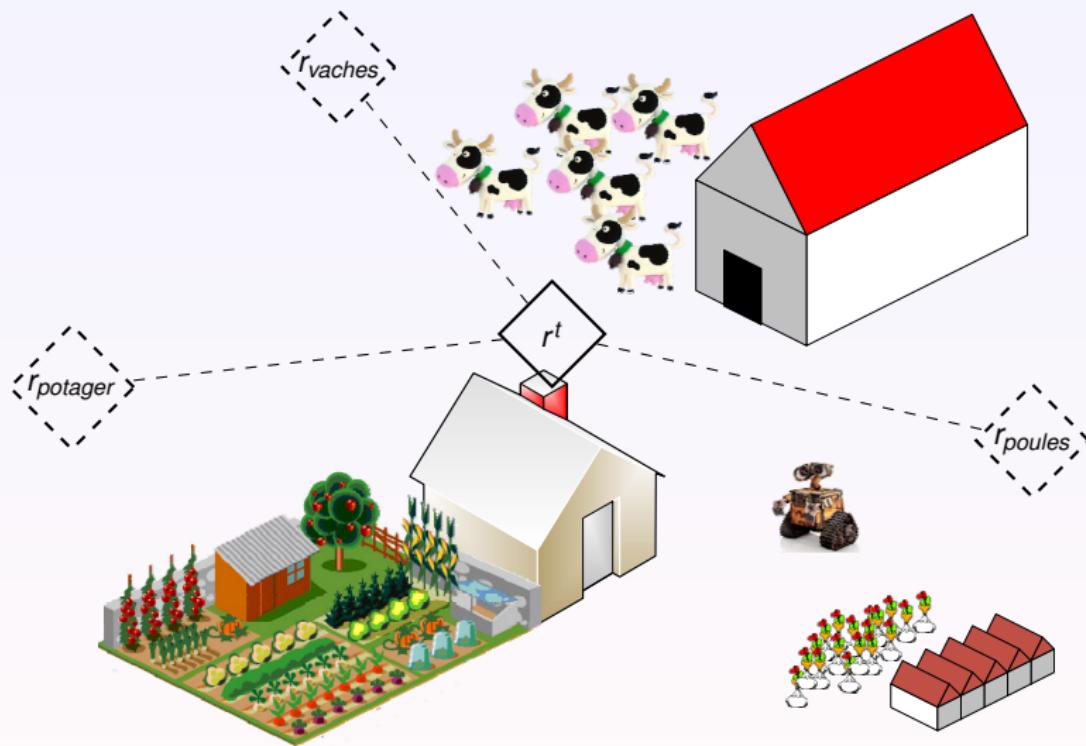
Exemple introductif



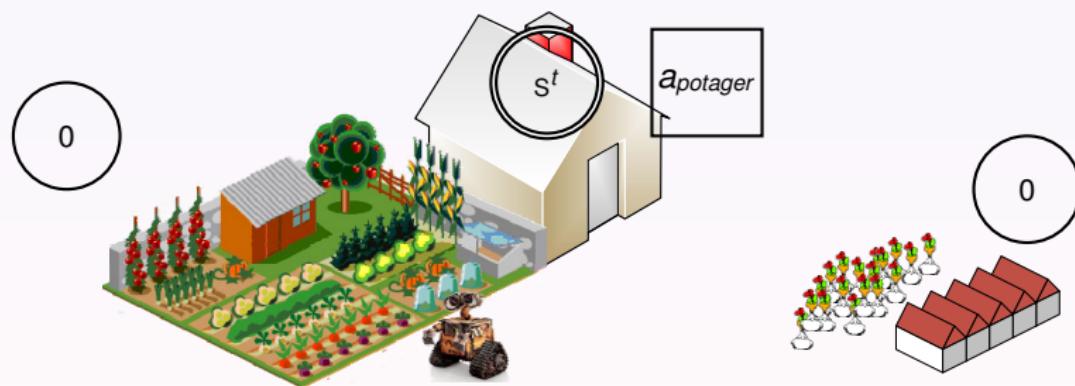
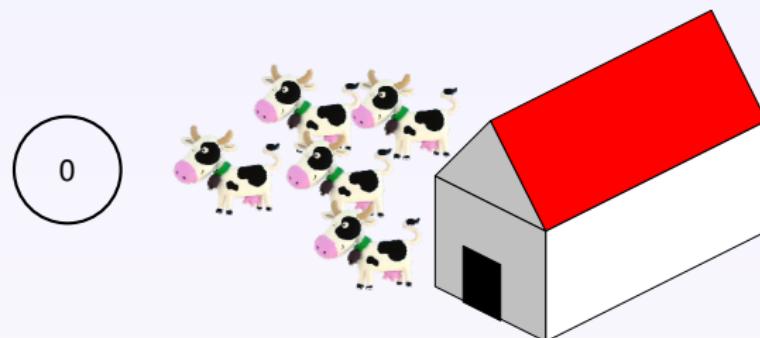
Exemple introductif



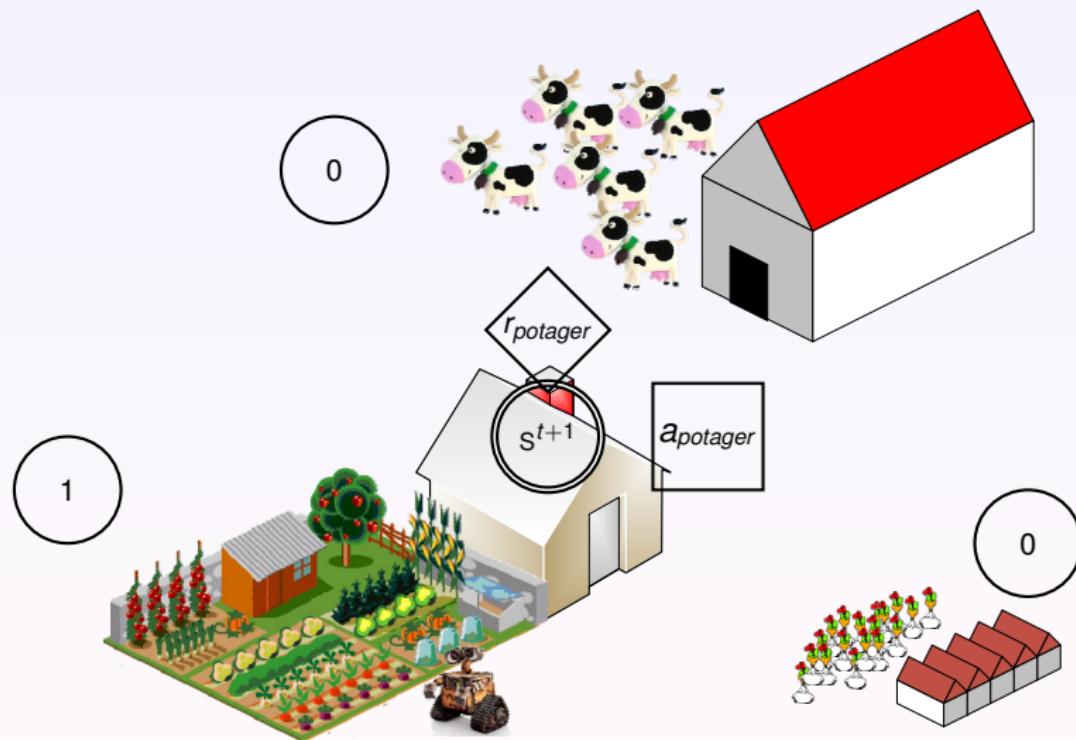
Exemple introductif



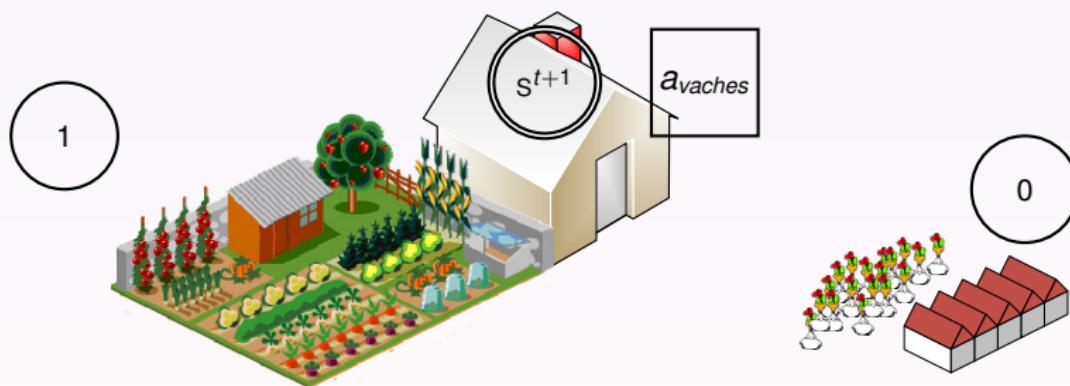
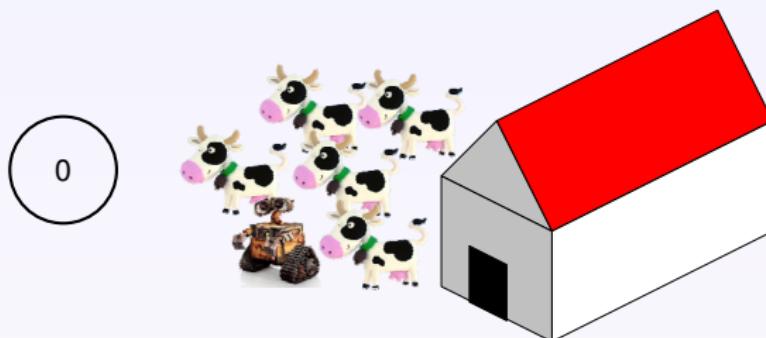
Exemple introductif



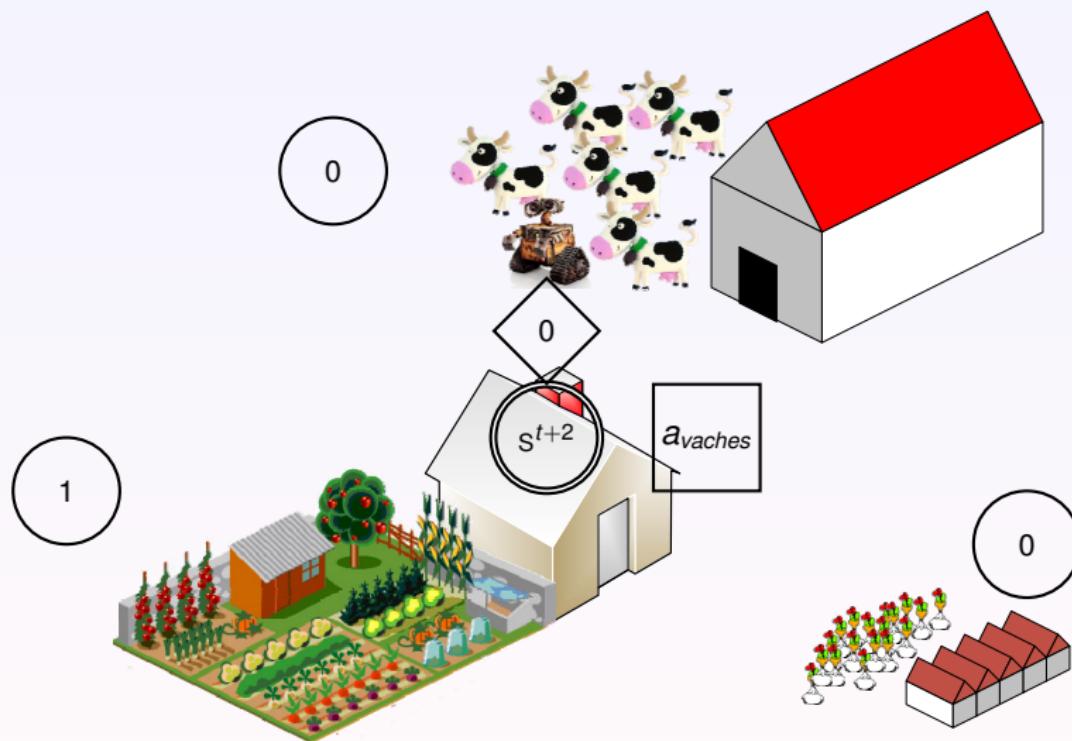
Exemple introductif



Exemple introductif

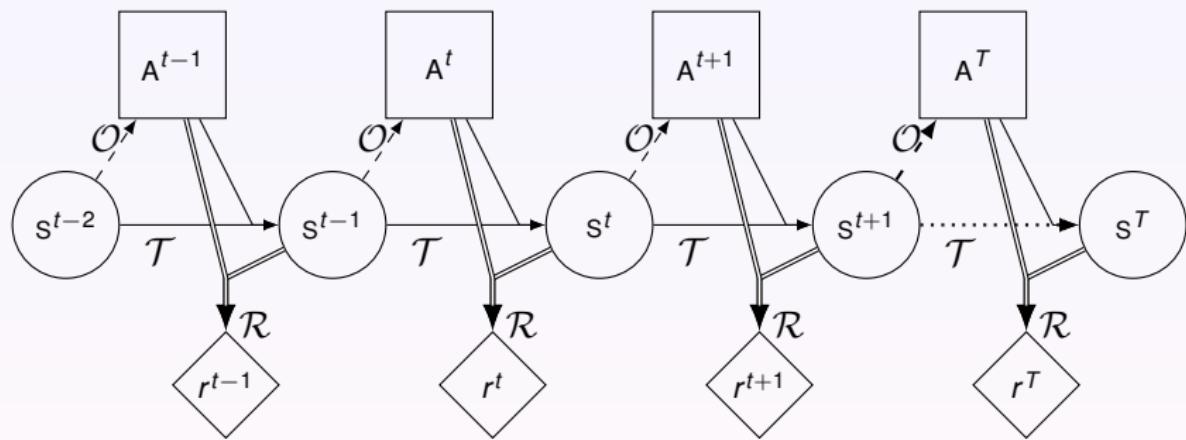


Exemple introductif



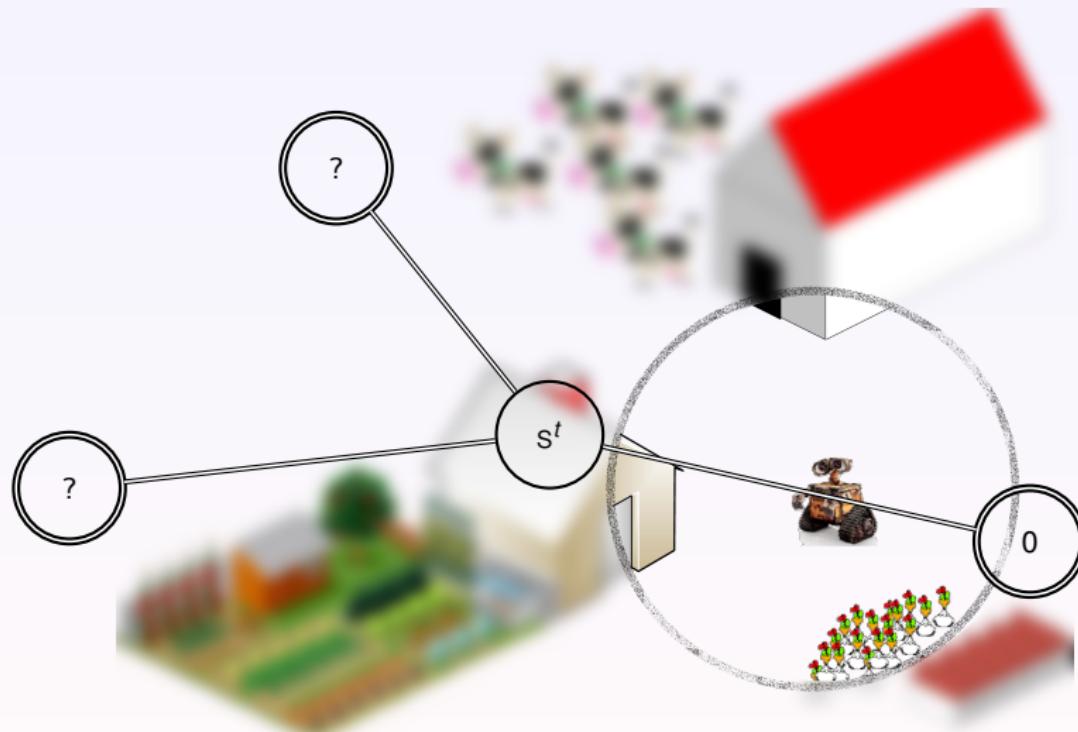
Exemple introductif

Représentation graphique d'un Processus décisionnel de Markov



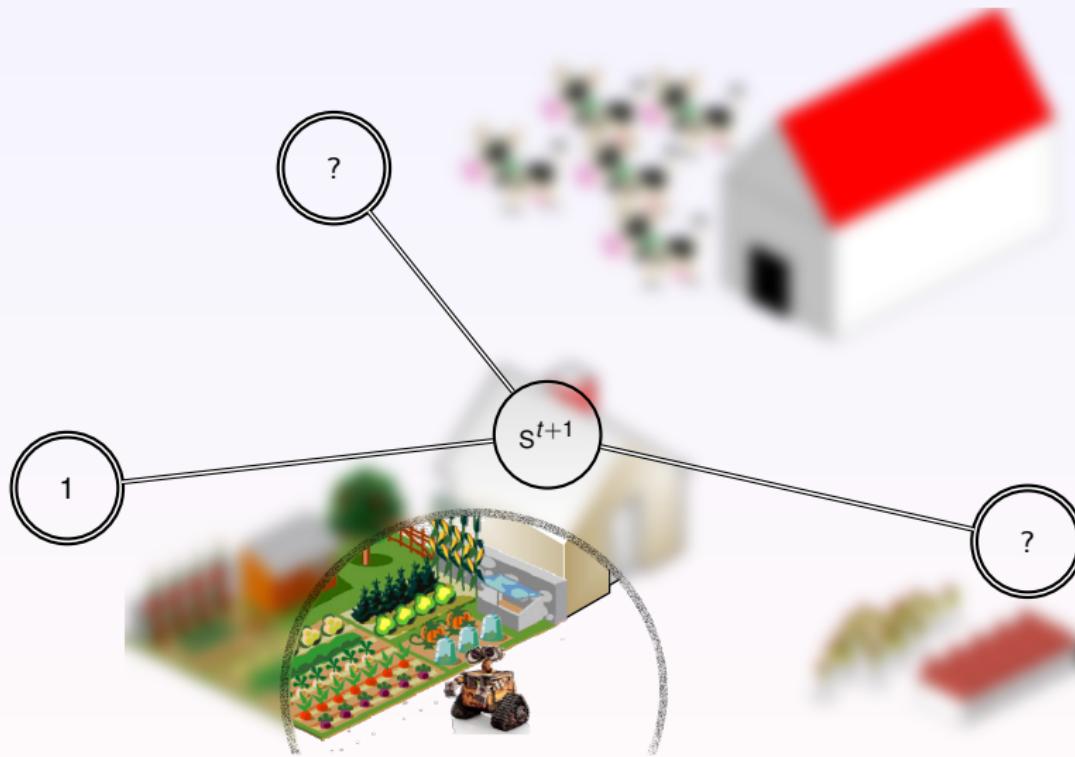
Exemple introductif

Et en partiellement observable ?



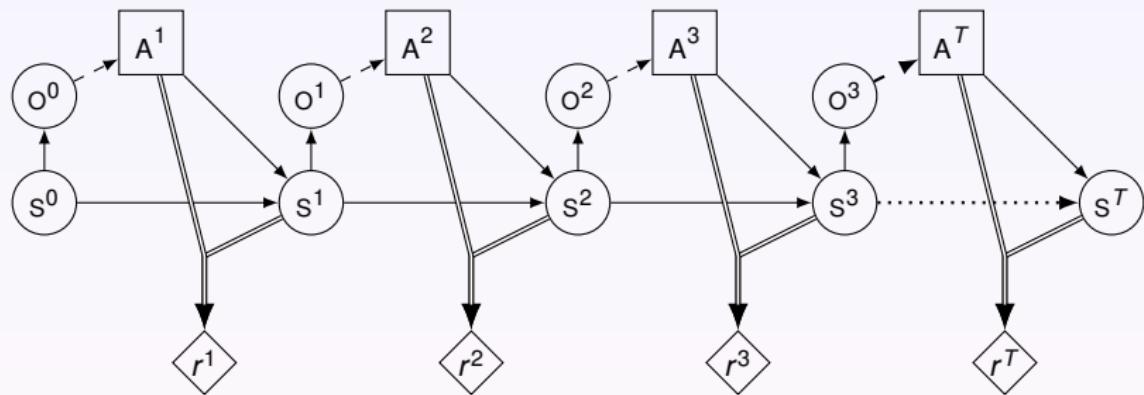
Exemple introductif

Et en partiellement observable ?



