

# Modèle de Conception de SMA coopératifs par planification réactive

I. Chadès et F. Charpillet

*chades@loria.fr*

*charp@loria.fr*

# Contexte général

---

- Systèmes Multi-Agents (SMA) :
  - ❖ SMA : agents interagissant dans un environnement afin de résoudre collectivement une tâche.
  - ❖ Agent : entité dotée de capacités de perception, d'action et de prise de décision.
- Nature des problèmes :
  - ❖ Dynamiques,
  - ❖ Distribués,
  - ❖ Connaissances imparfaites (incertitudes).
- Méthodes de résolution :
  - ❖ Phénomènes de coopération,
  - ❖ Contrôle décentralisé,
  - ❖ Résolution distribuée.

# Problématique

---

- Objectif :
  - ❖ Formaliser la conception d'agents coopératifs aux comportements réactifs à l'exécution (stimuli/réponse).
  
- Requièrè :
  - ❖ Gèrer l'incertitude (connaissances, comportements);
  - ❖ Éviter les conflits ;
  - ❖ Coordonner les actions des agents.
  
- Moyens :
  - ❖ Modèles décisionnels de Markov (élaboration de plans « réactifs »);
  - ❖ Inspiration biologique.

# Systemes Multi-Agents Coopératifs

- **SMA Coopératifs =**
  - ❖ Agents coopérants
  - +
  - ❖ Environnement
  - +
  - ❖ Interactions.

- **Propriété essentielle :**
  - ❖ **Localité** (situation)
    - ❖ des perceptions,
    - ❖ des prises de décisions.

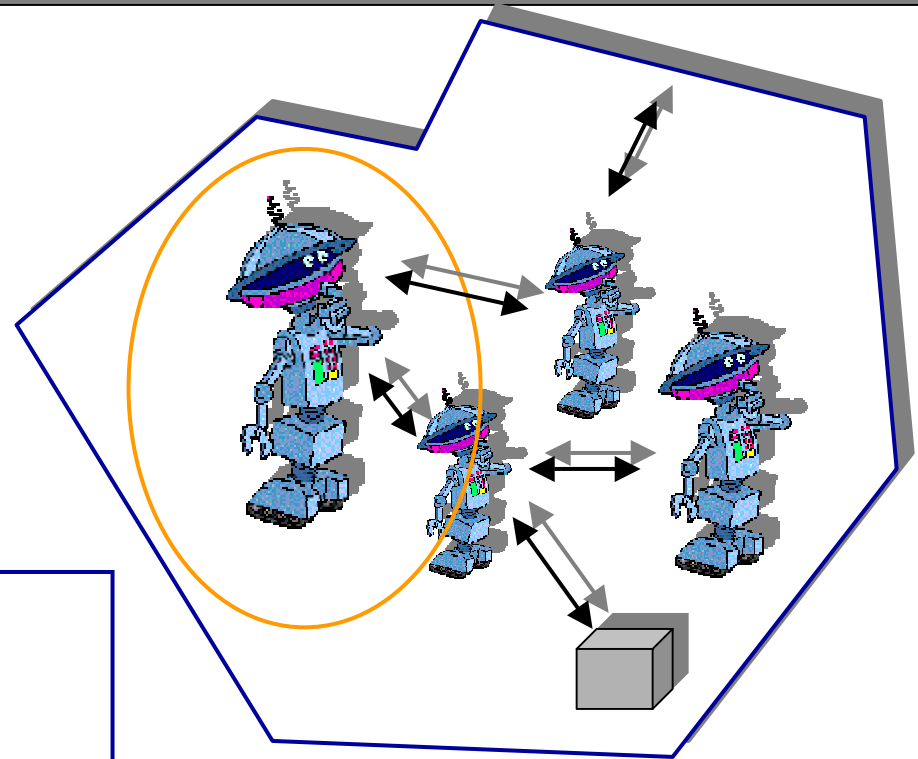


Figure 1 - SMA coopératifs

# Approches de conception

Dans la littérature ..

- Ascendante (SMA réactifs)
  - + Simplicité des mécanismes de résolution de chaque agent.
  - Choix des paramètres, formalisation.
  
