

Réunion d'équipe - 14 décembre 2012

Laëtitia Matignon

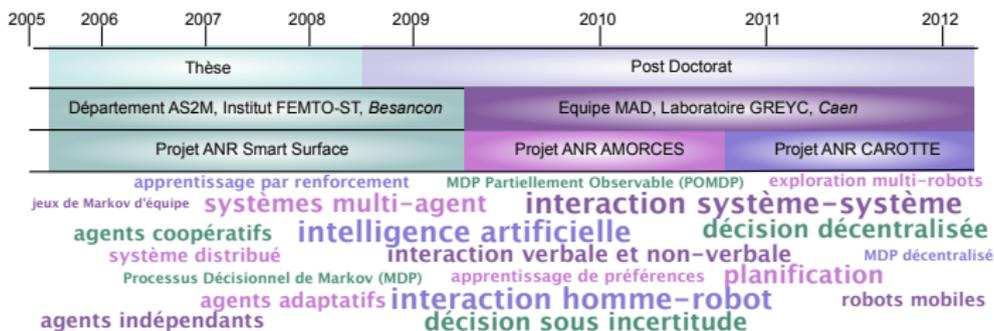
LIRIS - Équipe GrAMA
Université Claude Bernard Lyon 1



Plan

- 1 Thématiques
- 2 Contrôle adaptatif et décentralisé d'un système distribué
- 3 Coordination décentralisée d'agents décisionnels avec informations incomplètes
- 4 Modèle hybride de coopération homme-robot pour l'interaction verbale et non-verbale
- 5 Perspectives de recherche

Cursus



2005 : Diplôme d'ingénieur de l'ENSMM (*Besançon*)

École Nationale Supérieure de Mécanique et des Microtechniques

2005-2008 : Thèse de Doctorat : *Synthèse d'agents adaptatifs et coopératifs par apprentissage par renforcement - Application à la commande d'un système distribué de micromanipulation* directrice : N. Le Fort-Piat (Pr. 61eme FEMTO-ST), co-directeur : G. Laurent (MCF 61eme FEMTO-ST), rapporteurs : A.-I. Mouaddib (Pr. 27eme GREYC), J. Quinqueton (Pr. 27eme LIRMM)

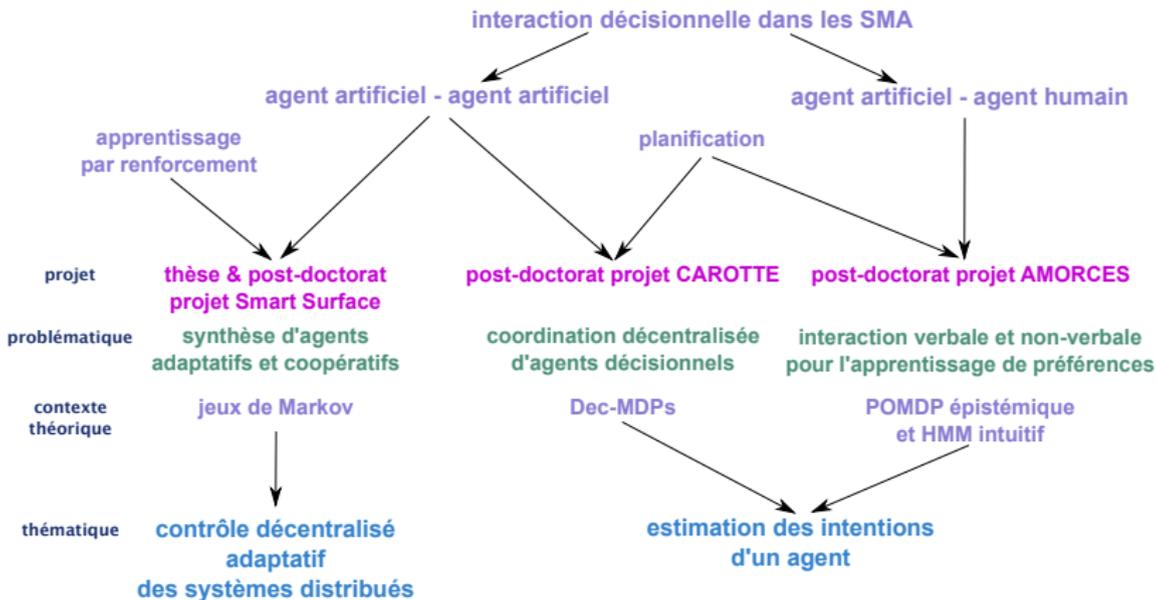
2008-2009 : Post-Doctorat - projet ANR *Smart Surface* - sous la direction de N. Le Fort-Piat.