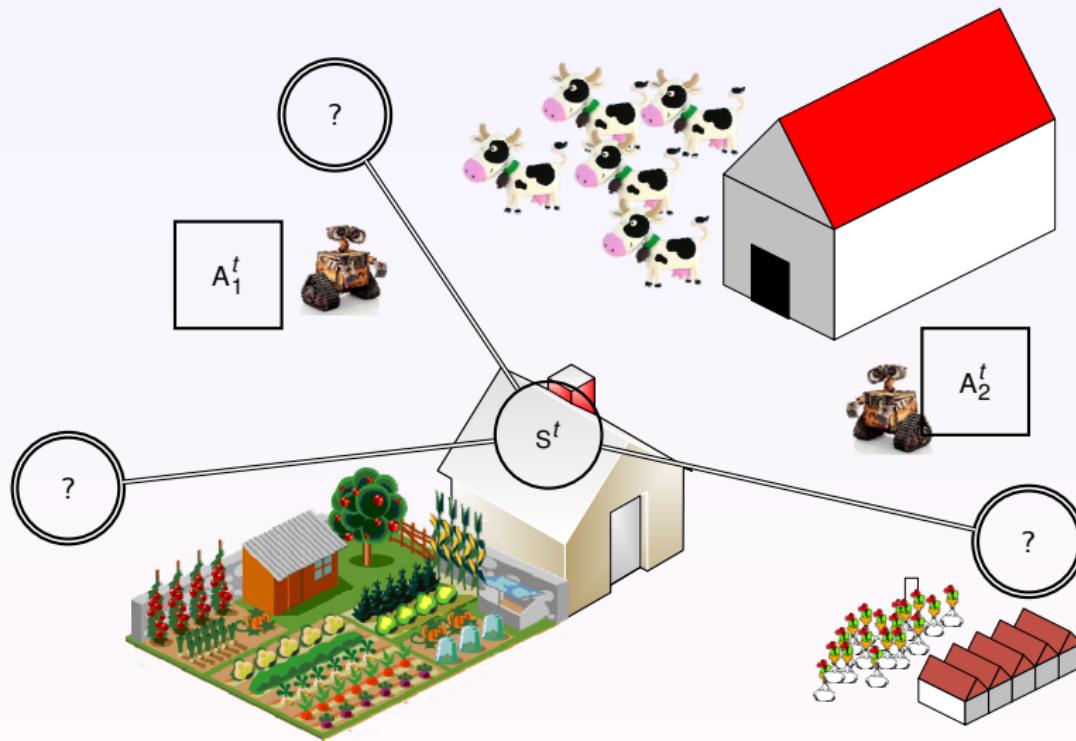
Exemple introductif

Représentation graphique en partiellement observable



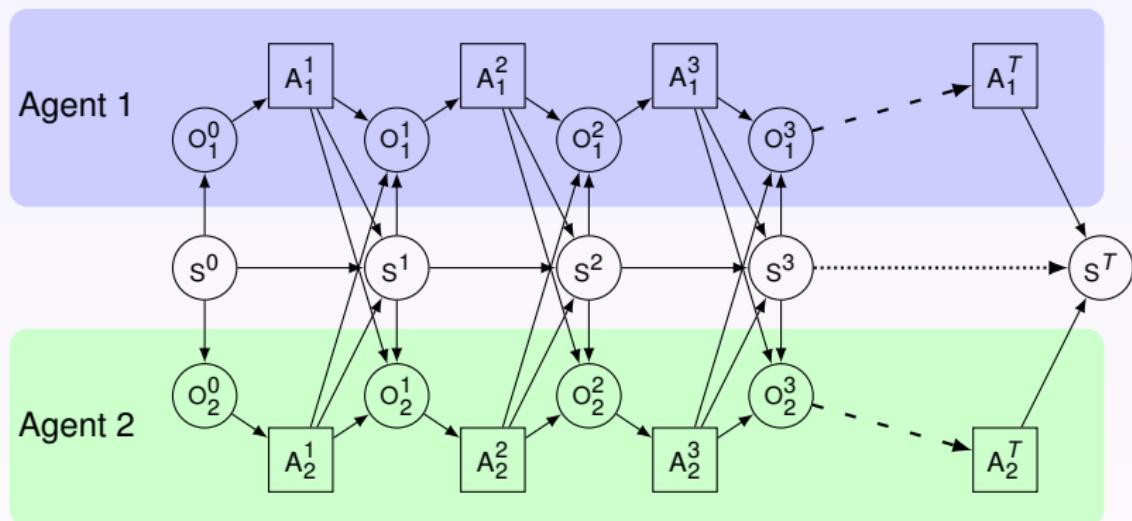
Exemple introductif

Et en multiagent ?



Exemple introductif

Représentation graphique en multiagent



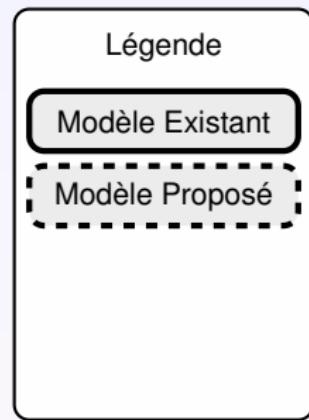
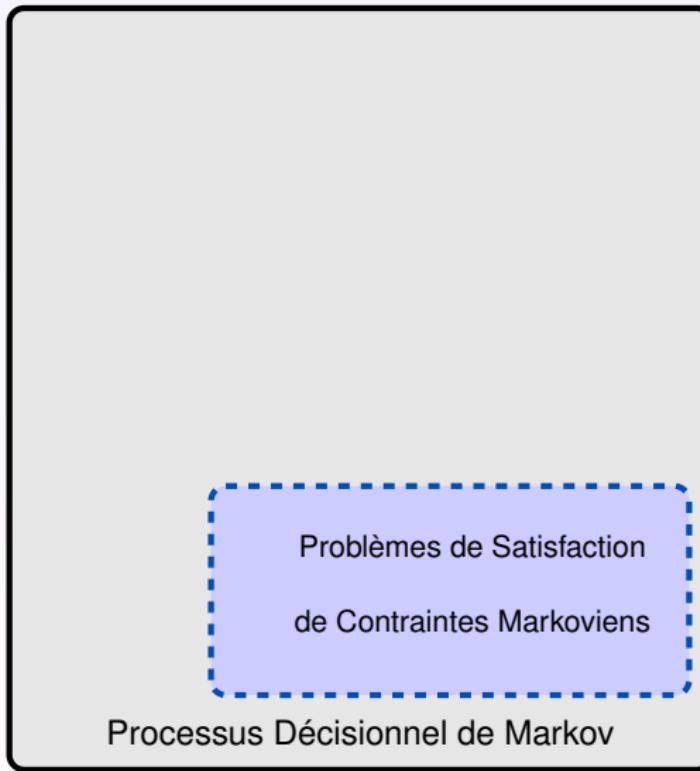
Contributions et plan de la présentation

Processus Décisionnel de Markov

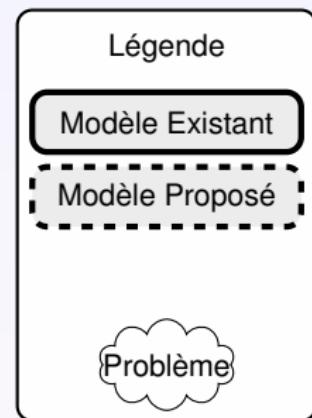
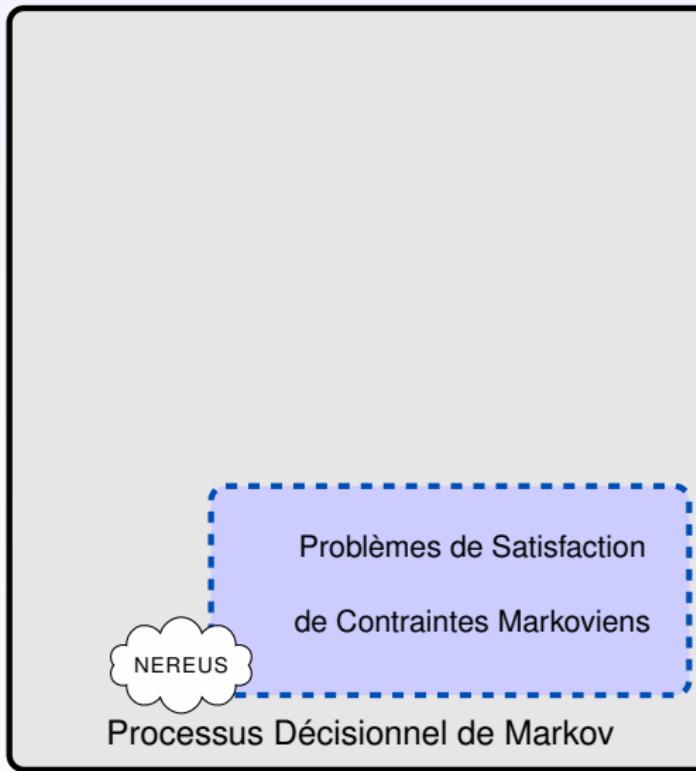
Légende

Modèle Existant

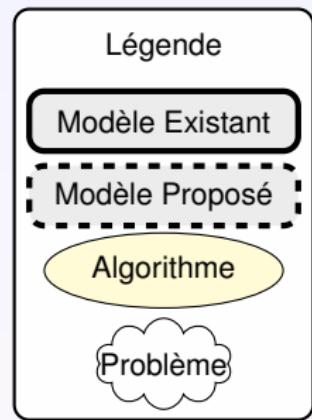
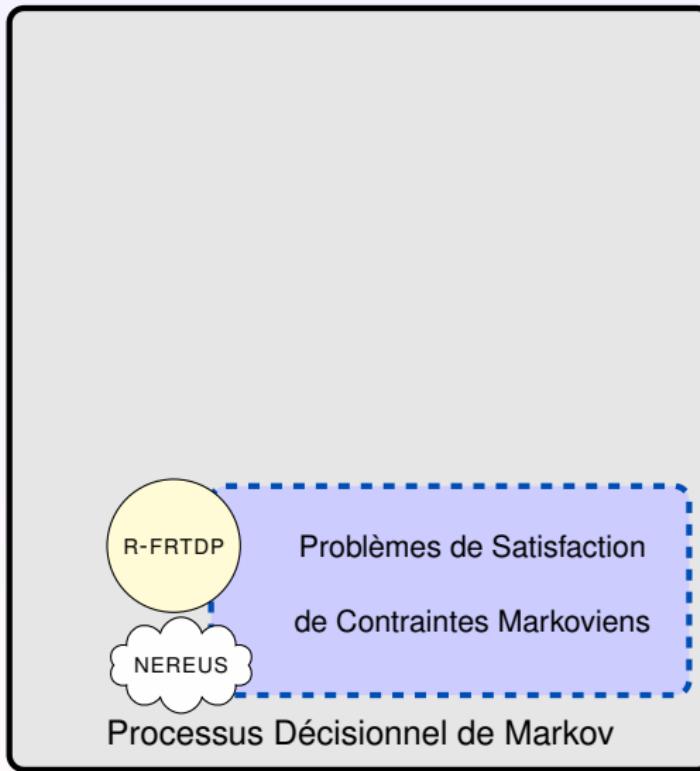
Contributions et plan de la présentation



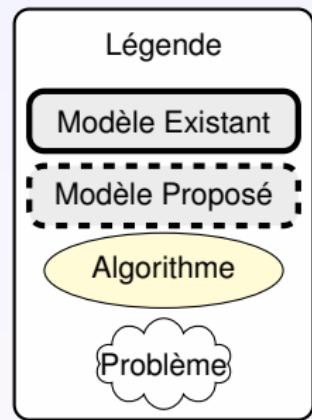
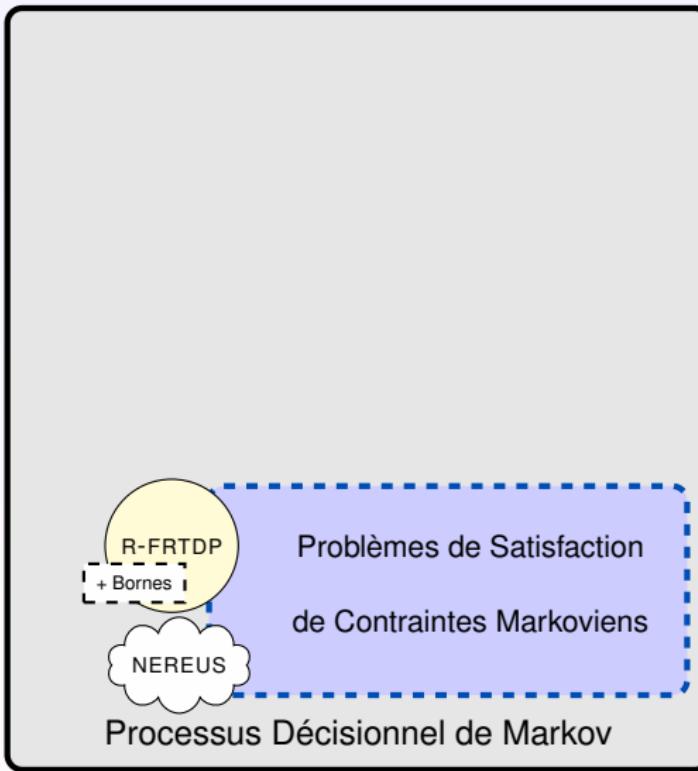
Contributions et plan de la présentation



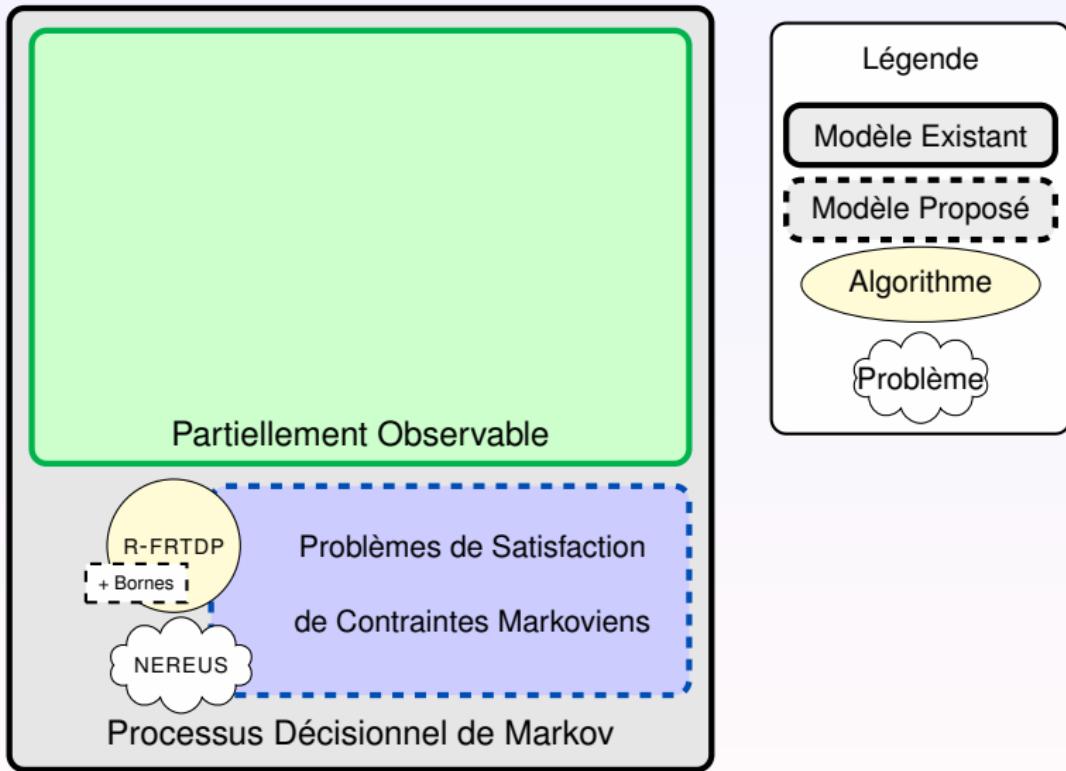
Contributions et plan de la présentation



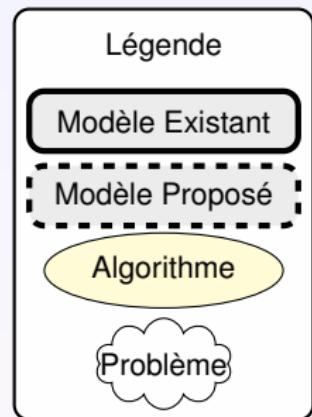
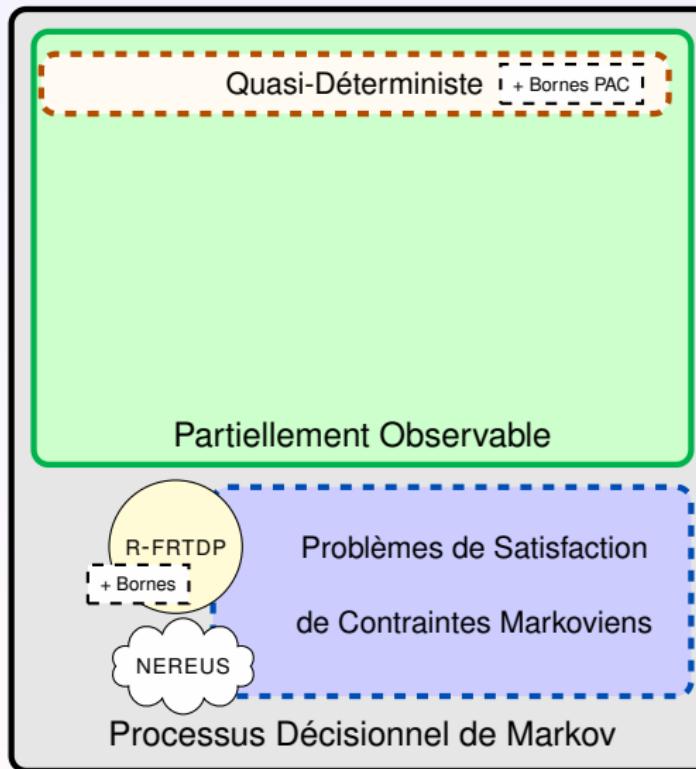
Contributions et plan de la présentation



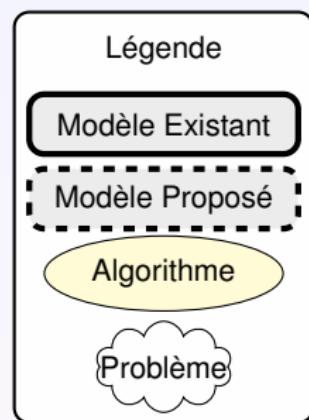
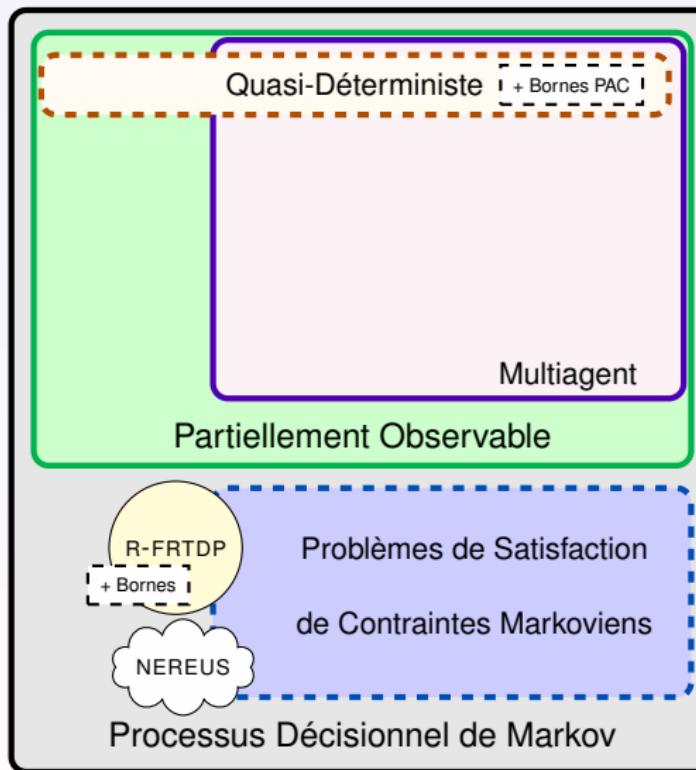
Contributions et plan de la présentation



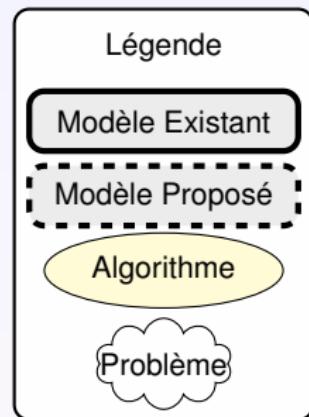
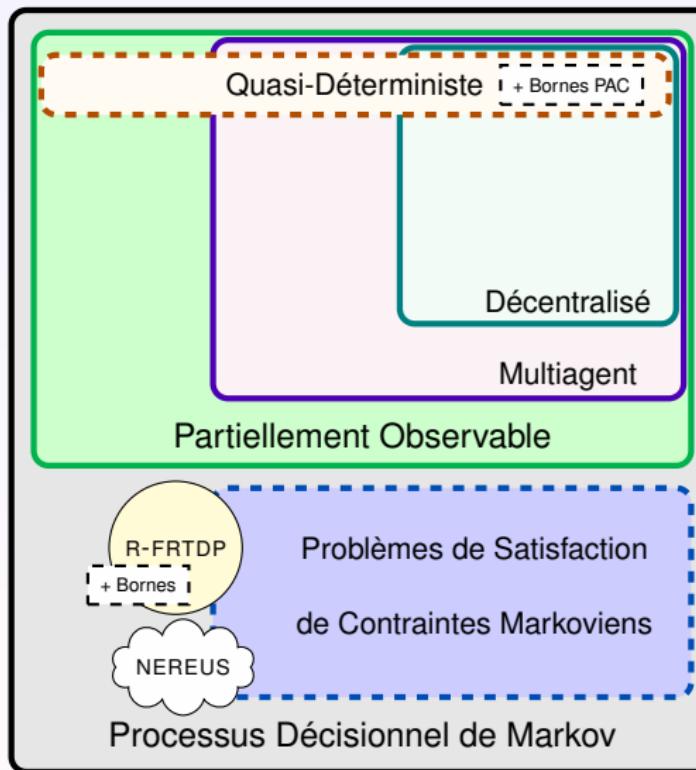
Contributions et plan de la présentation



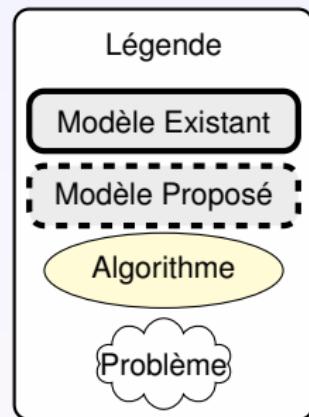
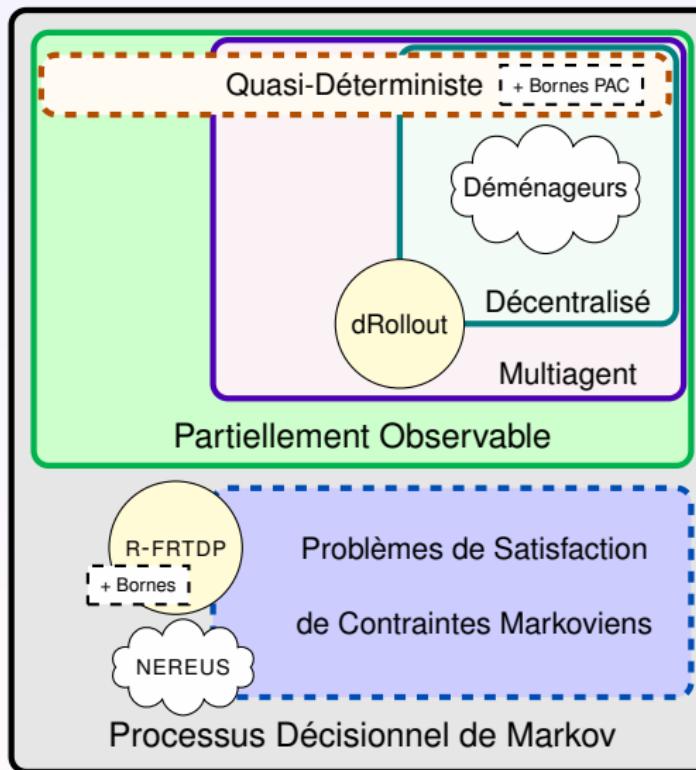
Contributions et plan de la présentation