- Descendante (SMA cognitifs)
  - + Adéquation, maîtrise.
  - Complexité des mécanismes de résolution.

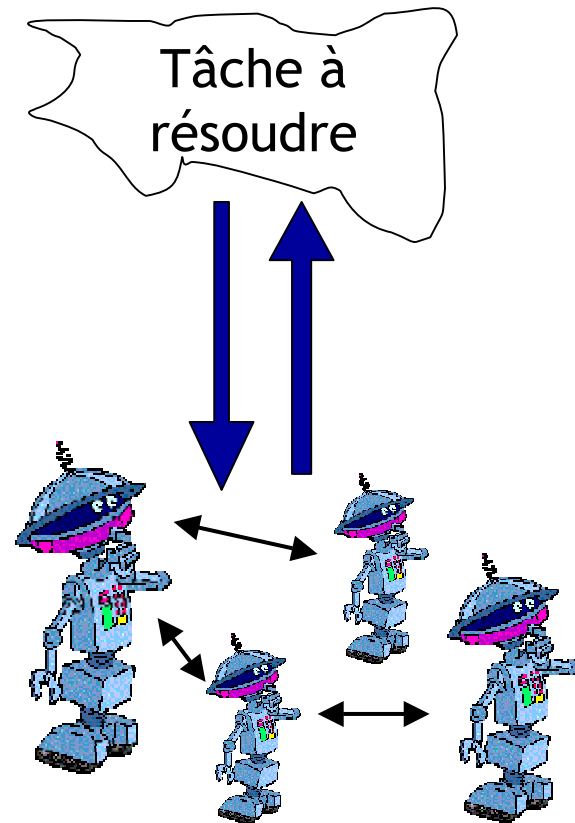


Figure 2 - Approche de conception

# Bilan

---

- Recherche d'une méthode de conception descendante d'agents coopératifs :
  - ❖ **Localité** (perception et prise de décision),
  - ❖ Incertitude,
  - ❖ Coordination des actions par élaboration de **plans réactifs**.
  
- **Modèles décisionnels de Markov** :
  - ❖ Calcul de plans réactifs, gestion de l'incertitude,
  - ❖ Résultats théoriques de convergence des algorithmes de résolutions.

# Processus Décisionnels de Markov (MDP ou PDM)

- MDP  $\langle S, A, T, R \rangle$ 
  - ❖  $S$  : ensemble fini d'états.
  - ❖  $A$  : ensemble fini d'actions.
  - ❖  $T : S \times A \rightarrow P(S)$
  - ❖  $R : S \times A \times S \rightarrow P(\mathbb{R})$

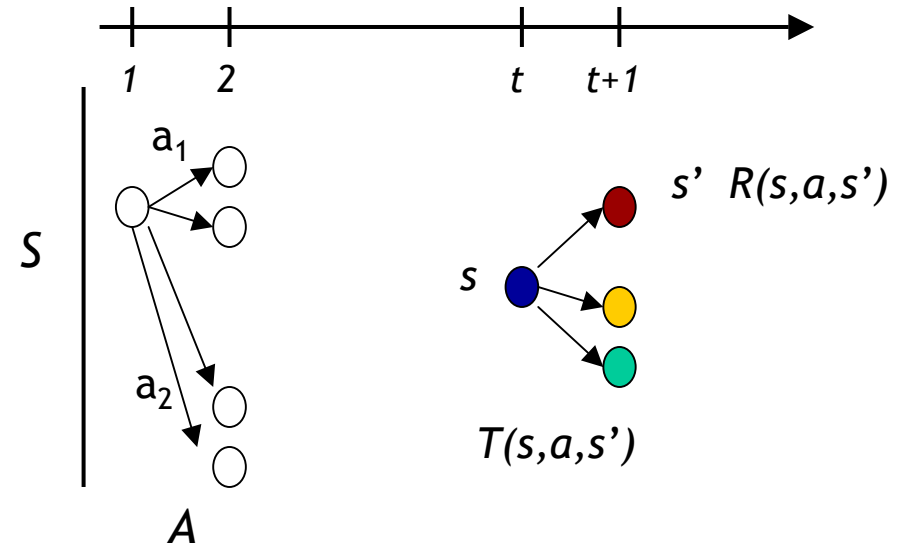


Figure 3 - Principe général d'un MDP

- Maximiser les récompenses

# Politique optimale

- Résoudre un problème MDP c'est trouver :
  - ❖  $\pi^* : S \rightarrow A$  telle que :
  - ❖  $\forall \pi, \forall s, V^{\pi^*}(s) \geq V^\pi(s)$
- Algorithmes de résolution [Puterman94] :
  - ❖ « *Value iteration* »
  - ❖ « *Policy iteration* »
- Complexité polynomiale (cas général) [Littman96] :
  - ❖  $|S|, |A|, B$ , et  $\gamma$