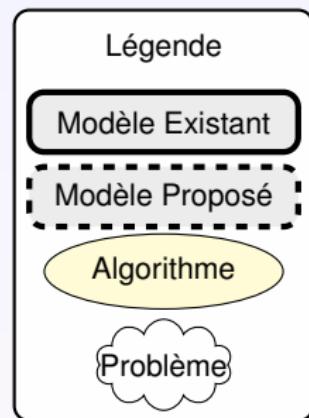
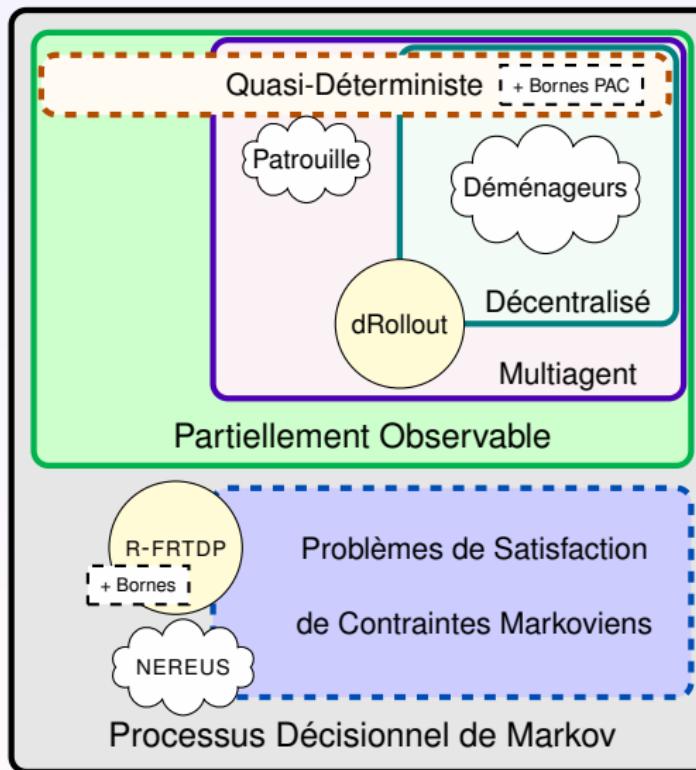
Contributions et plan de la présentation



Contributions et plan de la présentation

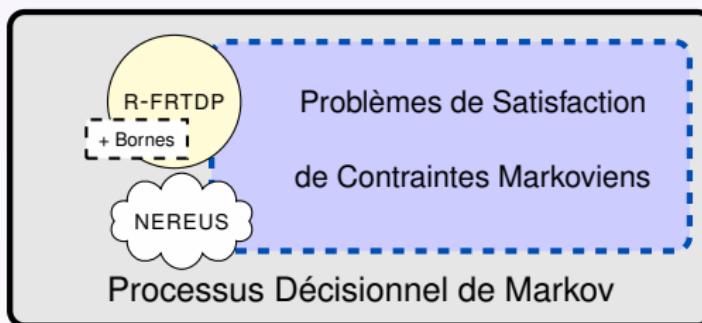


Contributions et plan de la présentation

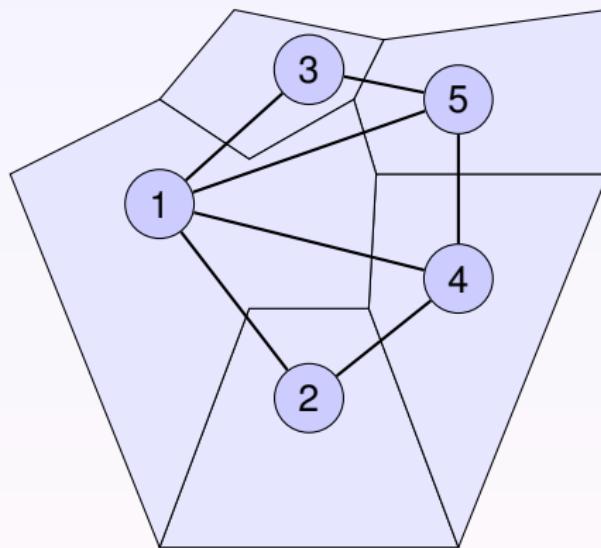


Plan

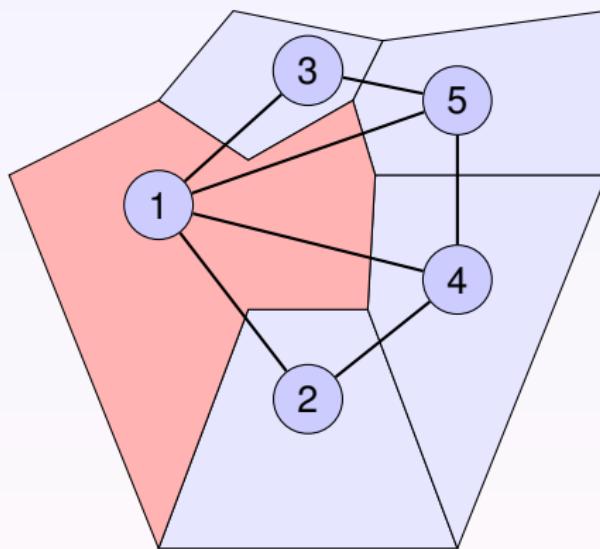
1^{ere} Partie : Contraintes et Modèles de Markov



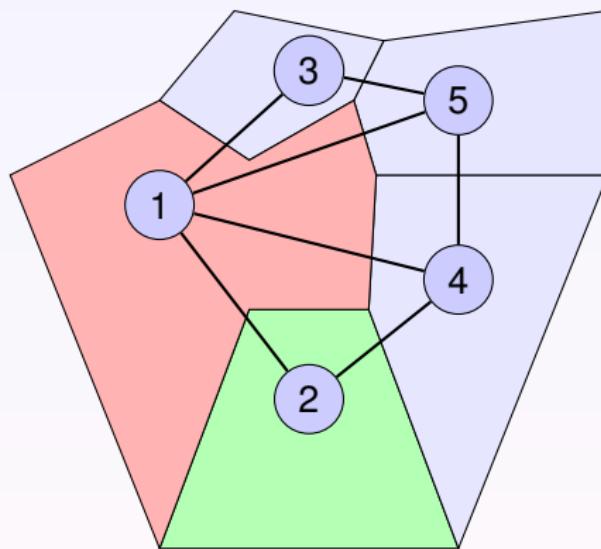
Problèmes de satisfaction de contraintes



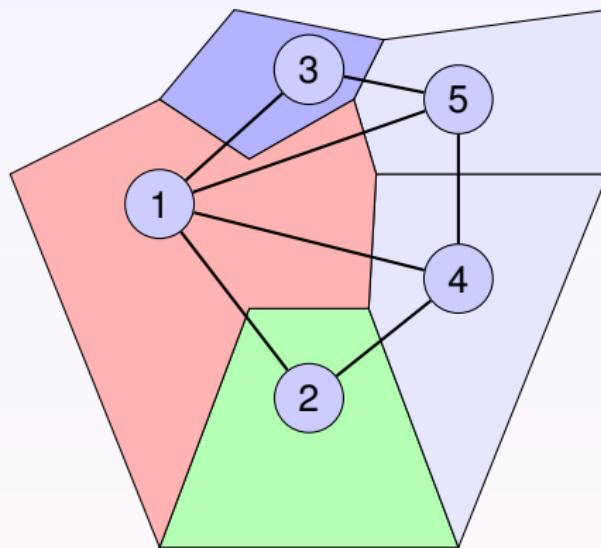
Problèmes de satisfaction de contraintes



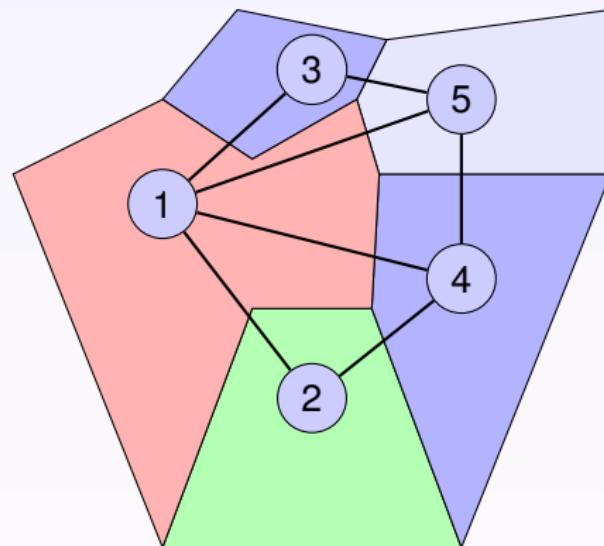
Problèmes de satisfaction de contraintes



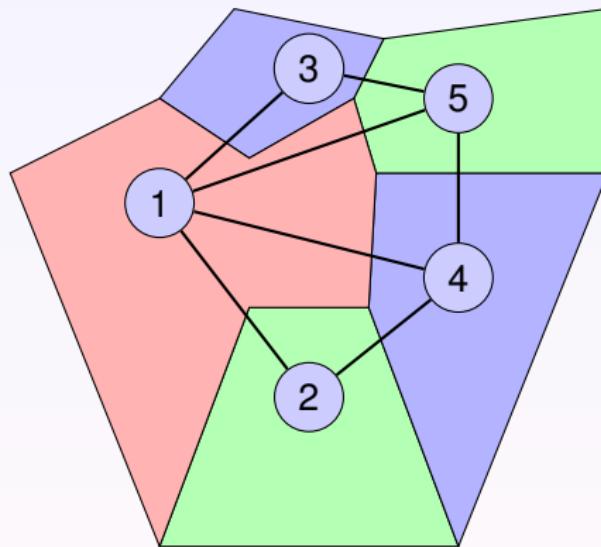
Problèmes de satisfaction de contraintes



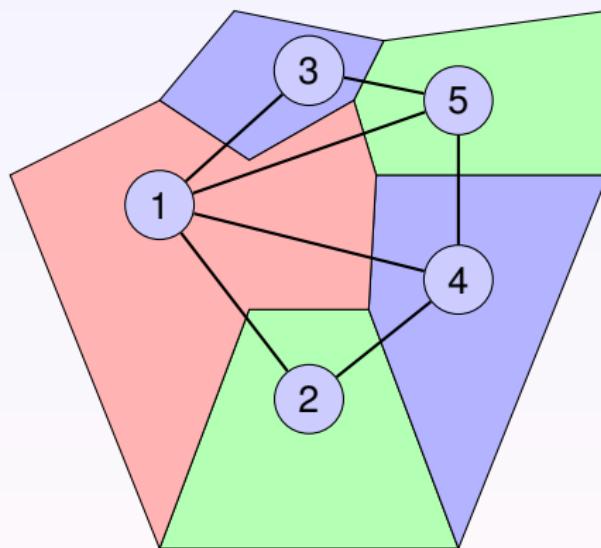
Problèmes de satisfaction de contraintes



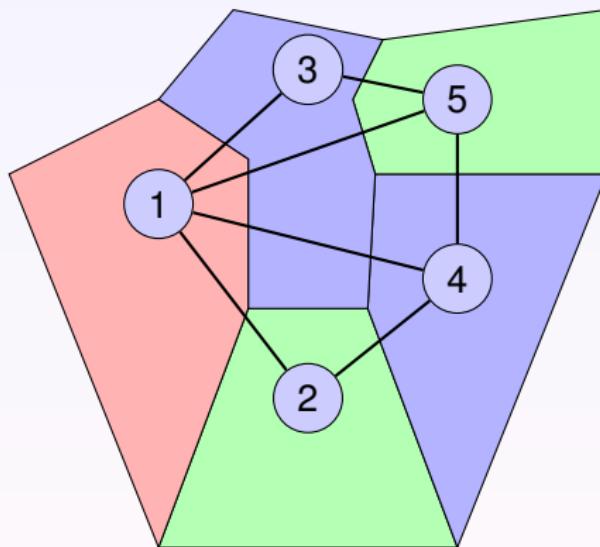
Problèmes de satisfaction de contraintes



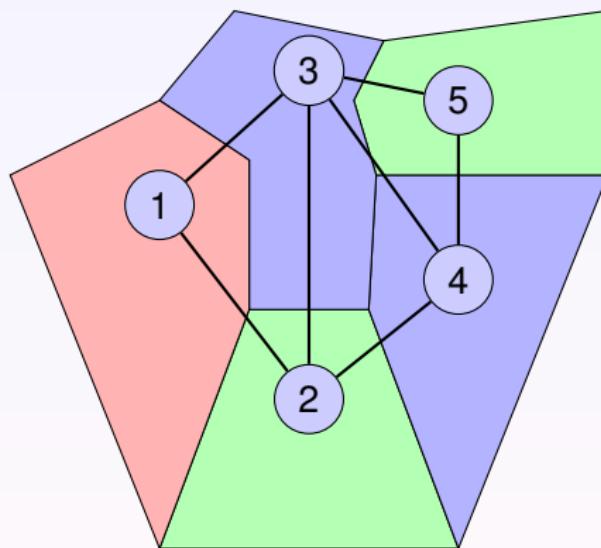
Problèmes de satisfaction de contraintes dynamiques



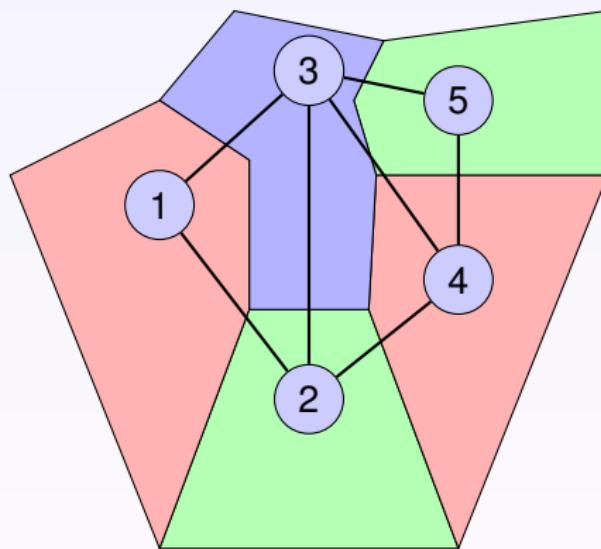
Problèmes de satisfaction de contraintes dynamiques



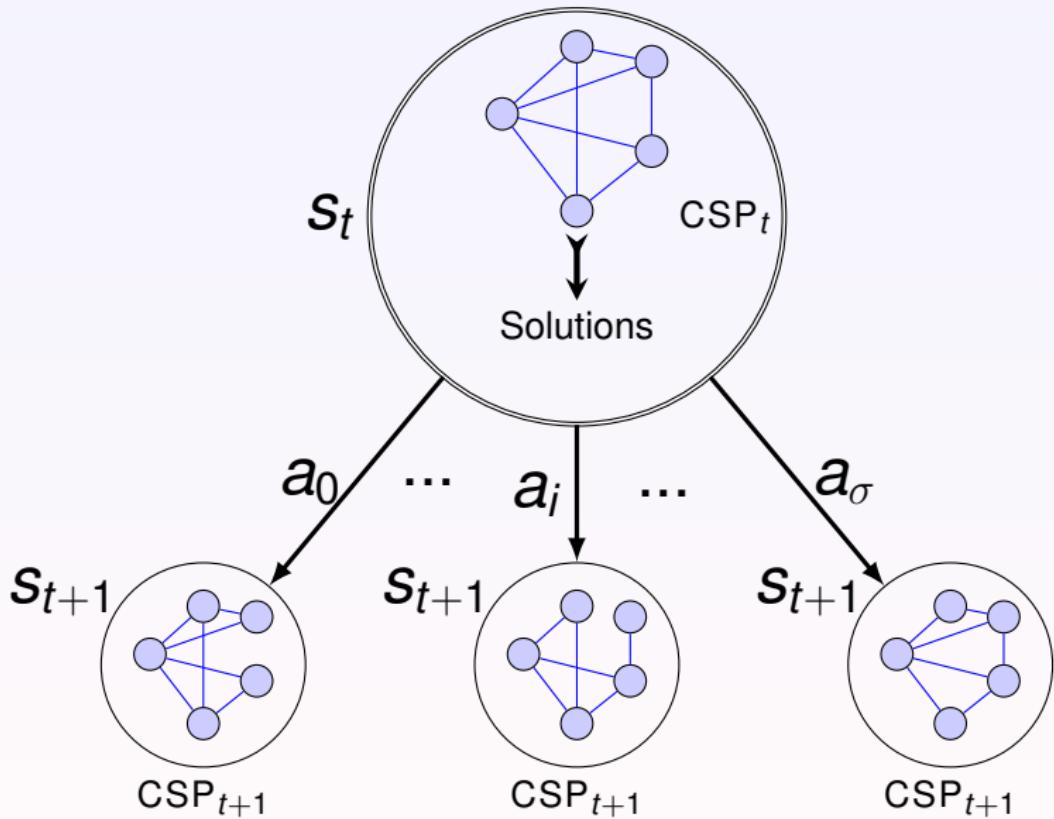
Problèmes de satisfaction de contraintes dynamiques



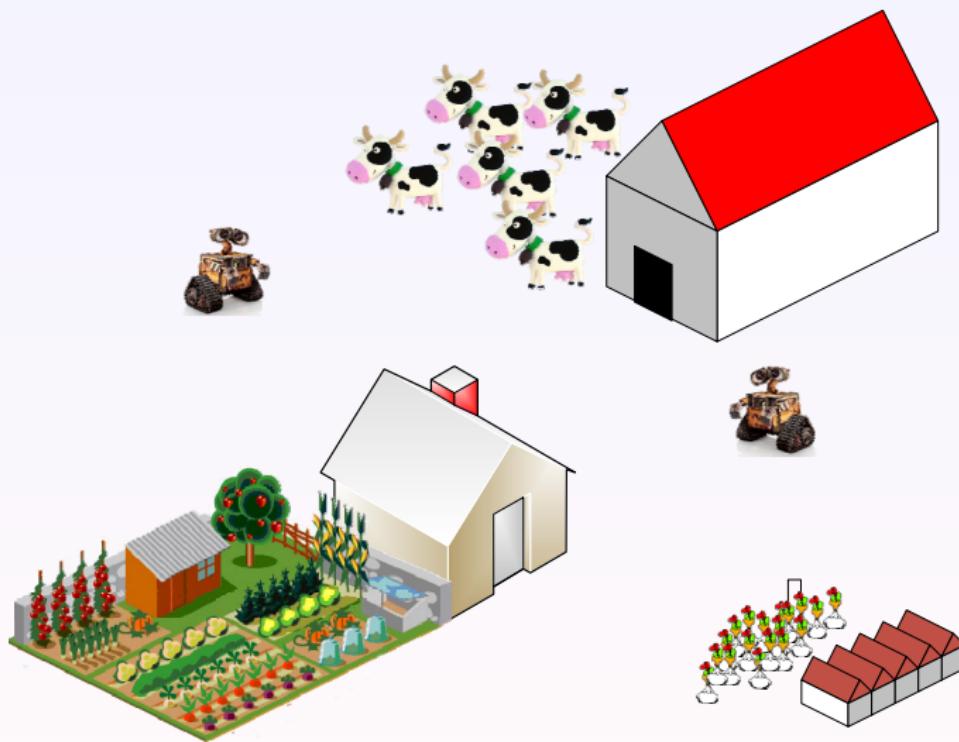
Problèmes de satisfaction de contraintes dynamiques