# Modèles décisionnels de Markov et complexités

Modèles	Nombre d'agents	Complexité	Implantation
MDP	1	Polynomiale	oui
MDP multi-agents (MMDP)	n	Polynomiale	Non, nombre d'agents > 3
POMDP	1	Non décidable	Résolutions approchées
DEC-POMDP DEC-MDP	n	Non décidable (NEXP à horizon fini)	X

Tableau 1 - Modèles décisionnels de Markov et complexités de résolution

# Bilan MDM

---

- Complexité des modèles décisionnels de Markov
  - ❖ MDP simplicité de résolution.
- Adapter le MDP
  - ❖ Localité ? Multi-agents ?
- Approximation d'un problème non markovien ?
  - ❖ Etude sur un phénomène biologique ... [Chades, 2003]
  - ❖ Résultats satisfaisants.

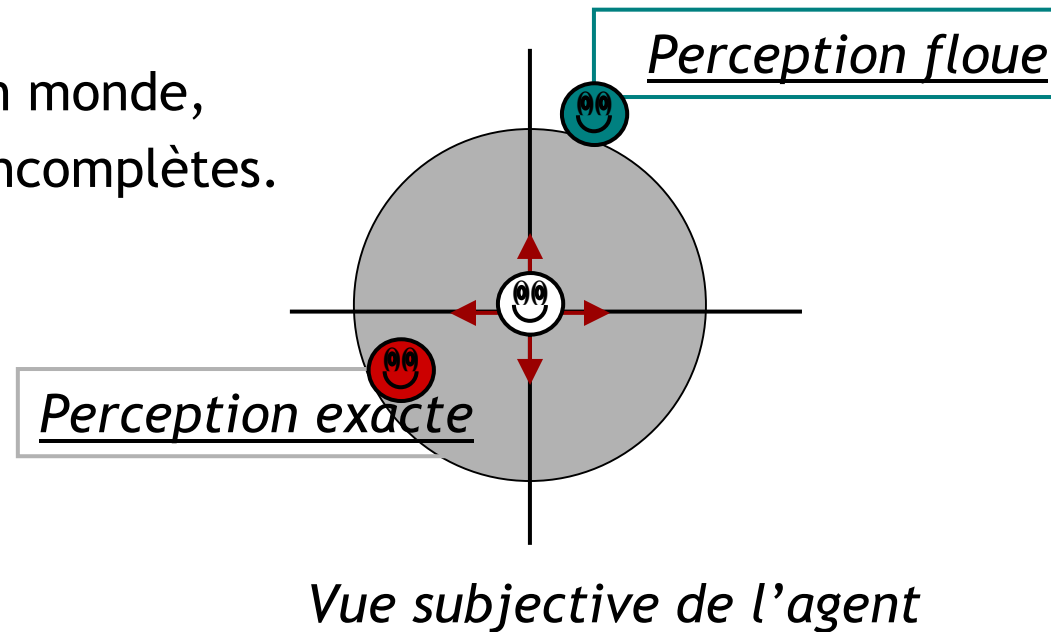
# Subjectivité et Empathie

---

- Notre modèle repose sur deux propriétés que nous allons exploiter :
  - ❖ **La subjectivité** : traite le problème de localité des perceptions et des prises de décisions.
  - ❖ **L'empathie** : traite le problème de la coordination d'actions.

# Subjectivité

- Subjectivité :
  - ❖ *Remplace le repère du système global d'un agent utilisant un MDP par un repère local.*
  - ❖ *Agent est au centre de son monde,*
  - ❖ *Perçoit des observations incomplètes.*
- MDP subjectifs



# MDP Subjectif

- MDP Subjectif :
  - ❖  $O_i$  ensemble de perceptions
  - ❖  $A_i$  ensemble des actions.
  - ❖  $T_i : O_i \times A_i \times O_i \rightarrow [0, 1]$
  - ❖  $R_i : O_i \rightarrow \mathbb{R}$
- Résolution ?
- Compromis entre précision et complexité

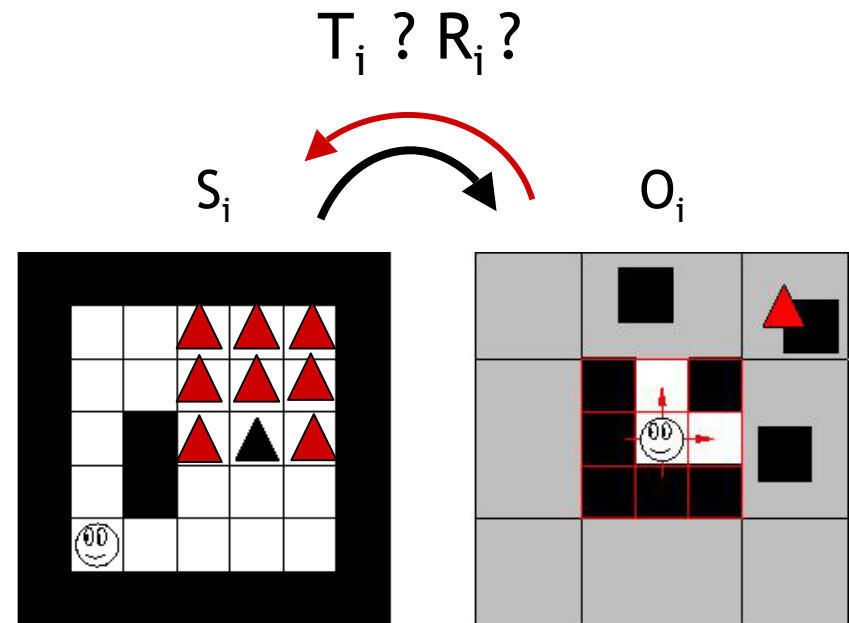


Figure 5 - *MDP subjectif*

# Effets de la subjectivité

- + Résolution d'un MDP de petite taille, complexité en temps constante.
- + Politique dédiée à un grand nombre d'environnements de topologies semblables.
- Politique peu adaptée aux environnements complexes (labyrinthe).