Problèmes de Markov avec satisfaction de contraintes

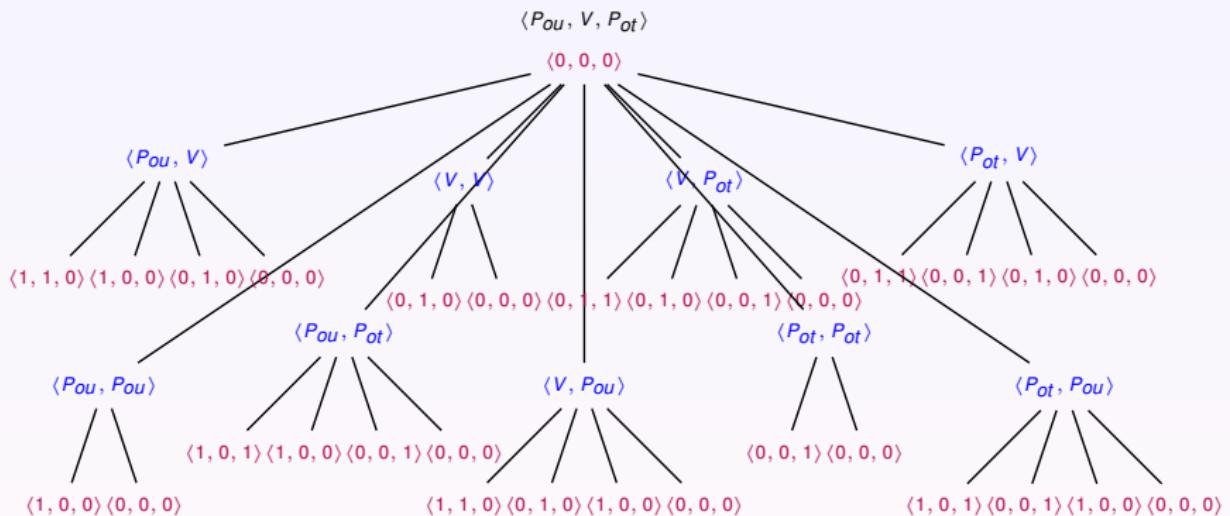


Exemple illustratif à la ferme



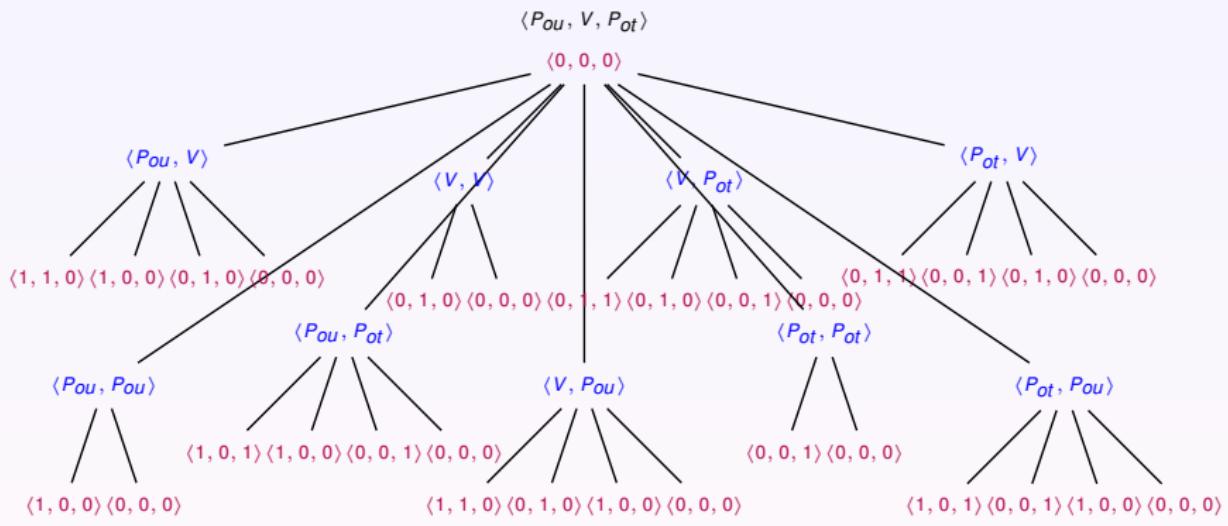
Exemple illustratif à la ferme

Arbre de résolution d'un MDP classique



Exemple illustratif à la ferme

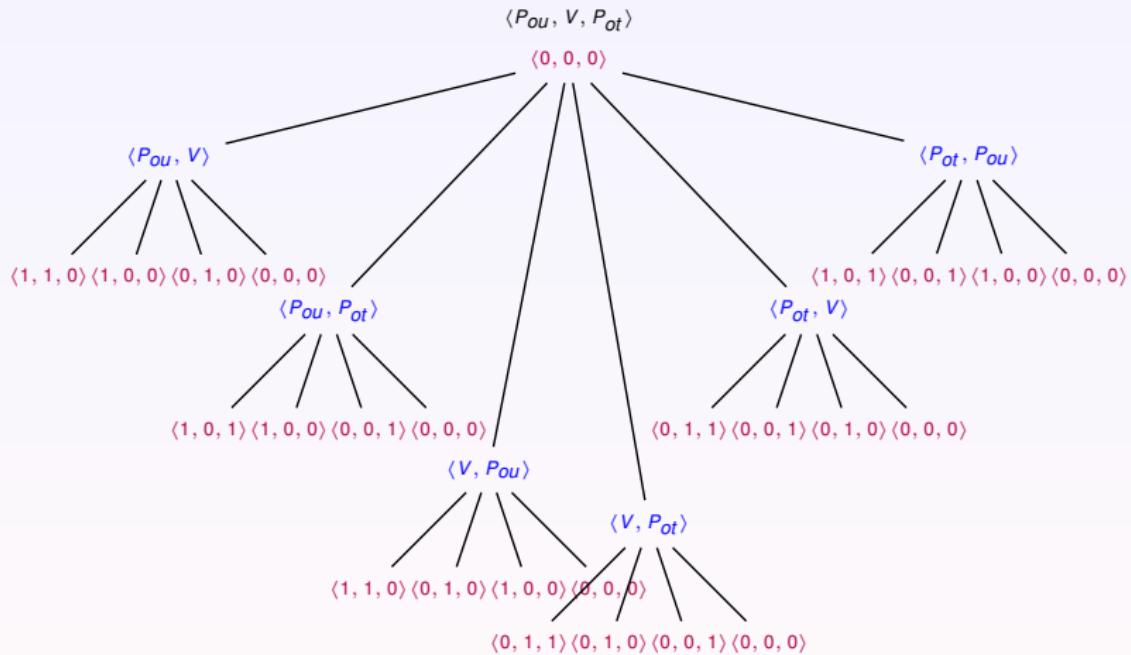
Arbre de résolution d'un MDP classique



7 états possibles selon les probabilités
pour 9 actions jointes

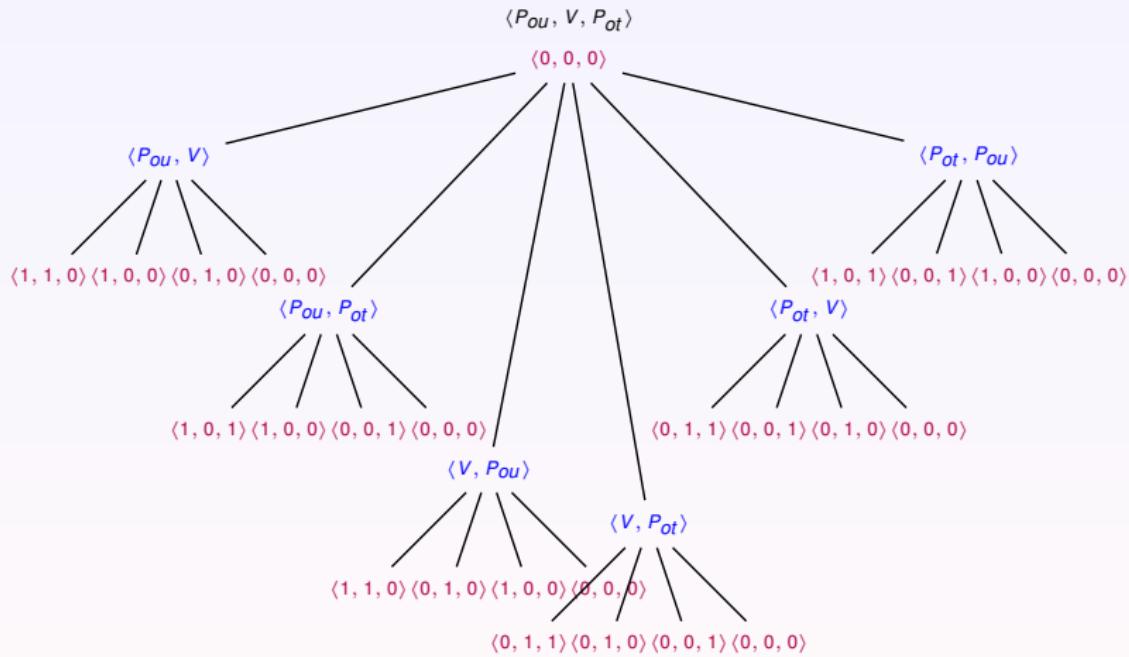
Exemple illustratif à la ferme

Arbre de résolution d'un MACSP avec *allDiff*



Exemple illustratif à la ferme

Arbre de résolution d'un MACSP avec *allDiff*



7 états possibles selon les probabilités
pour seulement 6 actions jointes

Résultats théoriques

Gain dans un problème d'allocation séquentiel

Dans un problème d'allocation de ressources à des tâches, le gain en facteur de branchement de l'utilisation de la contrainte *allDiff* qui imposent que toutes les ressources soient appliquées à des tâches différentes est égal à :

$$Gain = \frac{|T|^{|X|}}{C_{|T|}^{|X|}}$$

Résultats théoriques

Gain dans un problème d'allocation séquentiel

Dans un problème d'allocation de ressources à des tâches, le gain en facteur de branchement de l'utilisation de la contrainte *allDiff* qui imposent que toutes les ressources soient appliquées à des tâches différentes est égal à :

$$Gain = \frac{|T|^{|X|}}{C_{|T|}^{|X|}}$$

Complexité du modèle en pire cas

Trouver une politique pour un MACSP, qui obtiendrait une récompense espérée d'au moins C , est #P.

Résultats théoriques

Gain dans un problème d'allocation séquentiel

Dans un problème d'allocation de ressources à des tâches, le gain en facteur de branchement de l'utilisation de la contrainte *allDiff* qui imposent que toutes les ressources soient appliquées à des tâches différentes est égal à :

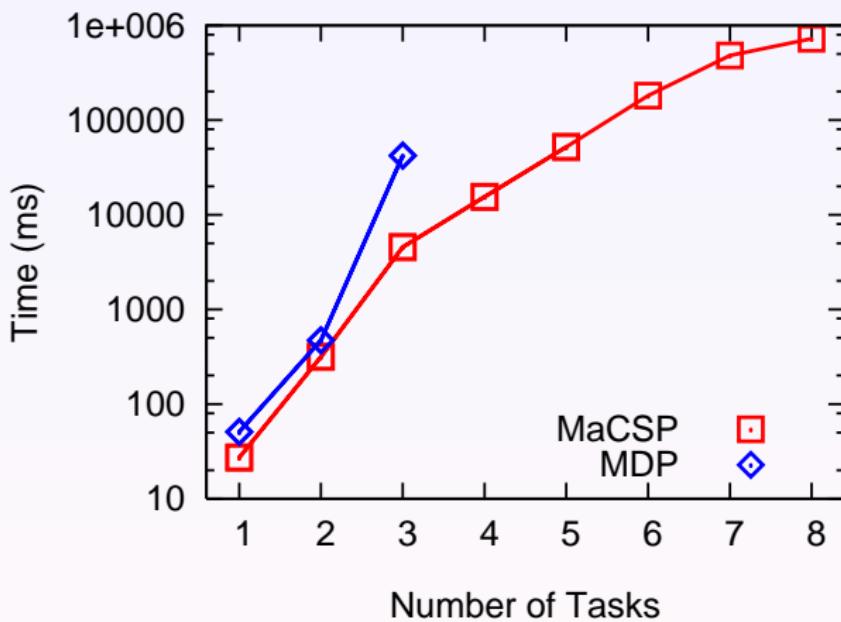
$$Gain = \frac{|T|^{|X|}}{C_{|T|}^{|X|}}$$

Complexité du modèle en pire cas

Trouver une politique pour un MACSP, qui obtiendrait une récompense espérée d'au moins C , est #P.

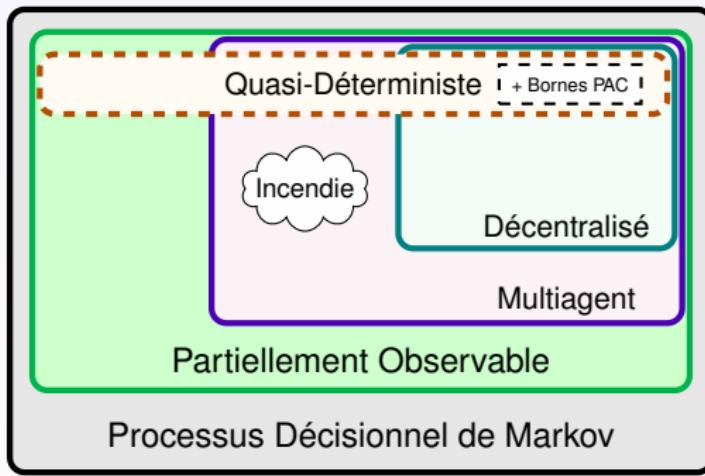
Mais cela fonctionne très bien *en pratique* puisque le nombre de variables est très inférieur au nombre d'états.

Résultats empiriques

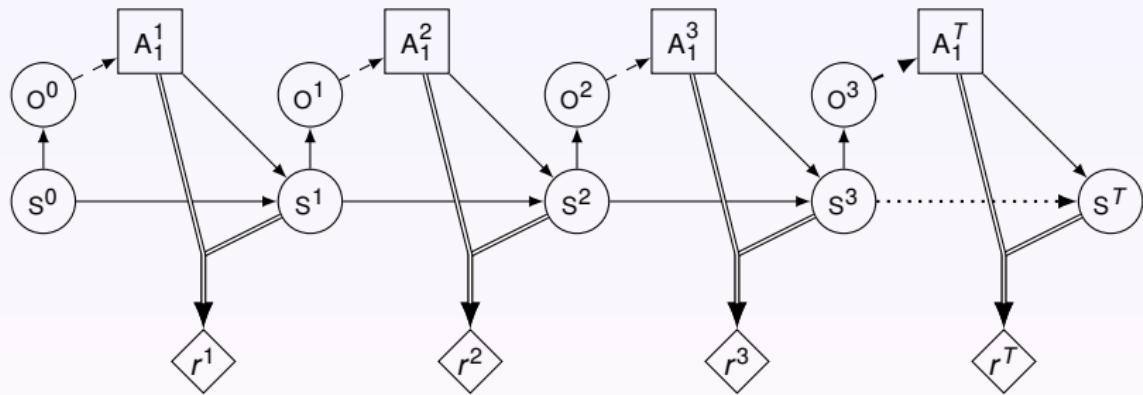


Plan

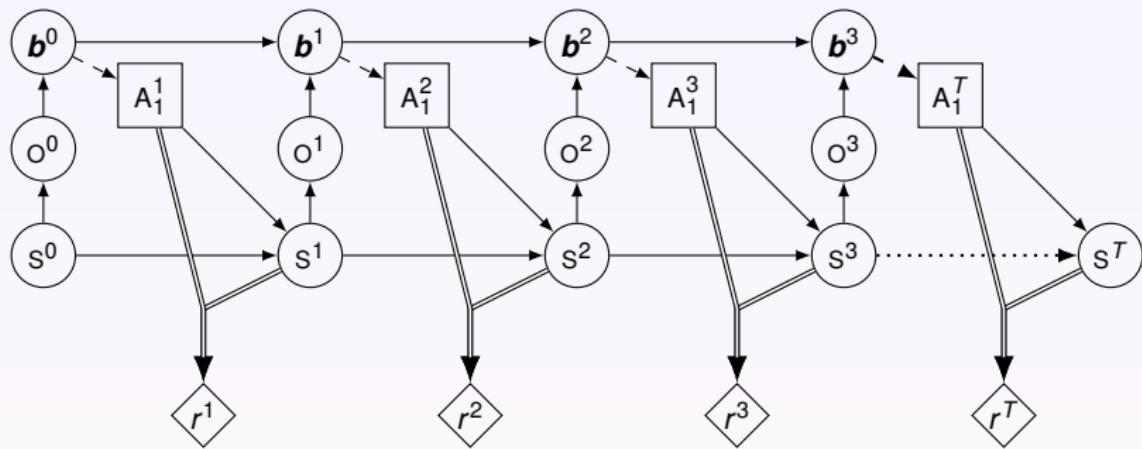
2^{eme} Partie : Contraintes et Observabilité dans les Modèles de Markov



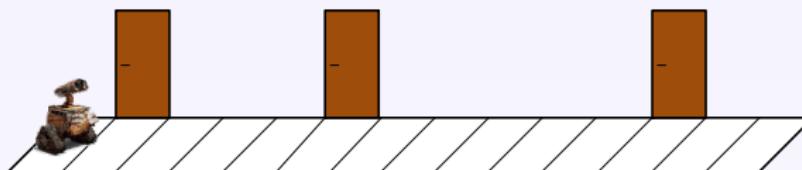
L'état de croyance



L'état de croyance



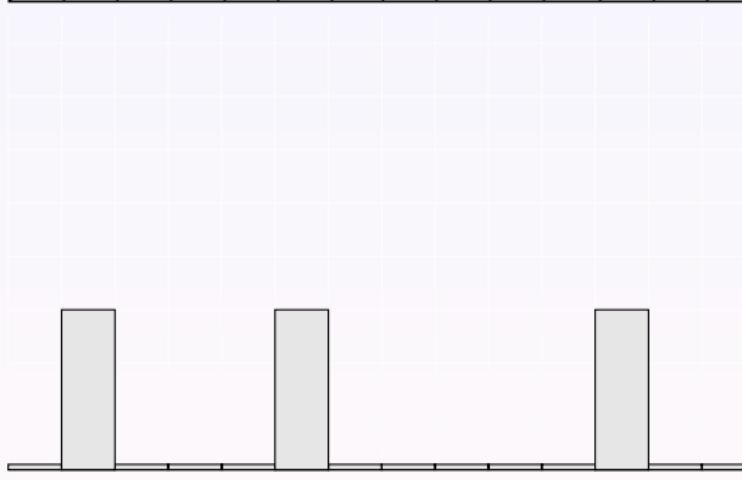
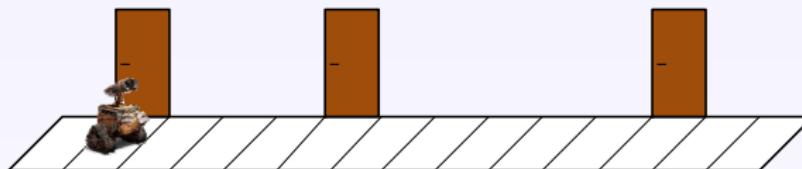
L'état de croyance : un exemple



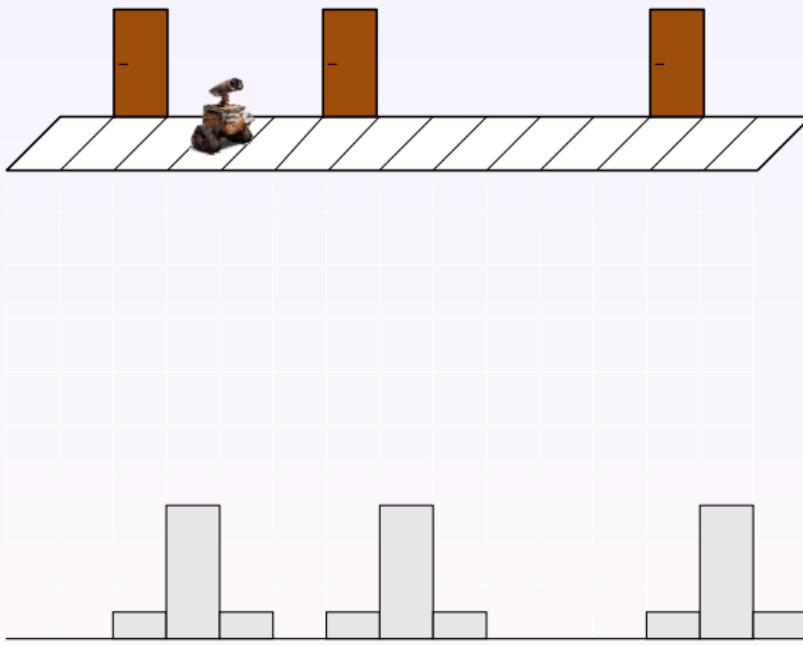
$b(s)$



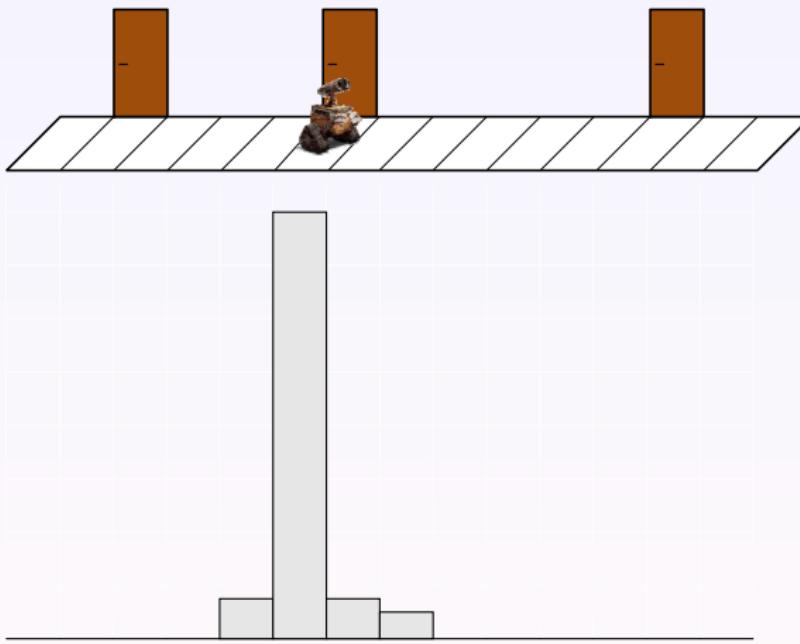
L'état de croyance : un exemple



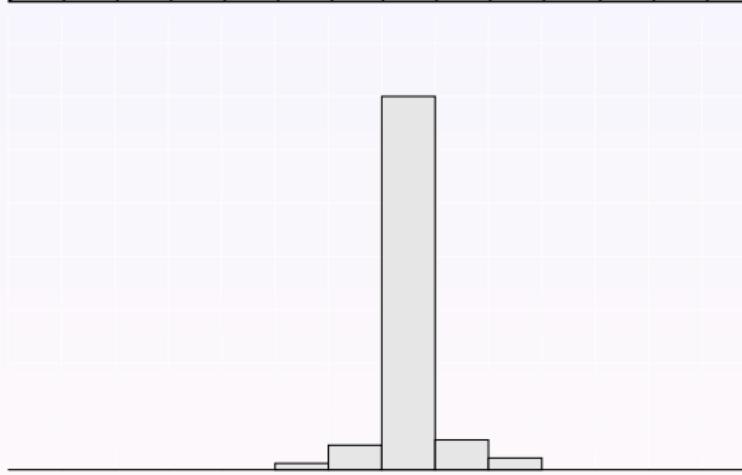
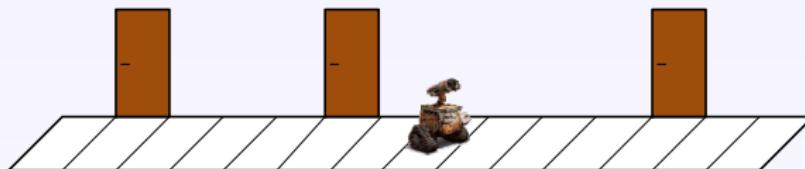
L'état de croyance : un exemple