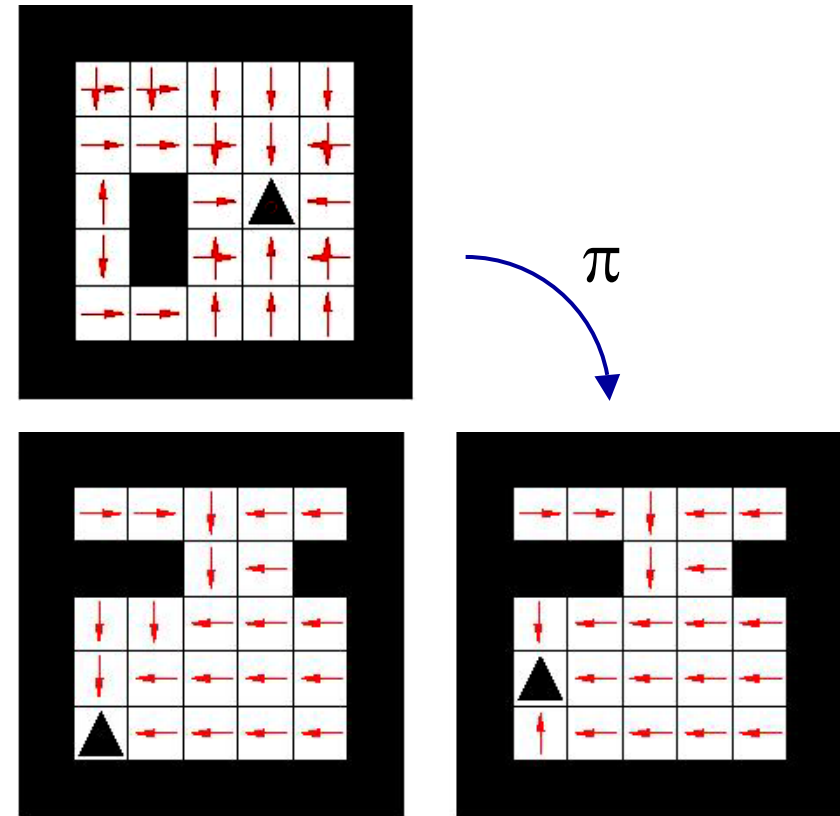


Figure 6 - Projection d'une politique subjective dans un environnement différent

# Empathie

## □ Empathie :

« *Habilité à percevoir, à identifier et à comprendre les sentiments ou émotions d'un autre [...] »* »

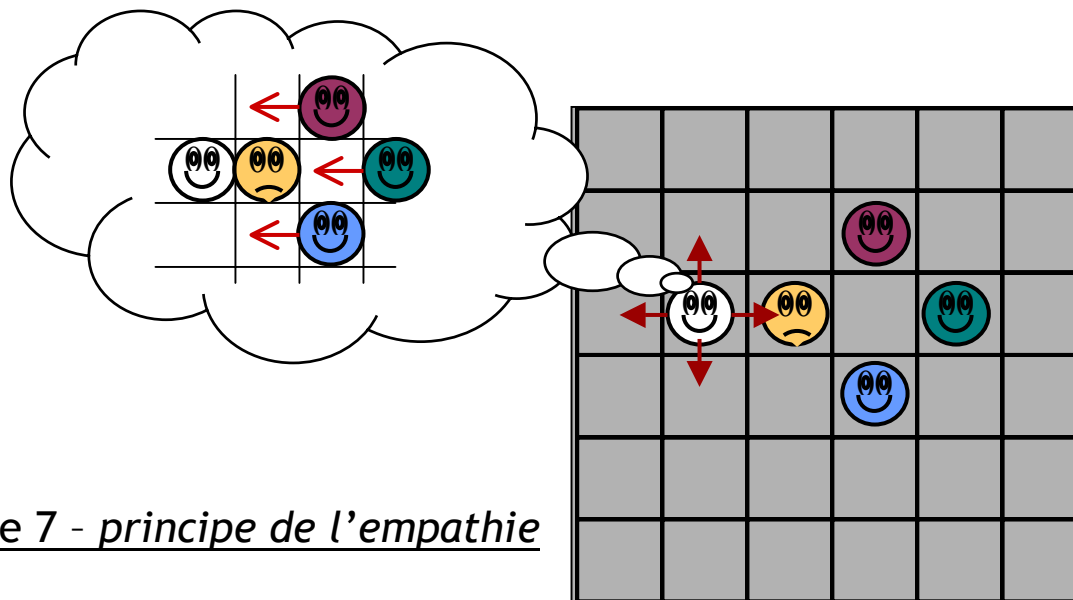


Figure 7 - principe de l'empathie

# Résoudre un MMDP : Co-évolution

- MDP multi-agent complètement observable :
  - ❖ Théorème 1 (stabilité de la solution optimale) [Chades et al, 2002]  
*Connaissant les politiques individuelles optimales de  $m$  agents, alors la résolution du MMDP comprenant les  $n-m$  agents donnera les politiques individuelles optimales restantes [...].*
  - ❖ Réduire la complexité de résolution.
  
- Exploitation du théorème ...
- Principe de co-évolution ...



# Algorithme itératif de Co-évolution alternatif

Entrée :  $\Pi^0 = (\pi_1^0, \dots, \pi_n^0)$

$t=0$ ;

**Répéter**

$m = \text{random}(n)$ ;