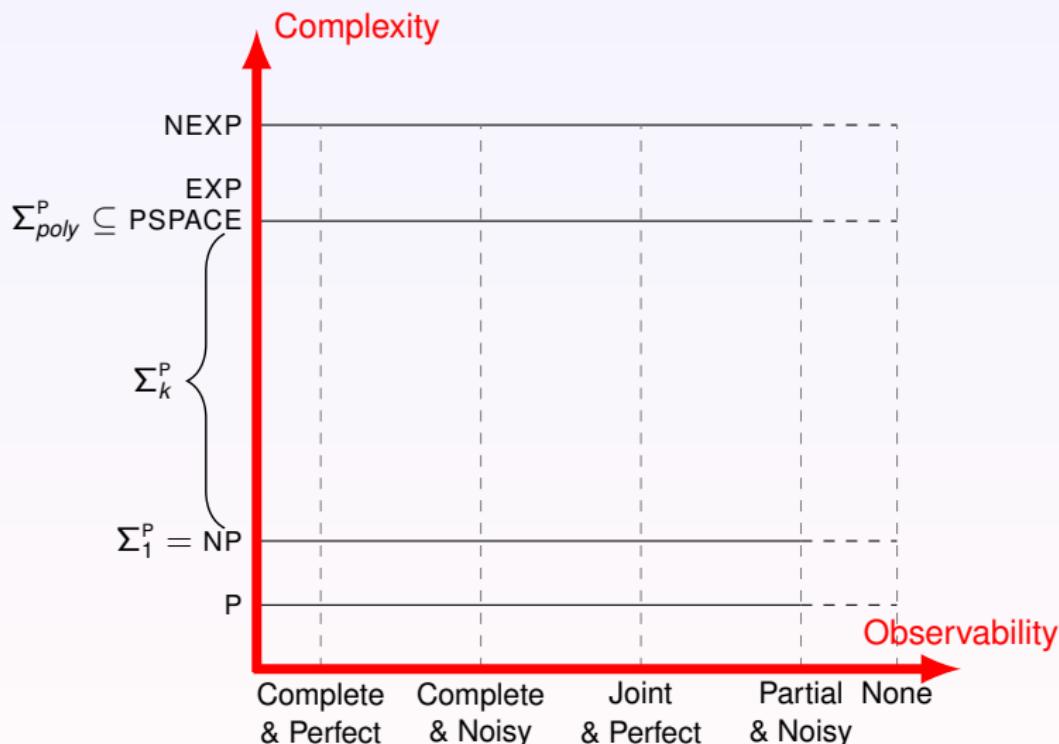
L'état de croyance : un exemple



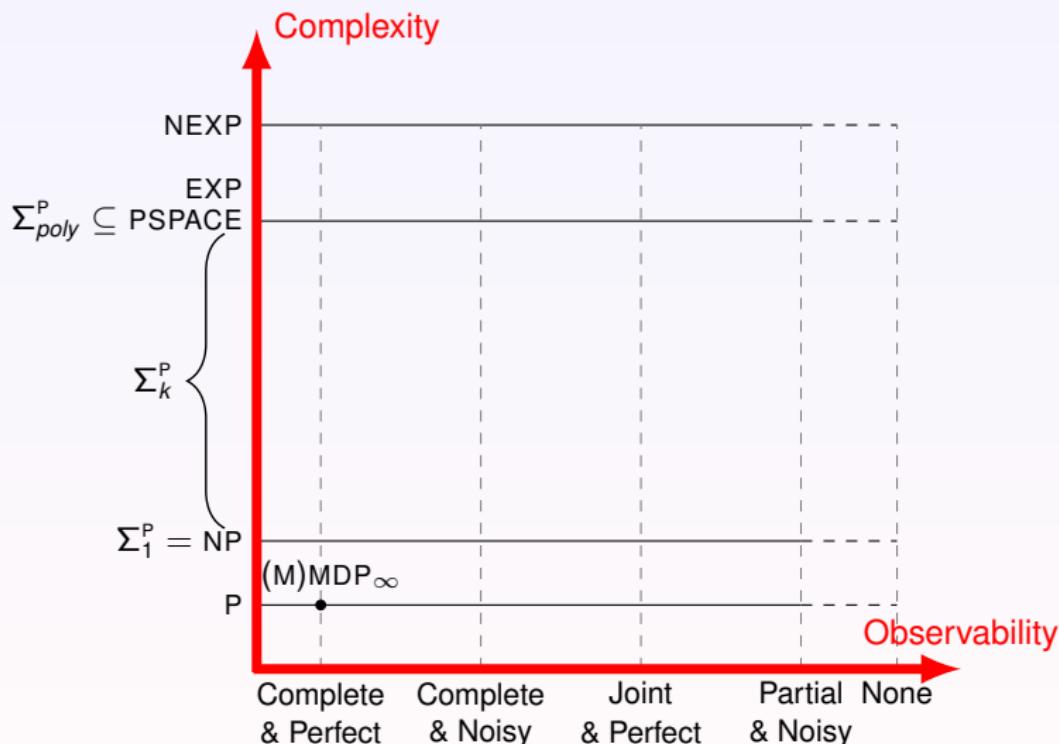
L'état de croyance : un exemple



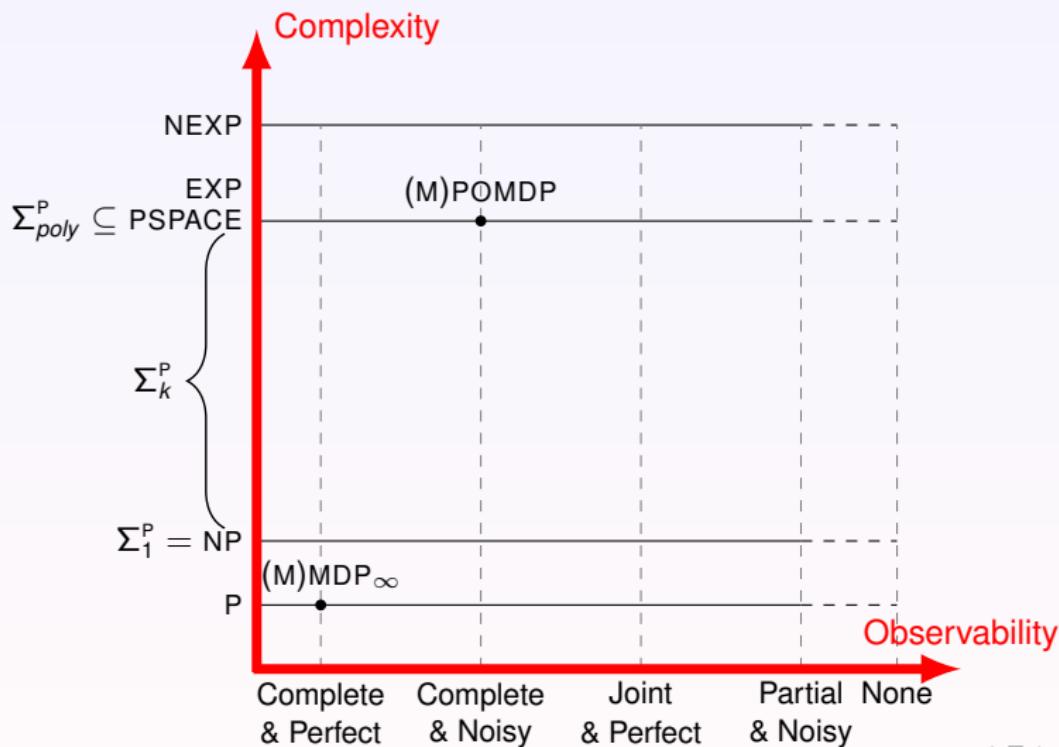
Complexité des modèles de Markov selon l'observabilité



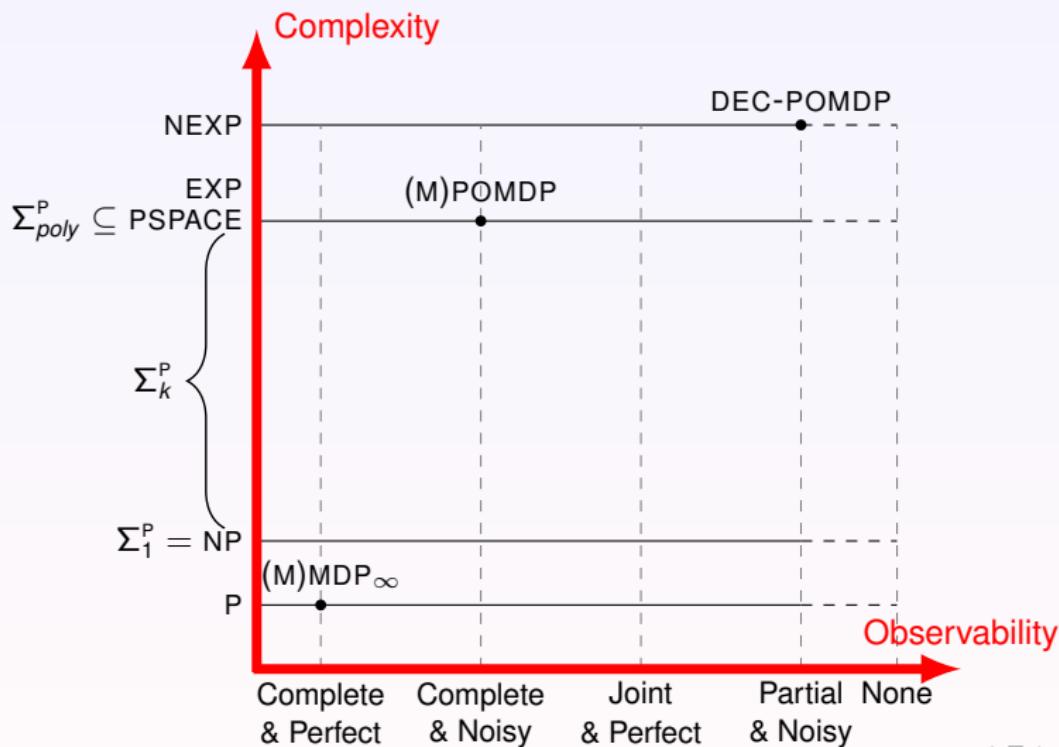
Complexité des modèles de Markov selon l'observabilité



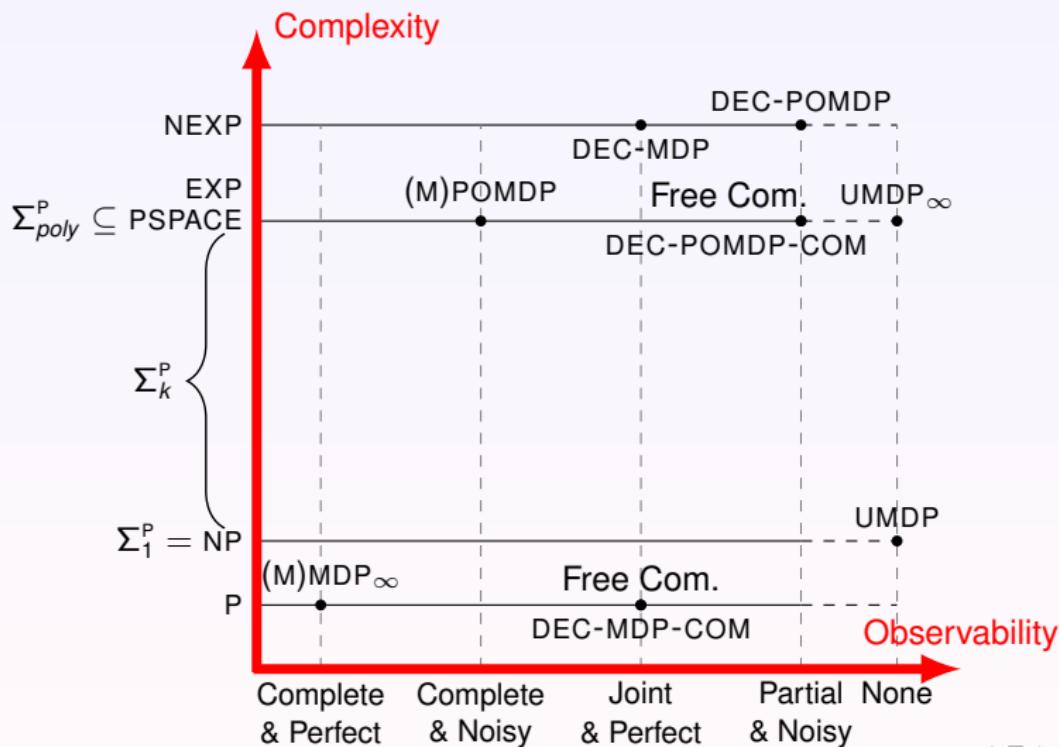
Complexité des modèles de Markov selon l'observabilité



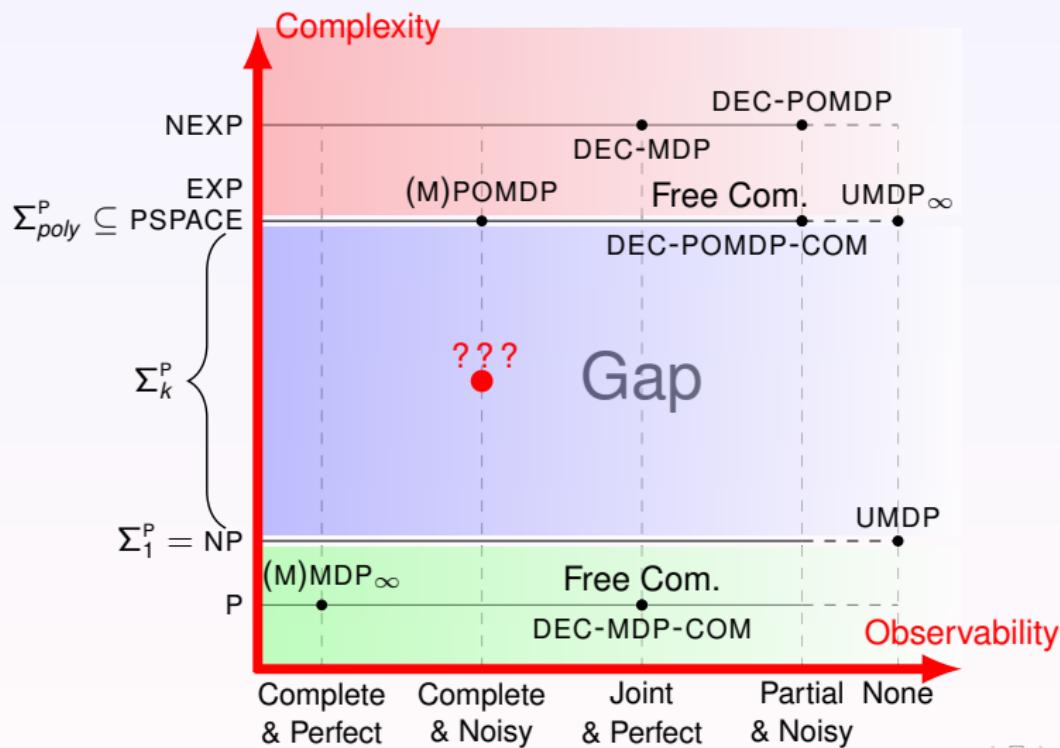
Complexité des modèles de Markov selon l'observabilité



Complexité des modèles de Markov selon l'observabilité



Complexité des modèles de Markov selon l'observabilité



Modèle avec observation bijective

Observabilité bijective

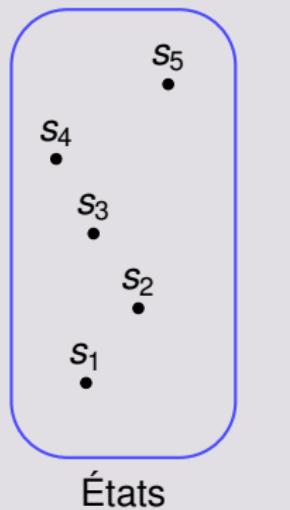
$\exists o_1 \in \Omega, \forall a \in \mathcal{A}, \forall s \in \mathcal{S}^{o_1},$

with $\mathcal{S}^{o_1} = \{s \in \mathcal{S} | \exists o_1 \in \Omega, P(o_1|s, a) > P(o|s, a), \forall o \neq o_1\},$

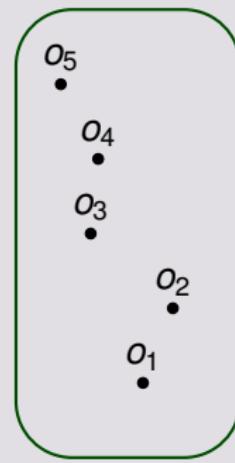
then $|\Omega| = |\mathcal{S}|$ and $|\mathcal{S}^{o_1}| = 1$

Modèle avec observation bijective

Observabilité bijective



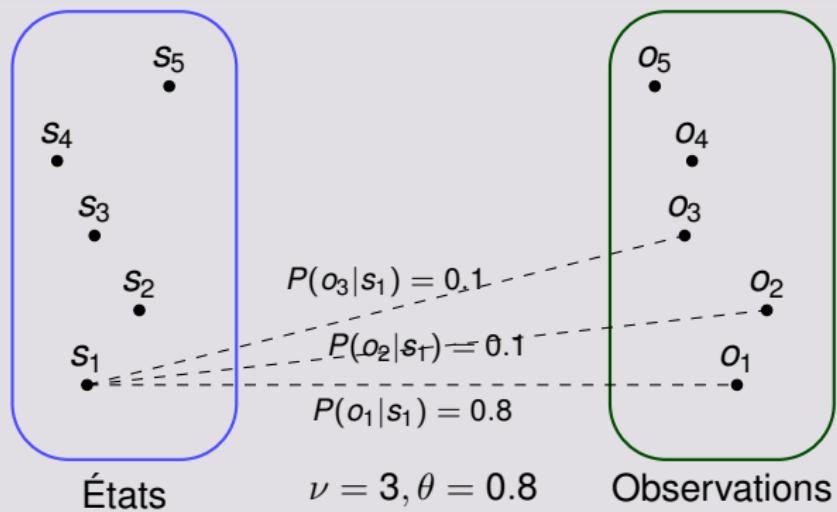
États



Observations

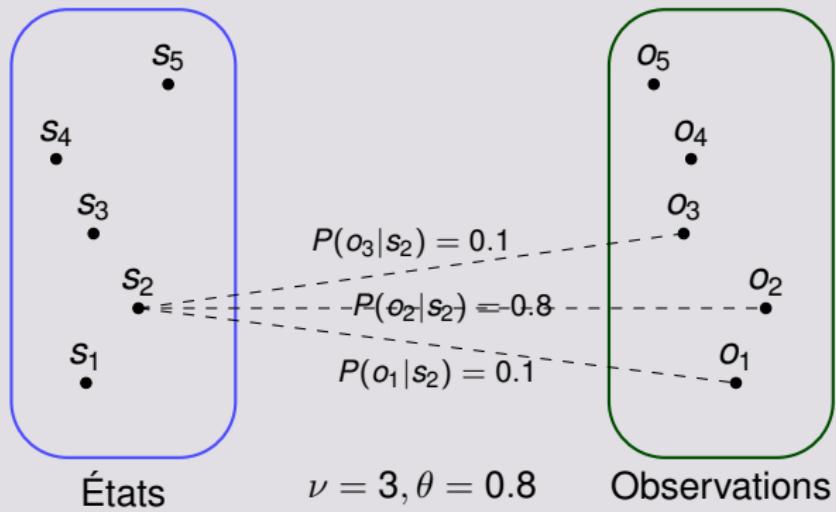
Modèle avec observation bijective

Observabilité bijective



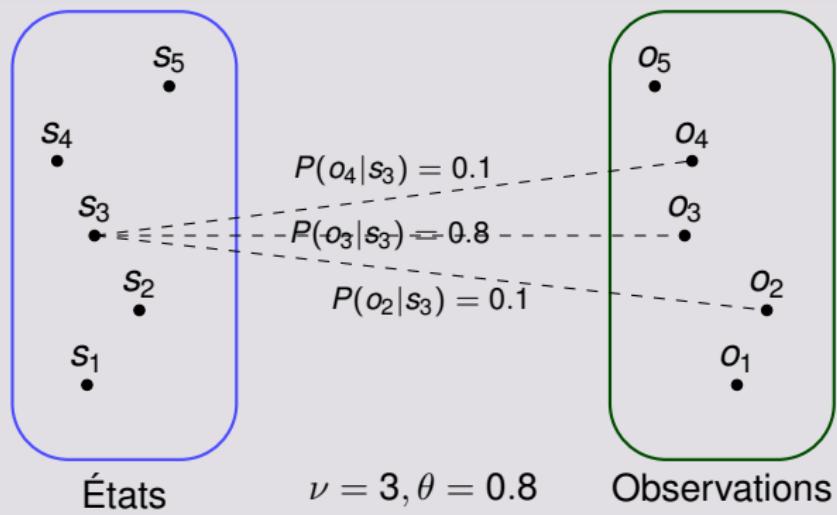
Modèle avec observation bijective

Observabilité bijective



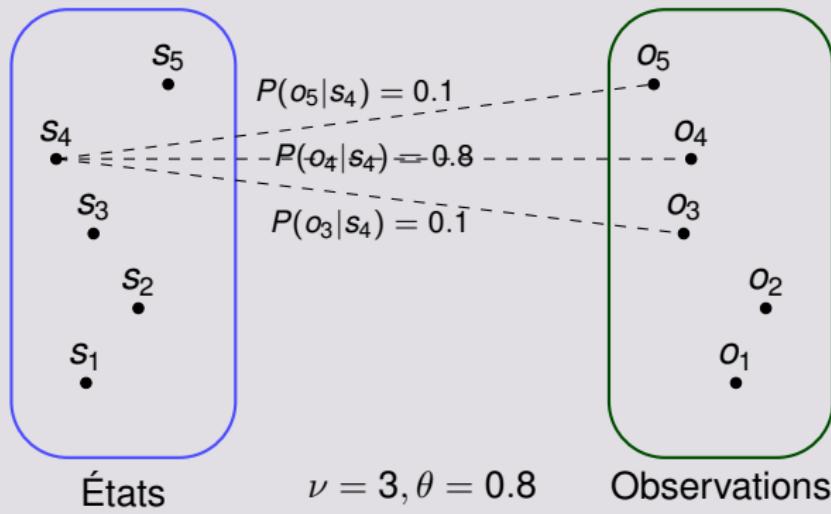
Modèle avec observation bijective

Observabilité bijective



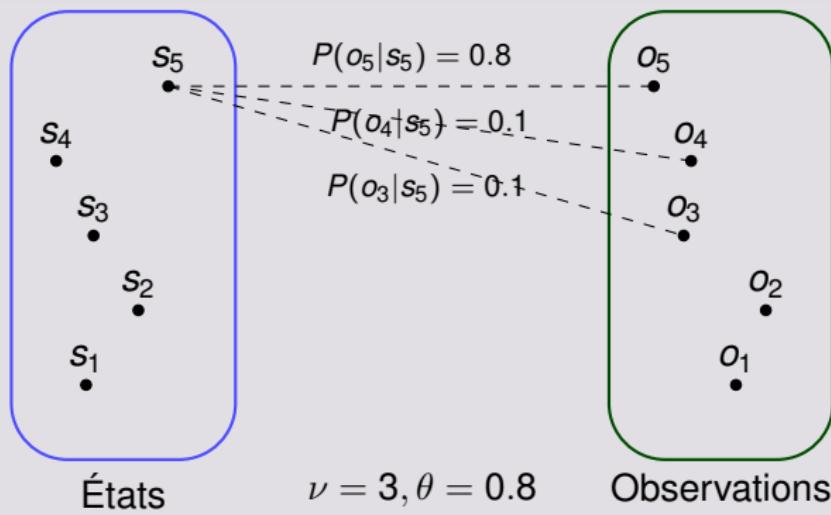
Modèle avec observation bijective

Observabilité bijective



Modèle avec observation bijective

Observabilité bijective



Résultats théoriques obtenus

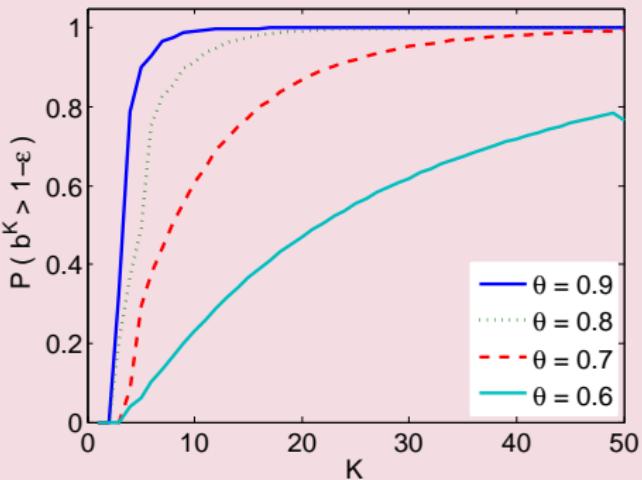
Convergence de l'état de croyance en probabilité

Sous l'hypothèse d'observation bijective, l'état de croyance \mathbf{b}^k est tel que $\mathbf{b}^k(s) \geq 1 - \varepsilon$ ssi

$$n \geq \frac{1}{2 \ln \frac{\nu\theta}{(1-\theta)}} \ln \left[\frac{1-\varepsilon}{\varepsilon} \left(1 + \nu^{1-\frac{k}{2}} \right) \right] + \frac{k}{2} \quad (1)$$

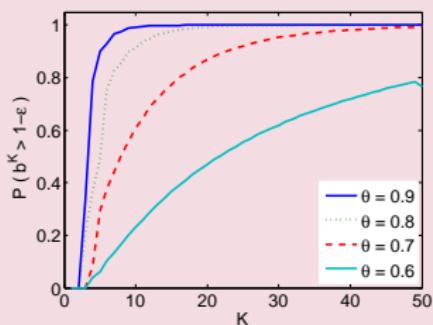
Résultats théoriques obtenus

Convergence de l'état de croyance en probabilité



Résultats théoriques obtenus

Convergence de l'état de croyance en probabilité

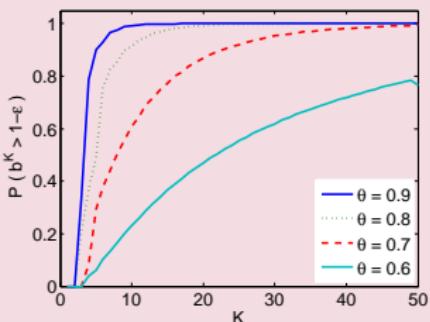


Réduction de la complexité en pire cas pour les POMDPs

Étant donné un POMDP à horizon infini avec actions déterministes et respectant l'hypothèse d'observabilité bijective, trouver une politique qui permette d'obtenir une récompense escomptée espérée donnée avec une forte probabilité est Σ_{2k-1}^P .

Résultats théoriques obtenus

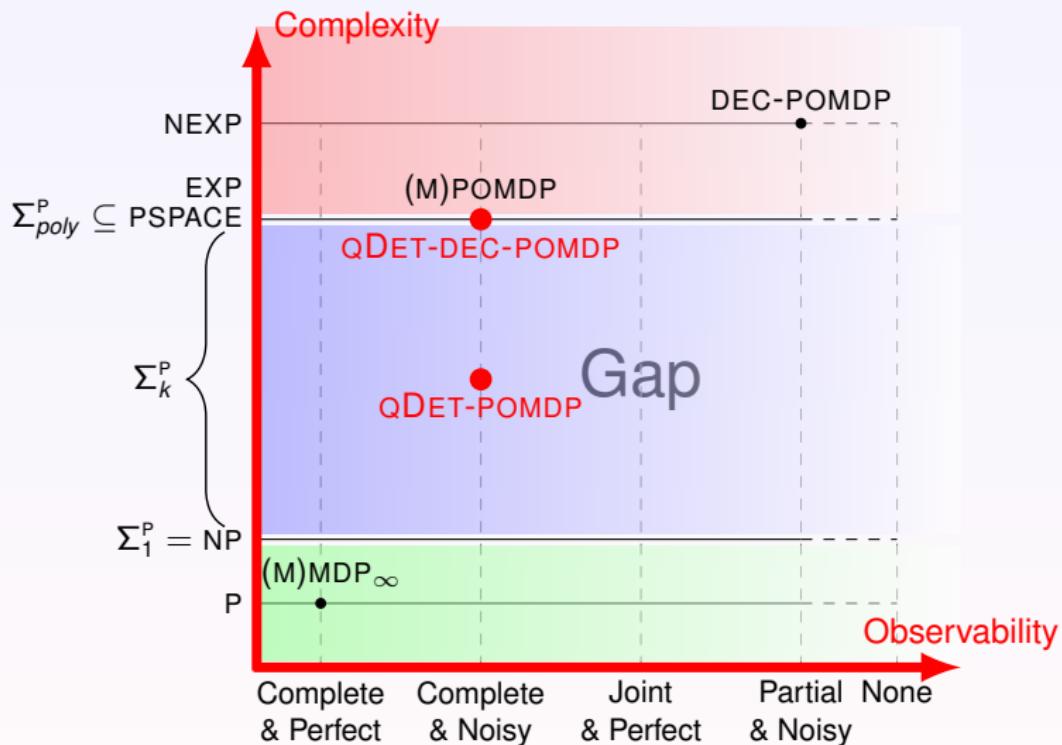
Convergence de l'état de croyance en probabilité



Réduction de la complexité en pire cas pour les DEC-POMDPs

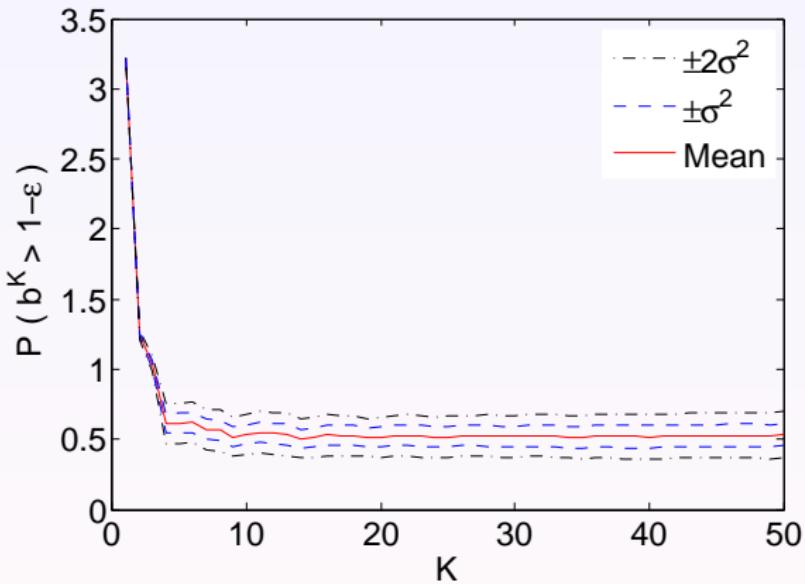
Étant donné un DEC-POMDP à horizon infini avec actions déterministes et respectant l'hypothèse d'observabilité bijective, trouver une politique qui permette d'obtenir une récompense escomptée espérée donnée avec une forte probabilité est PSPACE.

Résultats théoriques obtenus



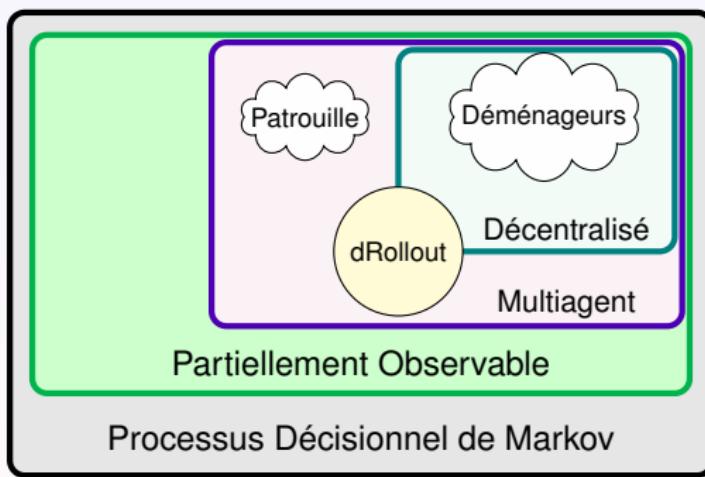
avec actions stochastiques

$$\sigma_a^2 = .05 \quad \sigma_0^2 = 0.12892 \quad \theta = 0.92214$$

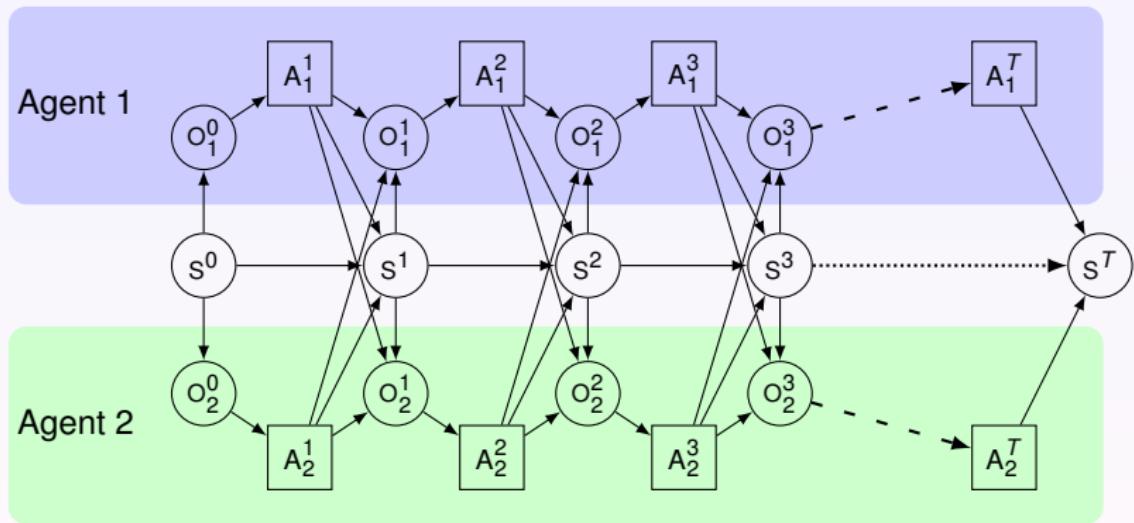


Plan

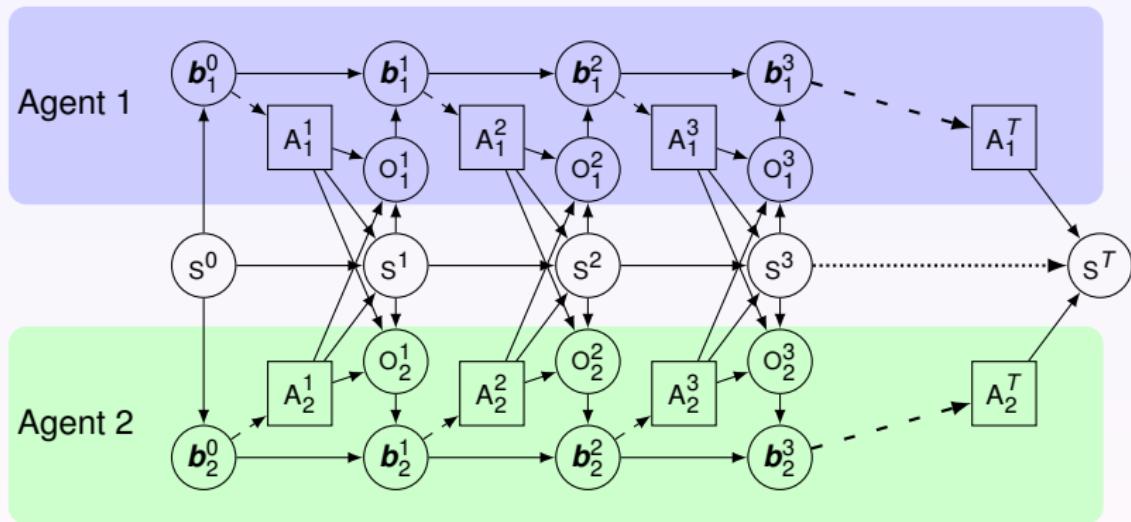
3^{eme} Partie : Observabilité dans les modèles de Markov décentralisés