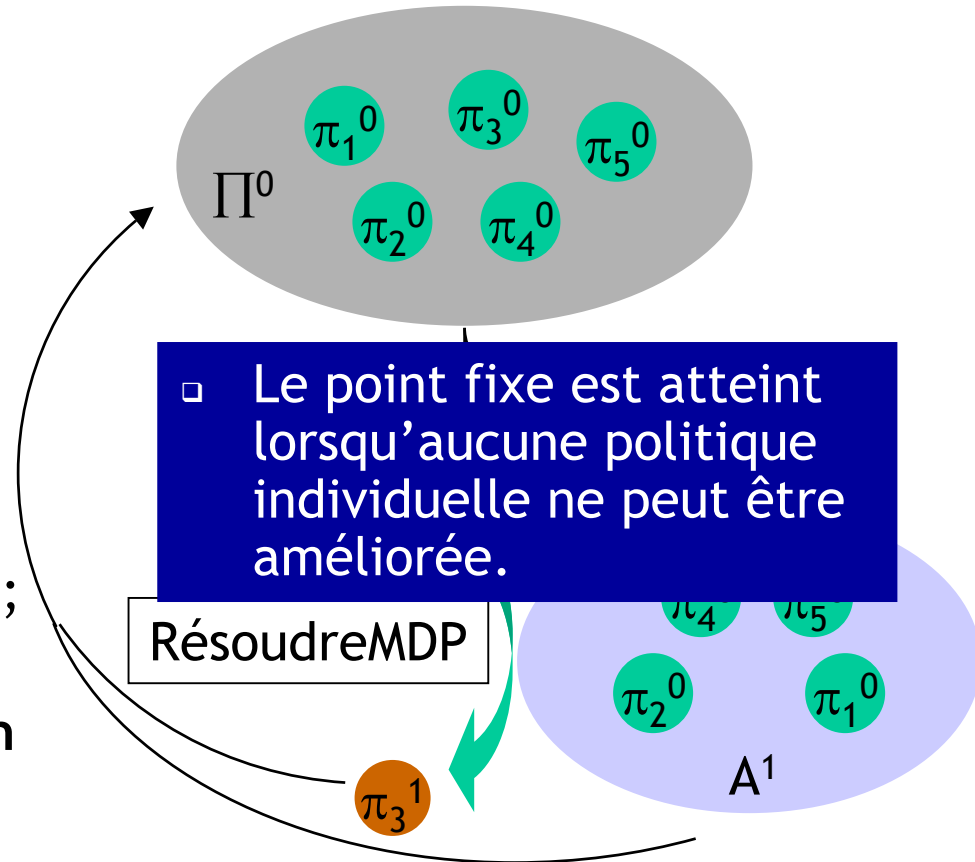
$A^{t+1} = (\pi_1^t, \dots, \pi_{m-1}^t, \pi_{m+1}^t, \dots, \pi_n^t);$   
 $\pi_1^t, \pi_{m+1}^t, \dots, \pi_n^t$ ;

$b^t = \pi_m^t$ ;

$b^{t+1} = \text{RésoudreMDP}(A^{t+1}, b^t)$ ;

$t=t+1$ ;

**Jusqu'à Convergence vers un point fixe qui est un équilibre de Nash (théorèmes 2 et 3).**



# Algorithme itératif de Co-évolution simultané

Entrée :  $\Pi^0 = (\pi_1^0, \dots, \pi_n^0)$

$t=0$ ;

**Répéter**

$m = \text{random}(n)$ ;

$A^{t+1} = (\pi_1^t, \dots, \pi_{m-1}^t, \pi_{m+1}^t, \dots, \pi_n^t)$ ;

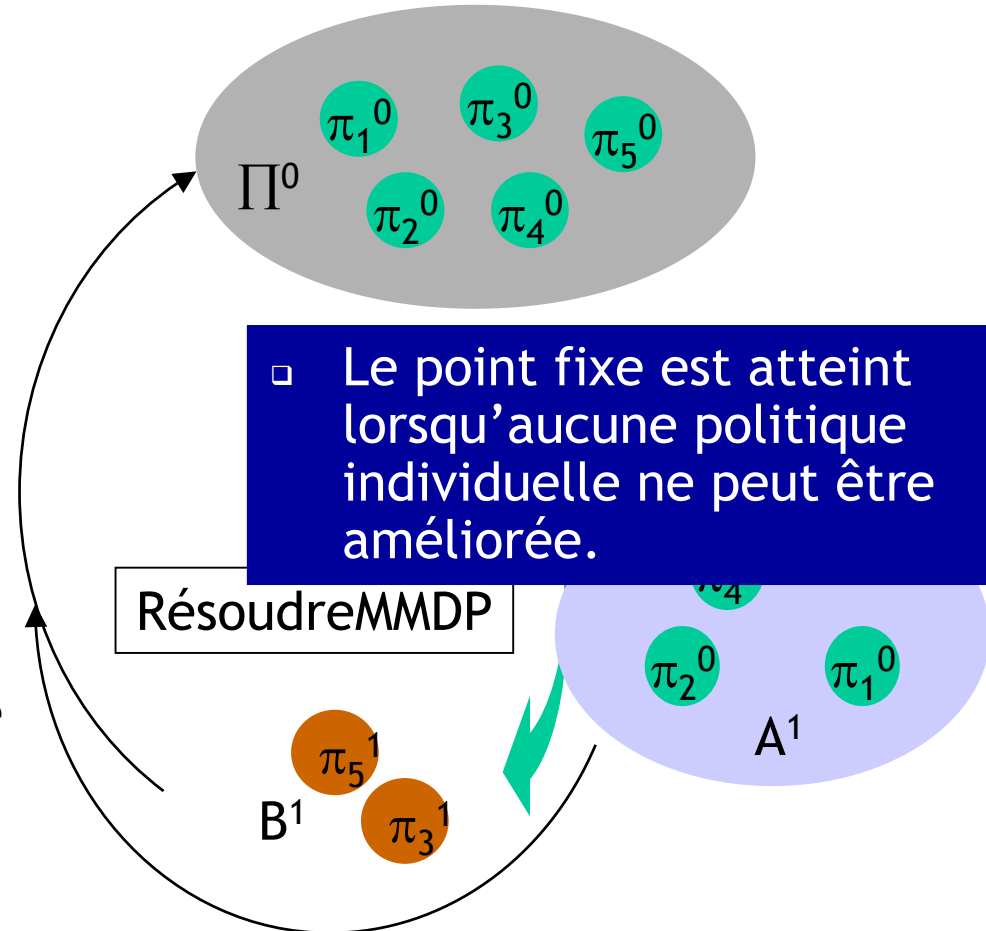
$B^t = \pi_m^t$ ;

$B^{t+1} = \text{RésoudreMMDP}(A^{t+1}, B^t)$ ;

$t=t+1$ ;

**Jusqu'à Convergence vers un point fixe qui est un équilibre de Nash (théorèmes 2 et 4).**

Sortie :  $\Pi^t = (\pi_1^t, \dots, \pi_n^t)$



# Analyses des algorithmes

---

- L'algorithme simultané :
  - ❖ Apte à sortir des optimums locaux.
  - ❖ En moyenne converge en moins d'itérations.
  - ❖ Limité par la complexité d'un MMDP.
- L'algorithme alternatif :
  - ❖ faible complexité de la fonction *résoudreMDP*.
- Coordination implicite :
  - ❖ récompense globale.
- Intérêt de l'aléatoire :
  - ❖ Donne la possibilité de lancer plusieurs fois l'algorithme et d'obtenir des équilibres de Nash de meilleurs qualités.