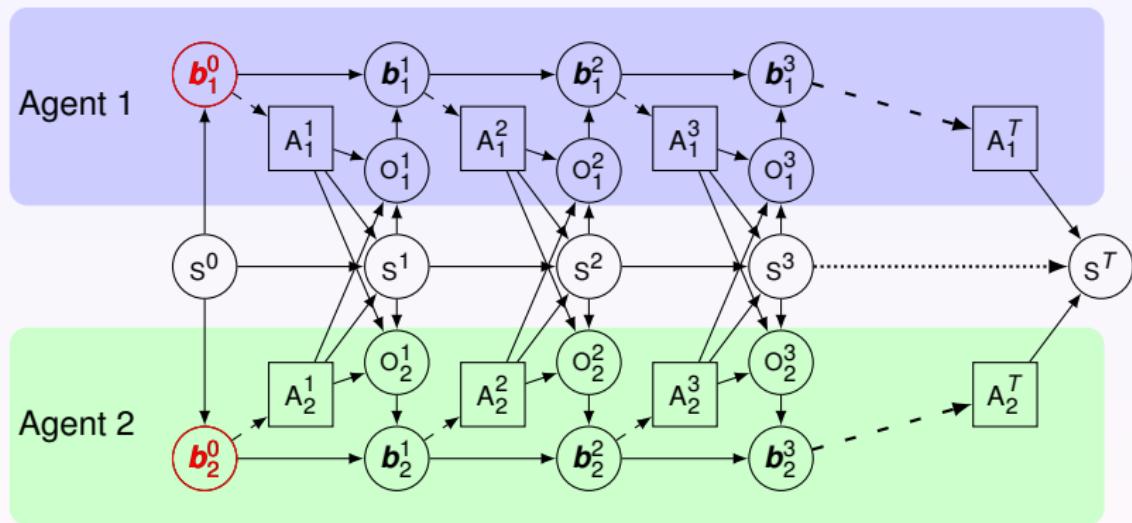
Bases sur les POMDPs décentralisés



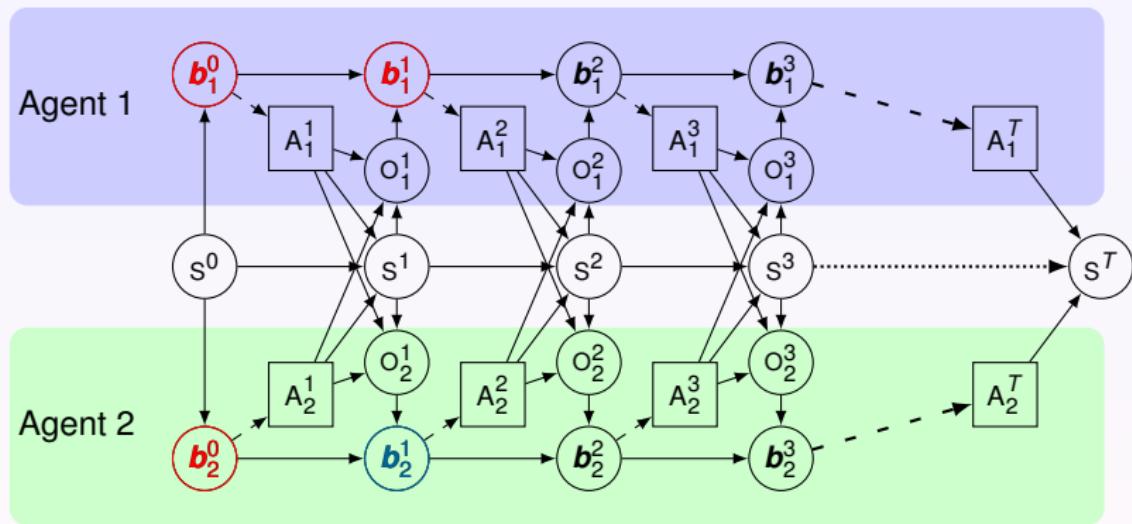
Bases sur les POMDPs décentralisés



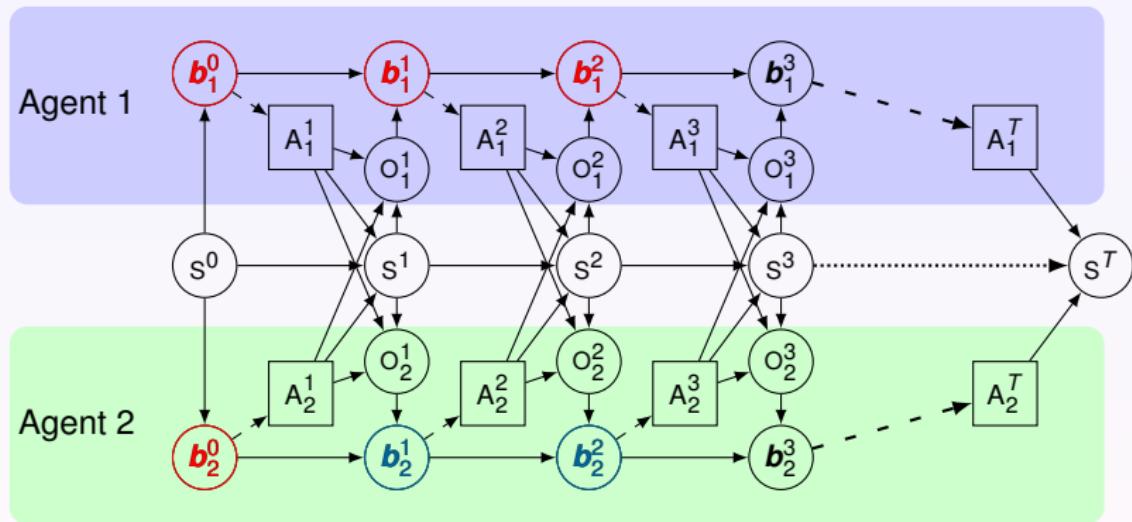
Bases sur les POMDPs décentralisés



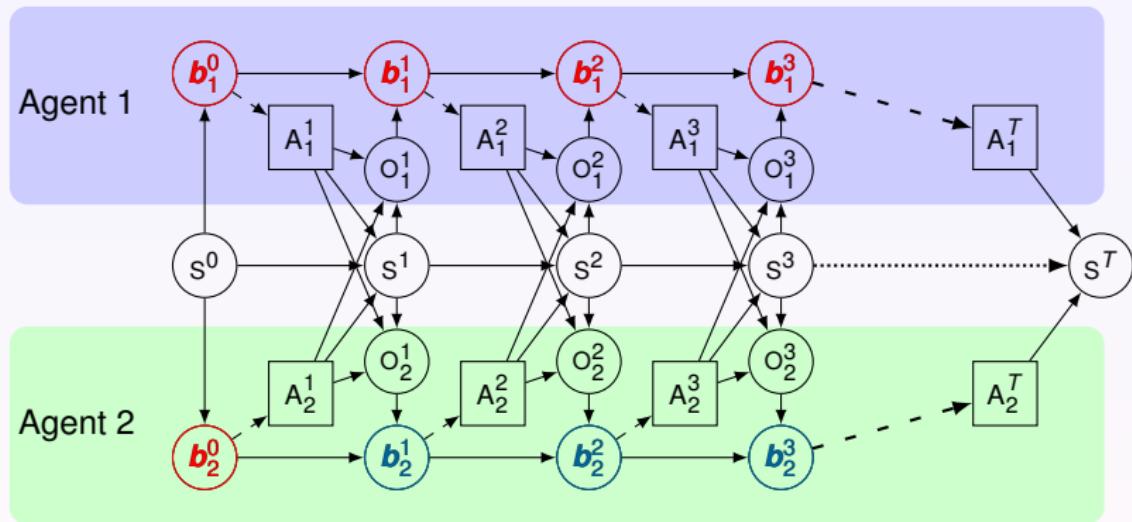
Bases sur les POMDPs décentralisés



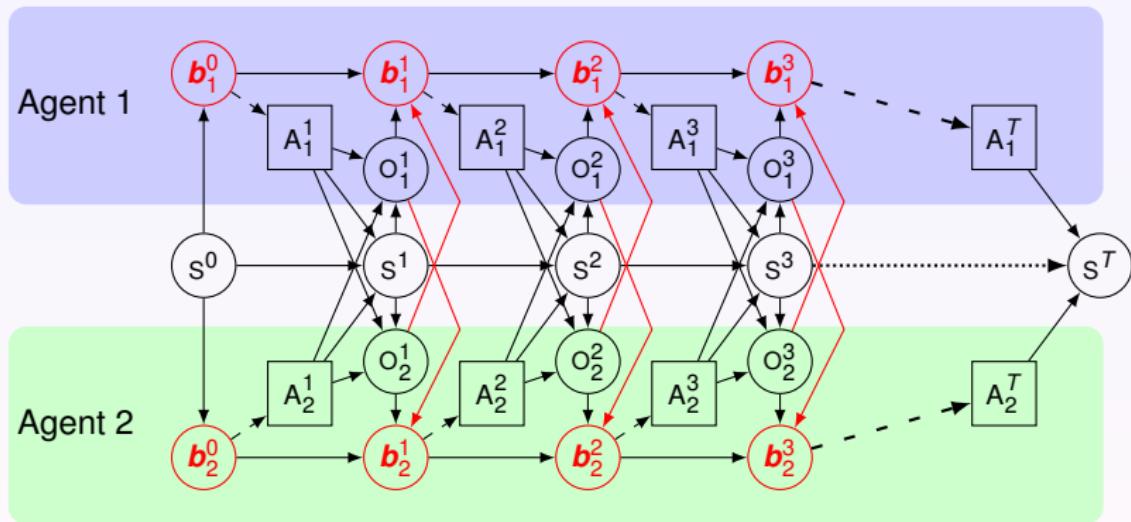
Bases sur les POMDPs décentralisés



Bases sur les POMDPs décentralisés



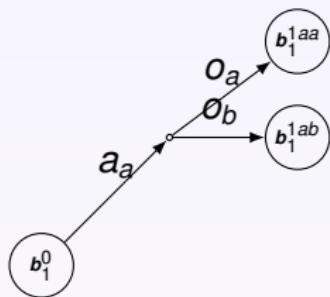
Bases sur les POMDPs décentralisés



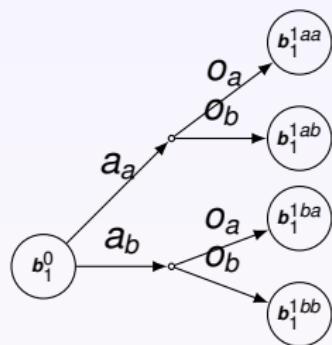
Algorithme de Rollout décentralisé

$$b_1^0$$

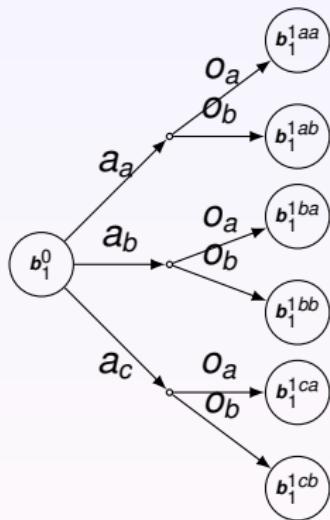
Algorithme de Rollout décentralisé



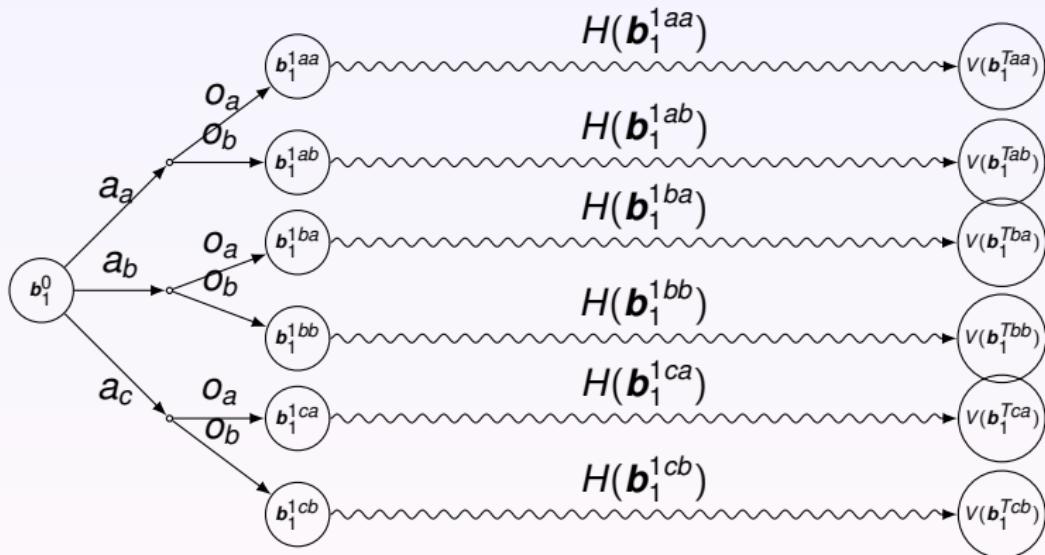
Algorithme de Rollout décentralisé



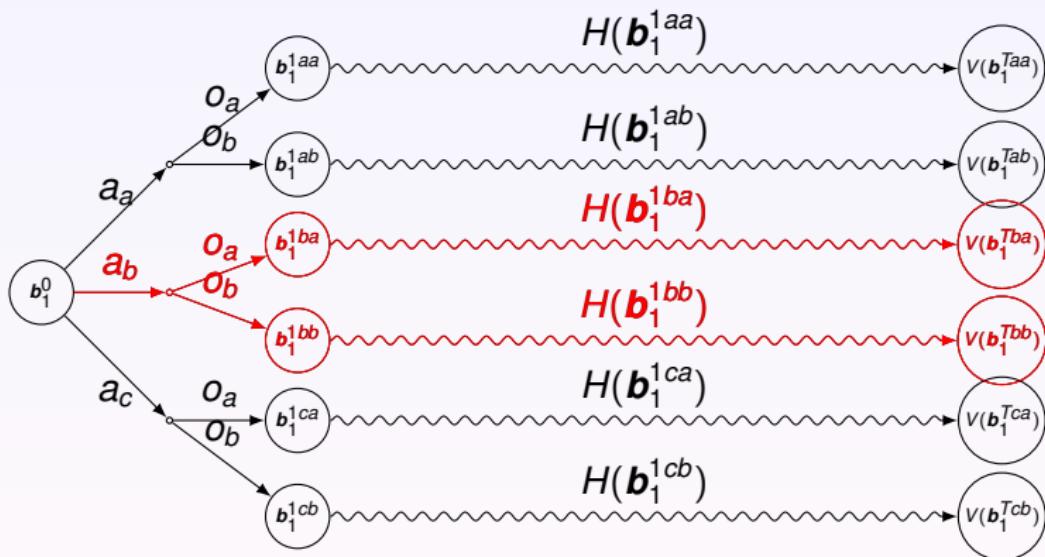
Algorithme de Rollout décentralisé



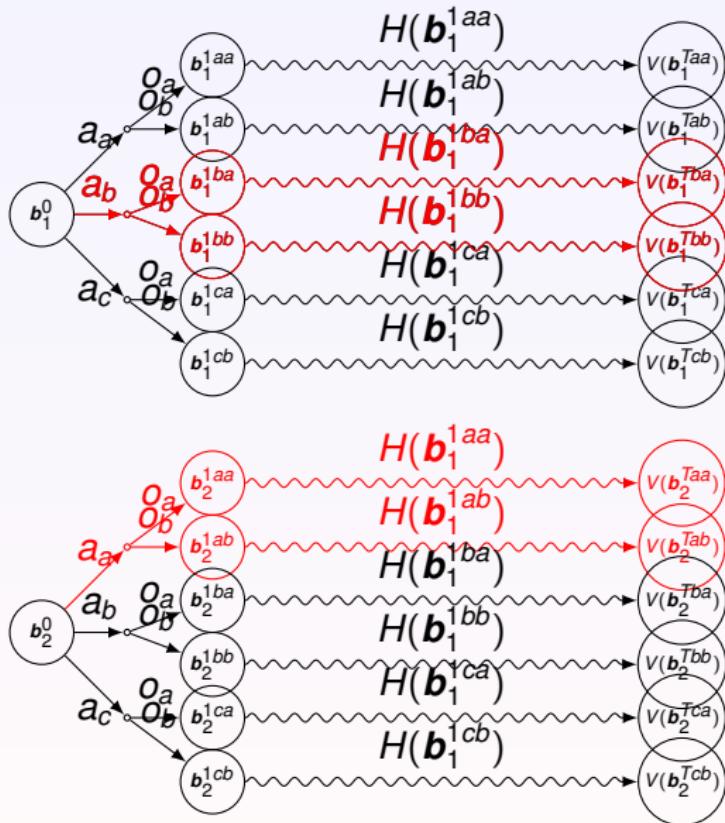
Algorithme de Rollout décentralisé



Algorithme de Rollout décentralisé



Algorithme de Rollout décentralisé



Planification en ligne avec observabilité décentralisée et sans communication

Observation la plus probable

$$\begin{aligned} \mathbf{o}_i^* &\leftarrow \langle o_i, \mathbf{o}_{-i}^* \rangle \\ \text{où } \mathbf{o}_{-i}^* &= \arg \max_{\mathbf{o}_{-i} \in \Omega_{-i}} \mathcal{O}(\langle o_i, \mathbf{o}_{-i} \rangle | \mathbf{a}, s') \mathbf{b}_i^t(s') \\ \mathbf{b}_i^{t+1}(s') &= \frac{\mathcal{O}(\mathbf{o}_i^* | \mathbf{a}, s') \sum_{s \in \mathcal{S}} \mathcal{T}(s' | s, \mathbf{a}) \mathbf{b}_i^t(s)}{\sum_{s, s' \in \mathcal{S}} \mathcal{O}(\mathbf{o}_i^* | \mathbf{a}, s') \mathcal{T}(s' | s, \mathbf{a}) \mathbf{b}_i^t(s)} \end{aligned} \quad (1)$$

Planification en ligne avec observabilité décentralisée et sans communication

Observation la plus probable

$$\begin{aligned}
 \mathbf{o}_i^* &\leftarrow \langle \mathbf{o}_i, \mathbf{o}_{-i}^* \rangle \\
 \text{où } \mathbf{o}_{-i}^* &= \arg \max_{\mathbf{o}_{-i} \in \Omega_{-i}} \mathcal{O}(\langle \mathbf{o}_i, \mathbf{o}_{-i} \rangle | \mathbf{a}, s') \mathbf{b}_i^t(s') \\
 \mathbf{b}_i^{t+1}(s') &= \frac{\mathcal{O}(\mathbf{o}_i^* | \mathbf{a}, s') \sum_{s \in \mathcal{S}} \mathcal{T}(s' | s, \mathbf{a}) \mathbf{b}_i^t(s)}{\sum_{s, s' \in \mathcal{S}} \mathcal{O}(\mathbf{o}_i^* | \mathbf{a}, s') \mathcal{T}(s' | s, \mathbf{a}) \mathbf{b}_i^t(s)} \quad (1)
 \end{aligned}$$

Observation espérée (*)

$$\mathbf{b}_i^{t+1}(s') = \sum_{\mathbf{o}_{-i} \in \Omega_{-i}} \frac{\mathcal{O}(\mathbf{o} | \mathbf{a}, s') \sum_{s \in \mathcal{S}} \mathcal{T}(s' | s, \mathbf{a}) \mathbf{b}_i^t(s)}{\sum_{s, s' \in \mathcal{S}} \mathcal{O}(\mathbf{o} | \mathbf{a}, s') \mathcal{T}(s' | s, \mathbf{a}) \mathbf{b}_i^t(s)}, \quad \mathbf{o} = \langle \mathbf{o}_i, \mathbf{o}_{-i} \rangle \quad (2)$$

Planification en ligne avec observabilité décentralisée et sans communication

Observation la plus probable

$$\boldsymbol{b}_i^{t+1}(s') = \max_{\boldsymbol{o}_{-i}} \boldsymbol{b}_i^{t+1}(s') \quad (1)$$

Observation espérée (*)

$$\boldsymbol{b}_i^{t+1}(s') = \sum_{\boldsymbol{o}_{-i} \in \Omega_{-i}} \frac{\mathcal{O}(\boldsymbol{o}|\boldsymbol{a}, s') \sum_{s \in \mathcal{S}} \mathcal{T}(s'|s, \boldsymbol{a}) \boldsymbol{b}_i^t(s)}{\sum_{s, s' \in \mathcal{S}} \mathcal{O}(\boldsymbol{o}|\boldsymbol{a}, s') \mathcal{T}(s'|s, \boldsymbol{a}) \boldsymbol{b}_i^t(s)}, \quad \boldsymbol{o} = \langle o_i, \boldsymbol{o}_{-i} \rangle \quad (2)$$

Planification en ligne avec observabilité décentralisée et sans communication

Observation la plus probable

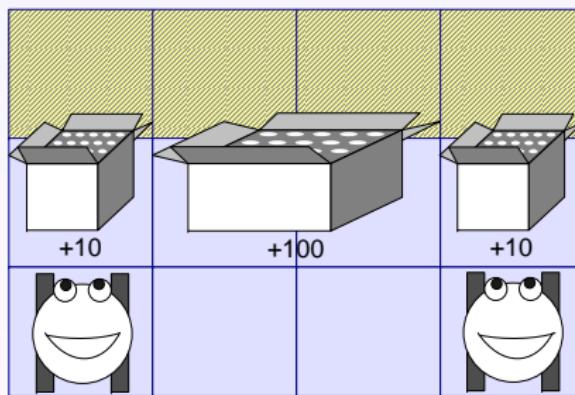
$$\boldsymbol{b}_i^{t+1}(s') = \max_{\boldsymbol{o}_{-i}} \boldsymbol{b}_i^{t+1}(s') \quad (1)$$

Observation espérée (\star)

$$\boldsymbol{b}_i^{t+1}(s') = \mathbb{E}_{\boldsymbol{o}_{-i}} \boldsymbol{b}_i^{t+1}(s') \quad (2)$$

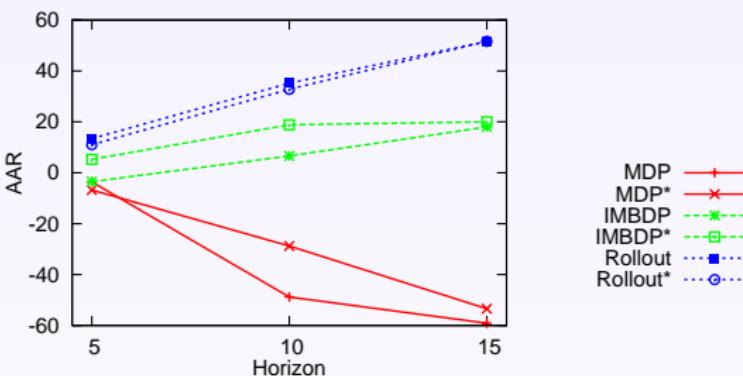
Résultats expérimentaux

Problème du déménageurs

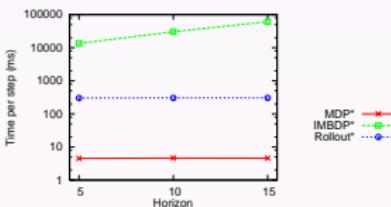


Résultats expérimentaux

Problème du déménageurs



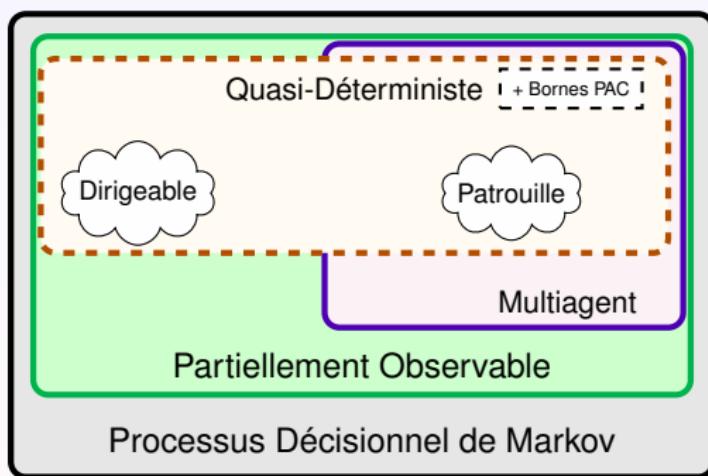
(a) Value vs Horizon



(b) Time (s) vs Horizon

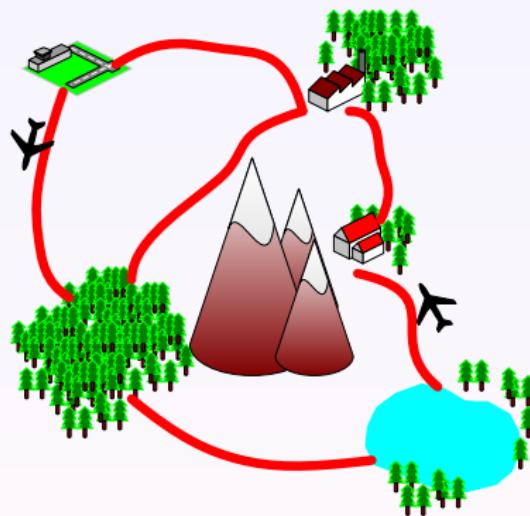
Plan

4^{eme} partie : Applications des outils développés



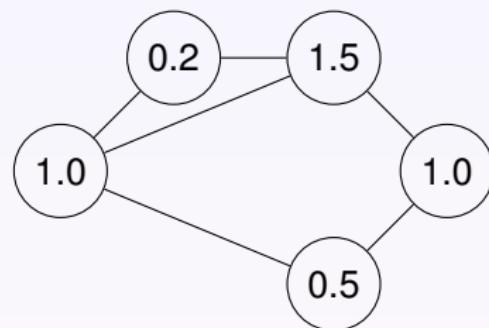
Extensions en cours

Groupe de patrouille autonome – J.S. Marier



Extensions en cours

Groupe de patrouille autonome – J.S. Marier



Extensions en cours

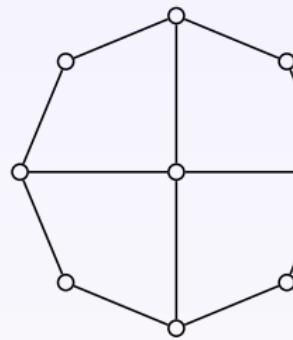
Groupe de patrouille autonome – J.S. Marier

Les agents :

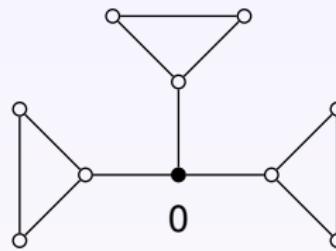
- Sont contraints dans leurs actions par un graphe de localisations à surveiller aussi souvent que possible ;
- Observent avec certitude leur position et communiquent celle-ci aux autres agents ;
- Peuvent communiquer leur état local et leur politique locale avec une fréquence suffisamment élevée ;
- N'ont pas accès à l'information qu'ils surveillent.

Extensions en cours

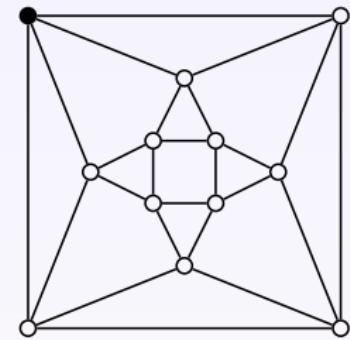
Groupe de patrouille autonome – J.S. Marier



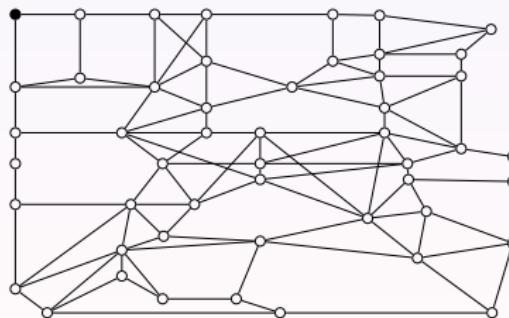
(c) Wheel



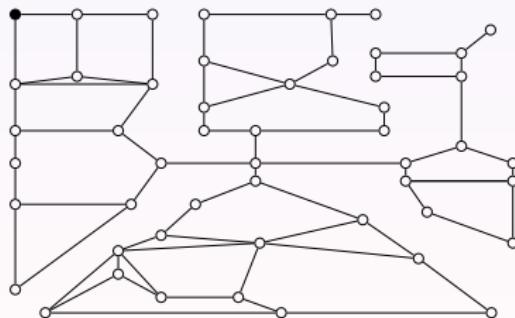
(d) Cloverleaf



(e) Cuboctahedron



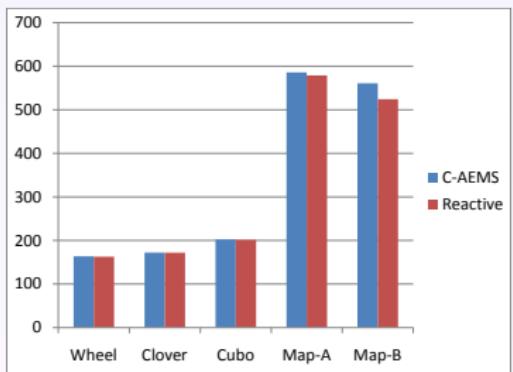
(f) Map-A



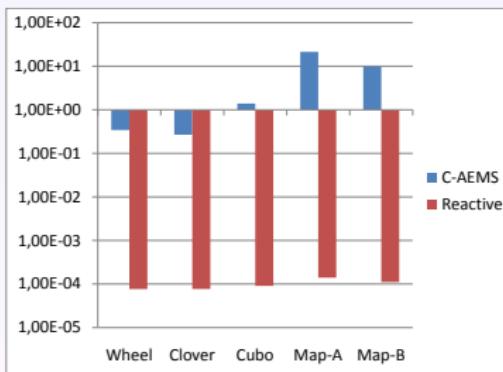
(g) Map-B

Extensions en cours

Groupe de patrouille autonome – J.S. Marier



(h) AAR



(i) Time(s)

Extensions en cours

Apprentissage de POMDPs continus – P. Dallaire

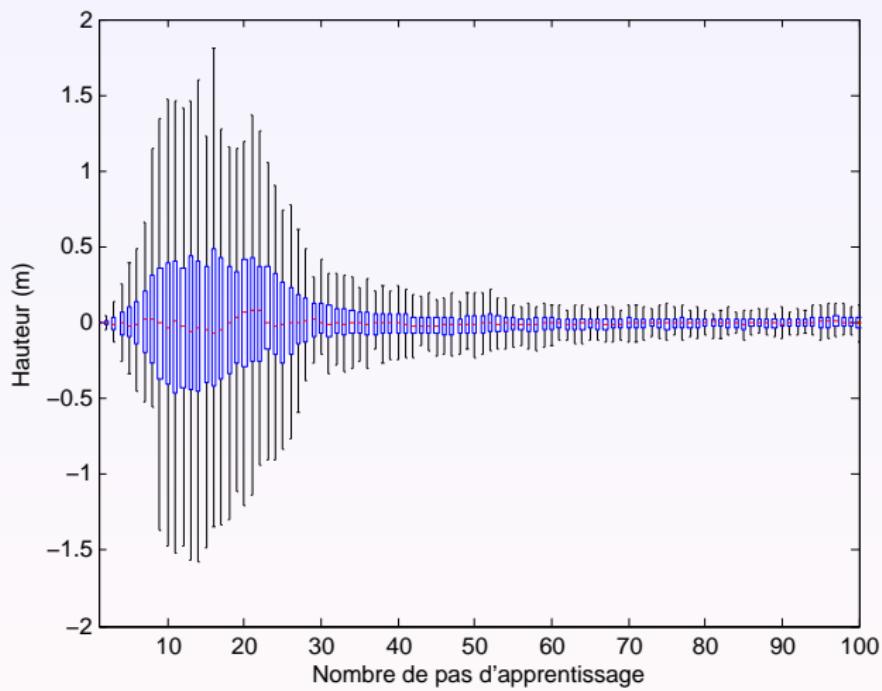


L'agent :

- Doit maintenir une hauteur relative de zero ;
- Observe sa hauteur en continu additionnée d'un bruit Gaussien (en cloche) ;
- Choisis un action monter/descendre en continu ;
- *Ne connaît pas sa dynamique.*

Extensions en cours

Apprentissage de POMDPs continus – P. Dallaire



Extensions possibles

- Extension du modèle MACSP au multiagent
 - Représentations des interactions agents par contraintes
- Étude des modèles d'observabilité bijective avec transition stochastique
 - et étude de l'équilibre génération / absorption d'entropie
- Extension des modèles d'observabilité bijective aux POMDPs continus
- Étude de la communication pour la synchronisation en ligne des états de croyance
 - À quel moment et quoi communiquer ?

Extensions possibles

- Extension du modèle MACSP au multiagent
 - Représentations des interactions agents par contraintes
- Étude des modèles d'observabilité bijective avec transition stochastique
 - et étude de l'équilibre génération / absorption d'entropie
- Extension des modèles d'observabilité bijective aux POMDPs continus
- Étude de la communication pour la synchronisation en ligne des états de croyance
 - À quel moment et quoi communiquer ?

Extensions possibles

- Extension du modèle MACSP au multiagent
 - Représentations des interactions agents par contraintes
- Étude des modèles d'observabilité bijective avec transition stochastique
 - et étude de l'équilibre génération / absorption d'entropie
- Extension des modèles d'observabilité bijective aux POMDPs continus
- Étude de la communication pour la synchronisation en ligne des états de croyance
 - À quel moment et quoi communiquer ?

Extensions possibles

- Extension du modèle MACSP au multiagent
 - Représentations des interactions agents par contraintes
- Étude des modèles d'observabilité bijective avec transition stochastique
 - et étude de l'équilibre génération / absorption d'entropie
- Extension des modèles d'observabilité bijective aux POMDPs continus
- Étude de la communication pour la synchronisation en ligne des états de croyance
 - À quel moment et quoi communiquer ?