# Conception de SMA

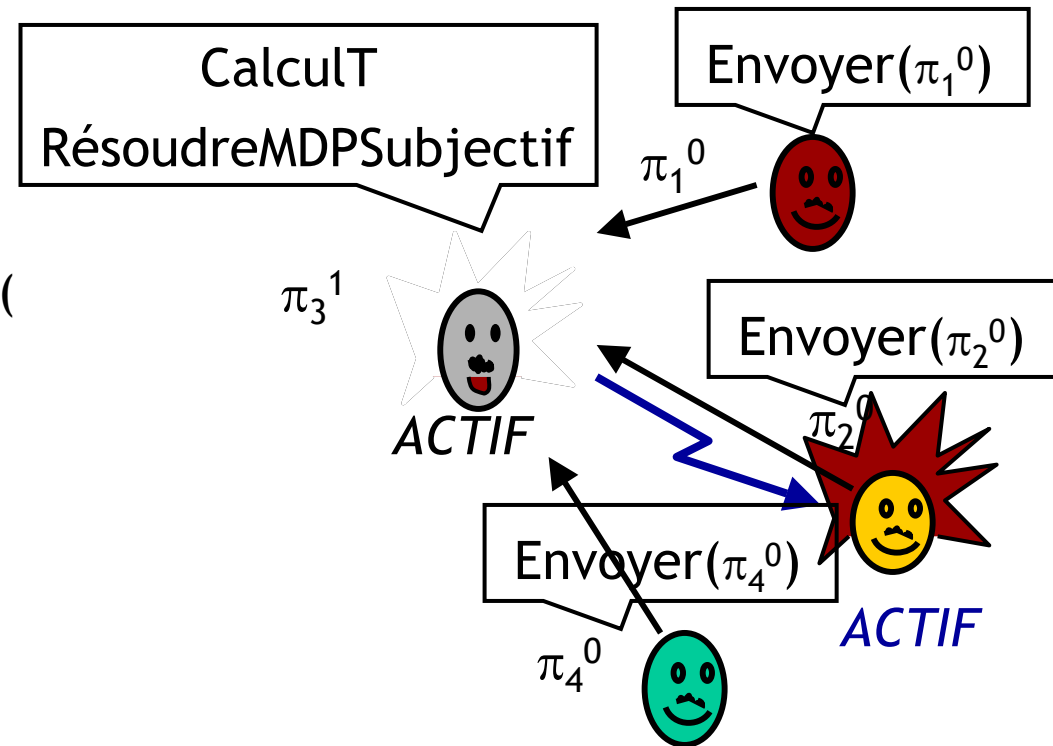
## Subjectivité et Empathie

---

- Adaptation de l'algorithme alternatif
- Utilisation de l'algorithme simultané nécessite une connaissance parfaite de l'environnement :
  - ❖ résoudre MMDP subjectif ?
- Conception décentralisée
  - ❖ On-line : agents communiquent leur plan intermédiaire.
- Conception centralisée
  - ❖ Off-line : adaptation directe.

# Algorithme décentralisé de conception alternative

t=0;  
Répéter  
  Si ACTIF Alors  
    Recevoir( $\{\Pi^t\} \setminus \pi_i^t$ );  
     $\pi_i^t = \text{RésoudreMDPSubjectif}(\text{CalculT}_i(\{\Pi^t\} \setminus \pi_i^t))$ ;  
  Sinon Envoyer( $\pi_i^t$ );  
  Fin Si  
  t=t+1;  
Jusqu'à convergence vers un point fixe ou t=Tmax



# Synthèse

---

- Famille d'algorithmes de conceptions :
  - ❖ centralisés ou décentralisés;
  - ❖ Populations hétérogènes/homogènes [*Chadès, 2003*];
  - ❖ Avec/Sans communication.
- Qualité des plans réactifs conçus :
  - ❖ Qualité des informations disponibles (limiter les ambiguïtés);
  - ❖ Résolution du MDP subjectif sera meilleure;
  - ❖ Convergence (éventuelle vers équilibre de Nash);
- Agents réactifs à l'exécution.

# Conclusions

---

- Modèle pour la conception de SMA coopératifs à l'aide de MDP :
  - ❖ Subjectivité : localité des perceptions des agents, étude de l'intérêt d'un MDP subjectif;
  - ❖ Empathie : coordination des agents, algorithmes de co-évolution simultanés et alternatifs (preuve de convergence - Nash);
  - ❖ Familles d'algorithmes de conception descendante de SMA aux comportements réactifs à l'exécution (preuve de convergence).
  - ❖ Coordination effective (expérimentations [Chades, 2003])
- Heuristique pour résoudre un DEC-POMDP (NEXP)

# Perspectives

---

- Comment répartir une récompense collective en récompenses individuelles ?
  - ❖ Garder les propriétés : « coopération » ...
- Automatiser la conception de SMA (compétitifs...)
  - ❖ Théorie des jeux
- SMA réactifs de conception ascendante et limites de la modélisation des MDP :
  - ❖ Adéquation : catégorie de problèmes en commun;
  - ❖ Et les autres ... ?



# Faim..

---

# Bibliographie

---

- [Puterman94]
  - ❖ L. Puterman. *Markov Decision Processes*. J. Wiley & Sons, 1994.
- [Littman96]
  - ❖ M. Littman. *Algorithms for Sequential Decision Making*. Ph. D Thesis, Department of Computer Science, Brown University, 1996.