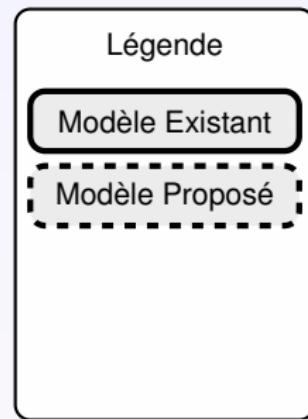
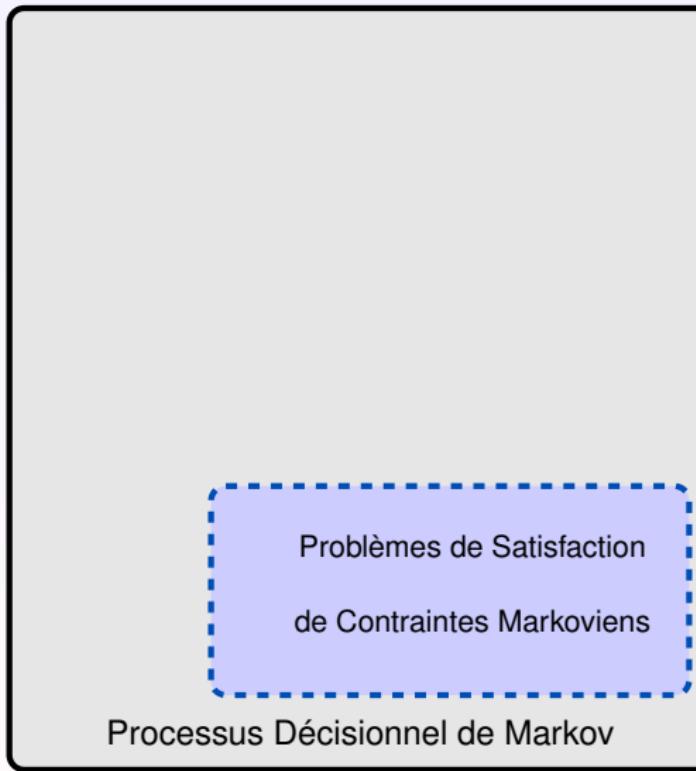
Rappel des Contributions

Processus Décisionnel de Markov

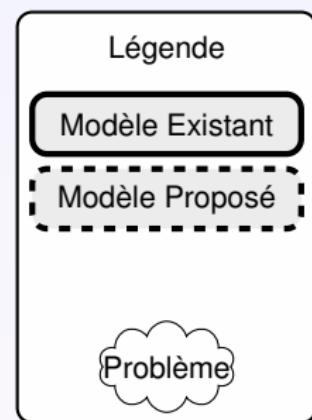
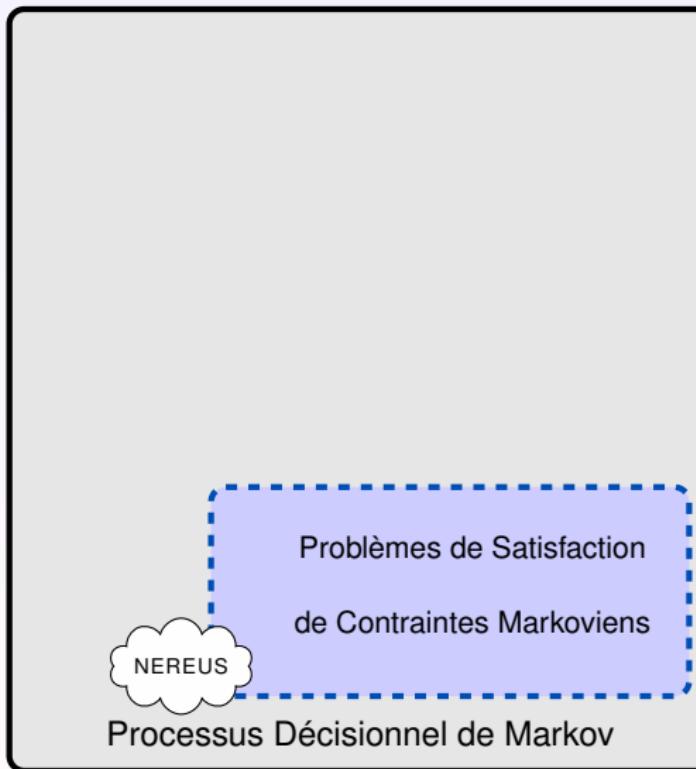
Légende

Modèle Existant

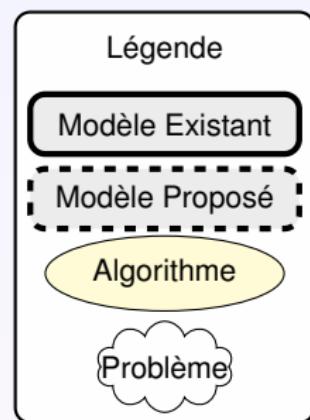
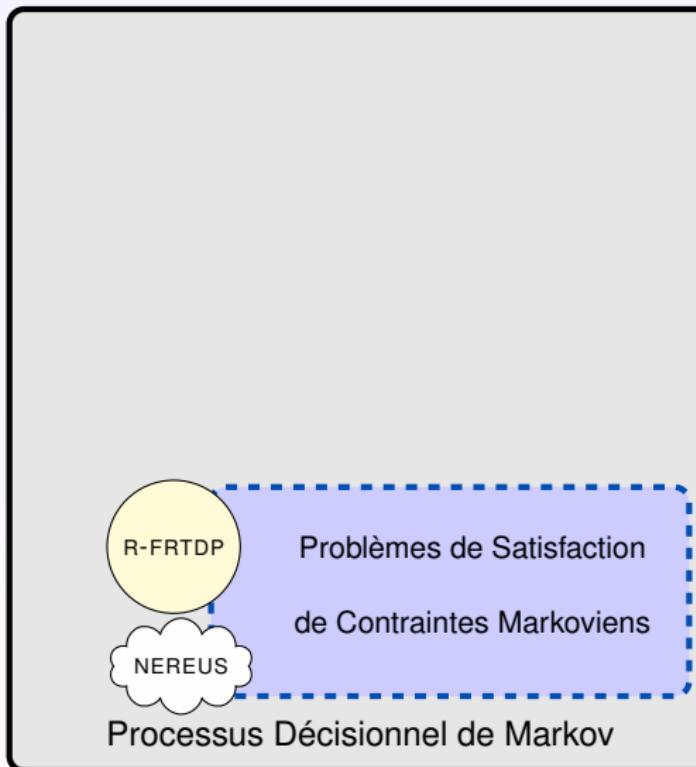
Rappel des Contributions



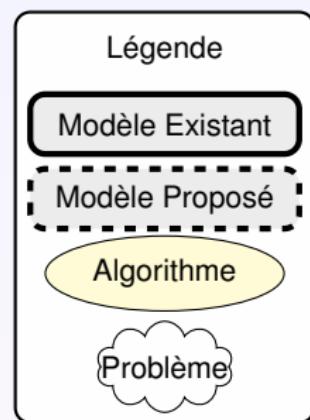
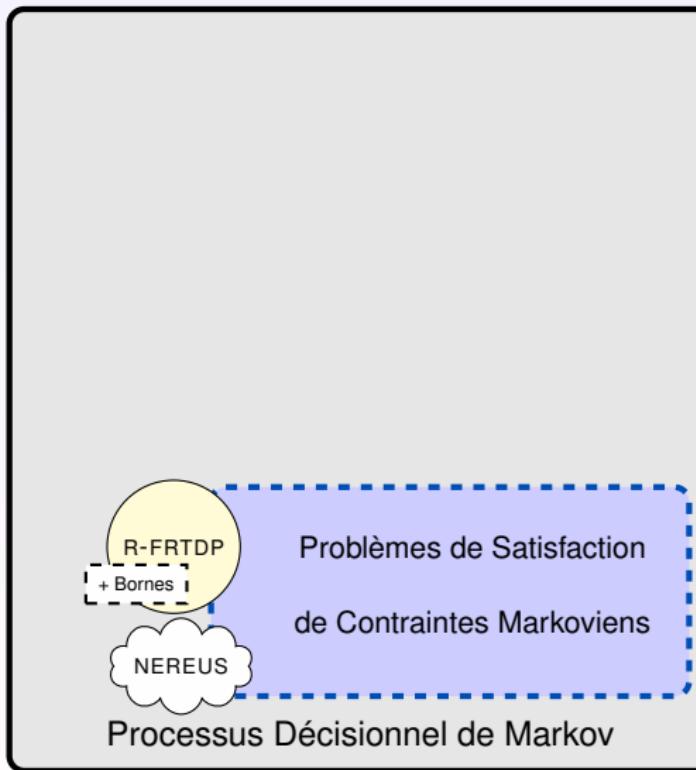
Rappel des Contributions



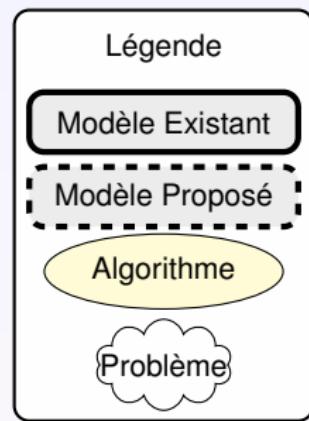
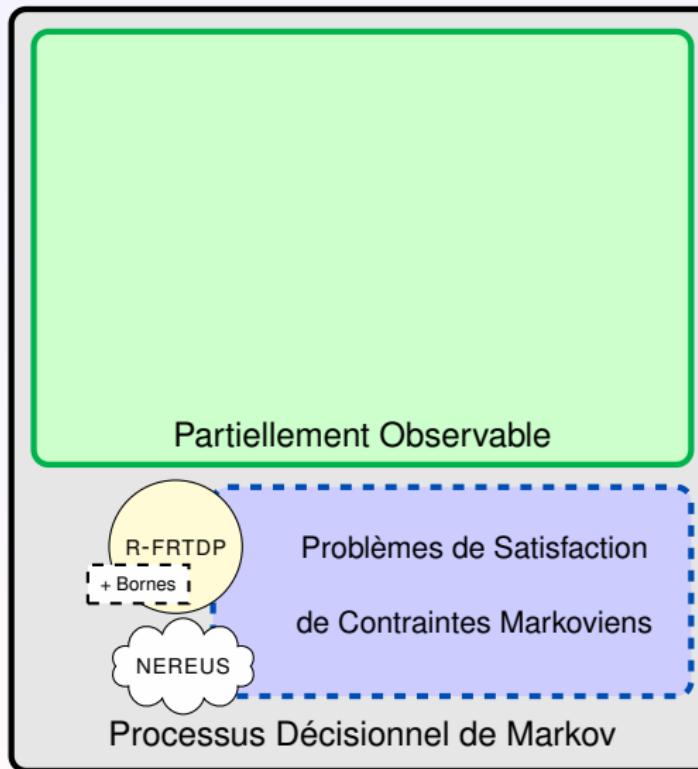
Rappel des Contributions



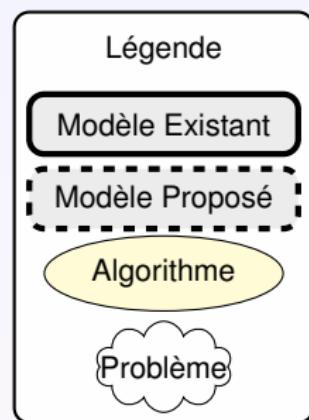
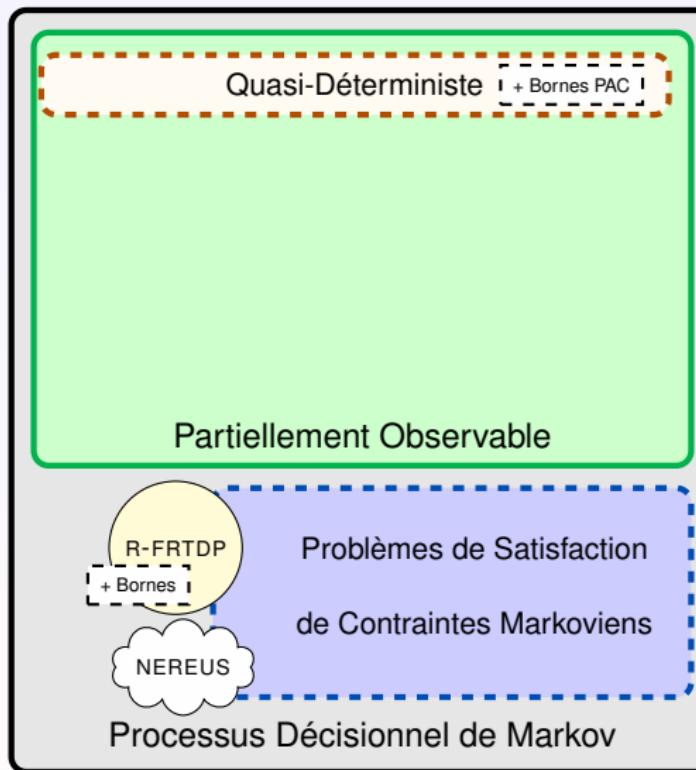
Rappel des Contributions



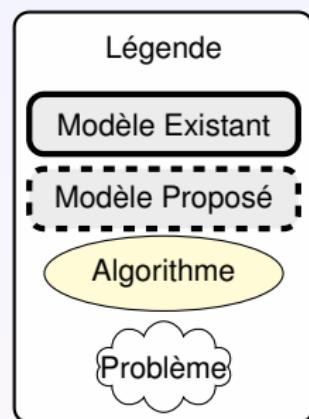
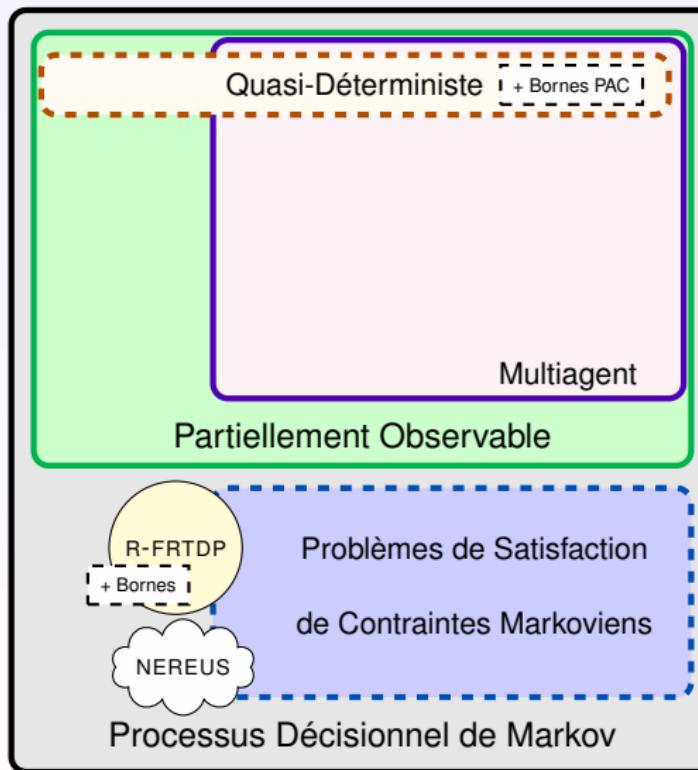
Rappel des Contributions



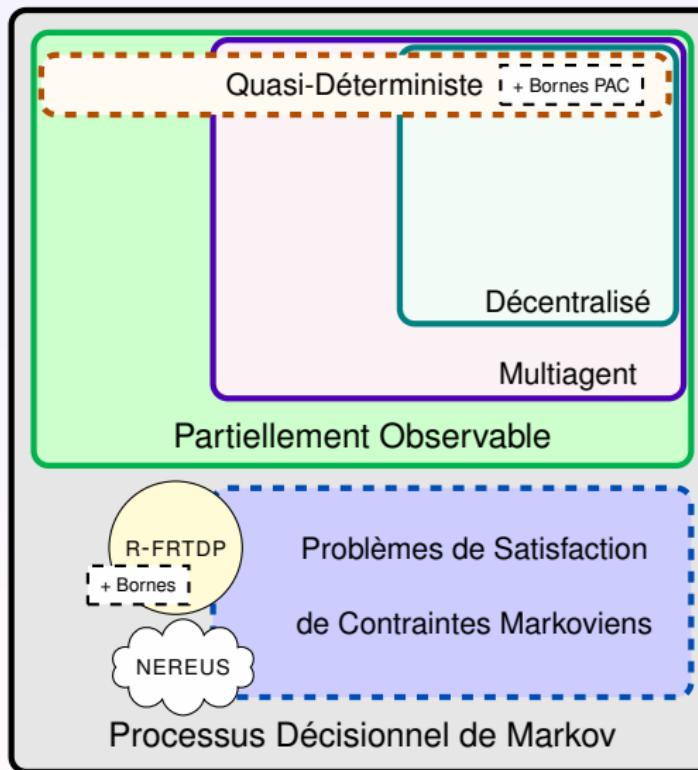
Rappel des Contributions



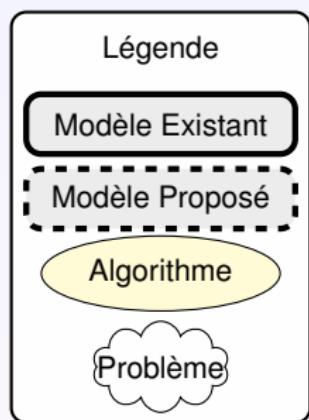
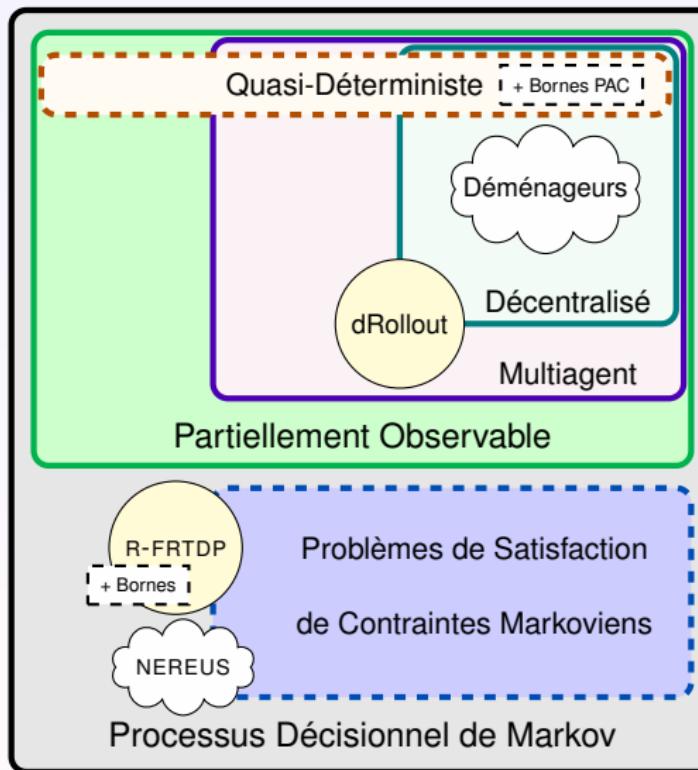
Rappel des Contributions



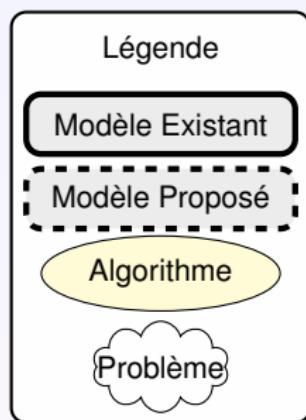
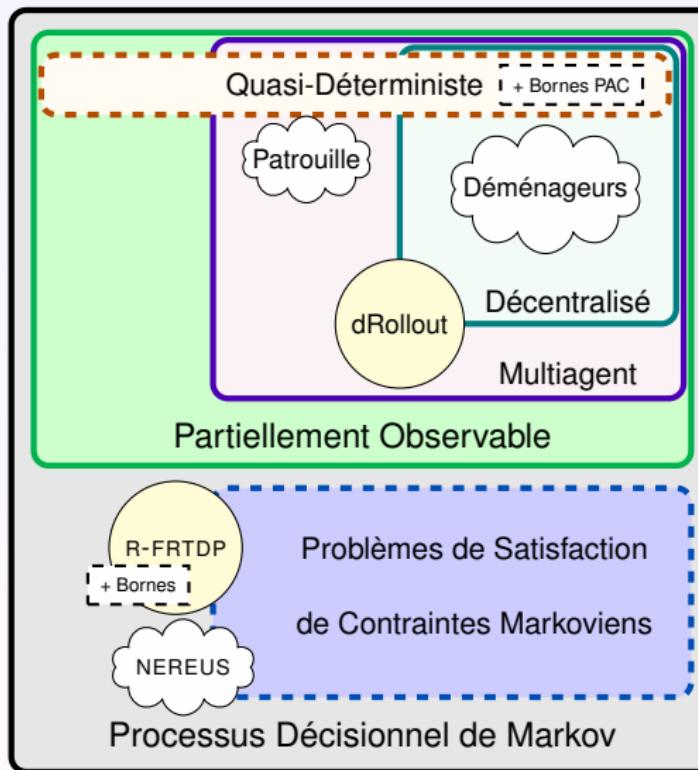
Rappel des Contributions



Rappel des Contributions



Rappel des Contributions



Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre
- ⇒ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre
- ⇒ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre
- ⇒ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre
- ⇒ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre
- ➡ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre
- ➡ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre

→ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Conclusion

- ✓ Abandon des concepts d'intentions, engagements. On ne raisonne que sur la valeur.
- ✓ Meilleur précision sur l'optimalité de la planification
- ✓ Pouvoir de modélisation très large
- ✗ Extrêmement sensible au modèle utilisé
- ✗ Extrêmement sensible à la définition de la fonction de récompense
- ✗ Problèmes très difficile à résoudre sans approximations
- ✗ Multiplication de modèles spécifiques plus faciles à résoudre
- ➡ Prendre en compte la spécificité du problème dans la modélisation est indispensable

Questions

Merci de votre attention

???



Bibliography

Articles comme premier auteur



Camille Besse and Brahim Chaib-draa.

An Efficient Model for Dynamic and Constrained Resource Allocation Problems.

In Proc. of the 2nd Inter. Workshop on Constraint Satisfaction Techniques for Planning and Scheduling Problems, September 2007.



Camille Besse and Brahim Chaib-draa.

Parallel Rollout for Online Solution of Dec-POMDP.

In Proc. of 21st Inter. FLAIRS Conf., May 2008.



Camille Besse and Brahim Chaib-draa.

Quasi-Deterministic Partially Observable Markov Decision Processes.

In Proc. of 16th Inter. Conf. On Neural Information Processing, pages 237–246, December 2009.



Camille Besse and Brahim Chaib-draa.

Quasi-Deterministic POMDPs and Dec-POMDPs.

In Proc. of 9th Inter. Conf. On Autonomous Agents and MultiAgent Systems (Extended Abstract), May 2010.



Camille Besse and Brahim Chaib-draa.

Quasi-Deterministic POMDPs and Dec-POMDPs.

In Proc. of 5th Inter. Workshop On Multiagent Sequential Decision Making in Uncertain Domains, May 2010.



Camille Besse, Pierrick Plamondon, and Brahim Chaib-draa.

R-FRTDP. A Real-Time DP Algorithm with Tight Bounds for a Stochastic Resource Allocation Problem.

In Proc. of the 20th Canadian Conf. on Artificial Intelligence, May 2007.

Bibliography

Articles comme second auteur



Patrick Dallaire, Camille Besse, and Brahim Chaib-draa.

Learning Gaussian Process Models from Uncertain Data.

In *Proc. of 16th Inter. Conf. On Neural Information Processing*, pages 433–440, December 2009.



Patrick Dallaire, Camille Besse, and Brahim Chaib-draa.

Approximate Inference of Latent Functions from Uncertain Data with Gaussian Processes.

Neurocomputing, page To appear, 201X.



Patrick Dallaire, Camille Besse, Stéphane Ross, and Brahim Chaib-draa.

Bayesian Reinforcement Learning in POMDPs with Gaussian Processes.

In *Proc. of the Inter. Conf. on Intelligent Robots and Systems*, 2009.



Jean-Samuel Marier, Camille Besse, and Brahim Chaib-draa.

A Markov Model for Multiagent Patrolling in Continuous Time.

In *Proc. of 16th Inter. Conf. On Neural Information Processing*, pages 648–656, December 2009.



Jean-Samuel Marier, Camille Besse, and Brahim Chaib-draa.

Solving the Continuous Time Multiagent Patrol Problem.

In *Proc. of 2010 IEEE Inter. Conf. on Robotics and Automation*, 2010.