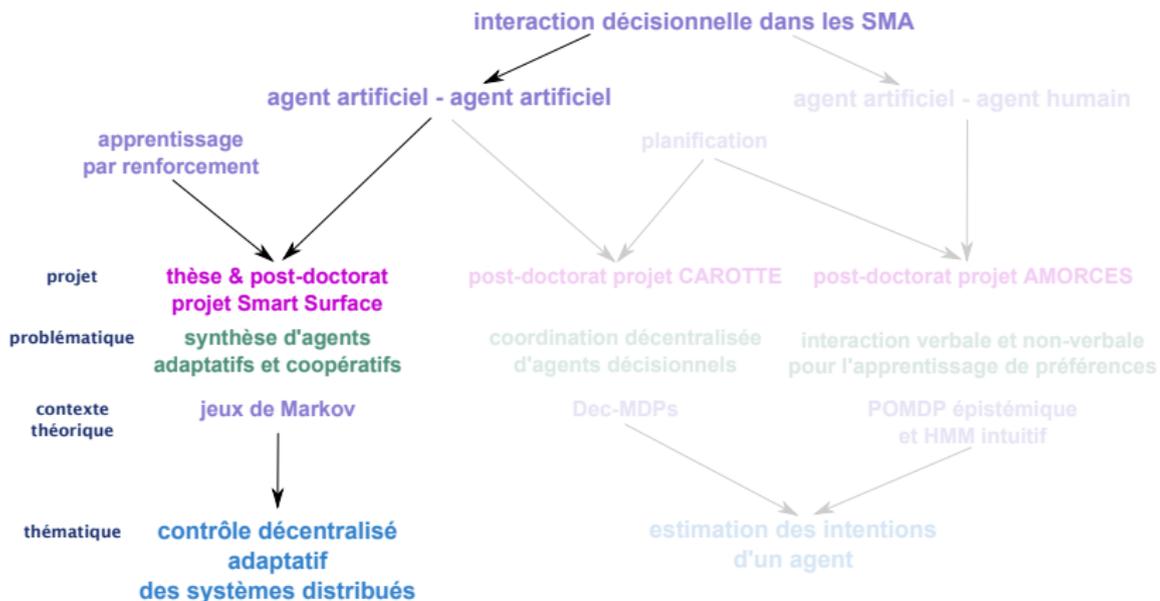
2009-2010 : Post-Doctorat - projet ANR *Algorithmes et MOdèles pour un Robot Collaboratif Eloquent et Social : Impact de la communication verbale et non-verbale sur la réalisation de missions coopératives* - en collaboration avec A.-I. Mouaddib et A.-B. Karami.

2010-2012 : Post-Doctorat - projet ANR *CArtographie par ROBoT d'un TErritoire : Stratégies d'exploration multi-robot* - en collaboration avec A.-I. Mouaddib et L. Jeanpierre.

Activités de recherche : Thématiques

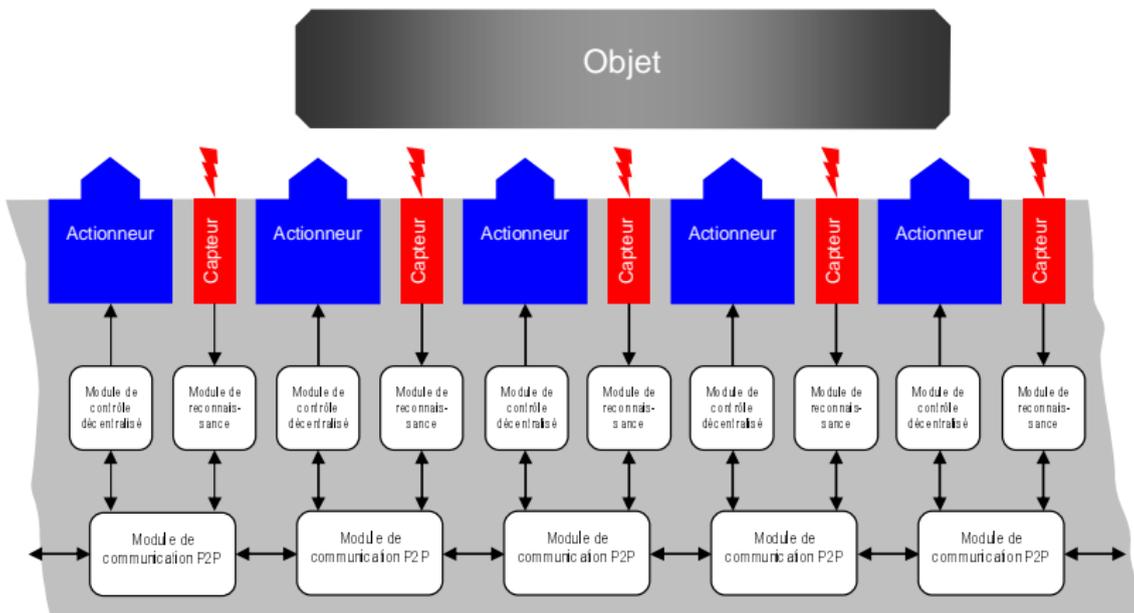
Thématique générale : Intelligence artificielle, prise de décision sous incertitude, applications à la robotique, modèles dérivés des modèles de Markov



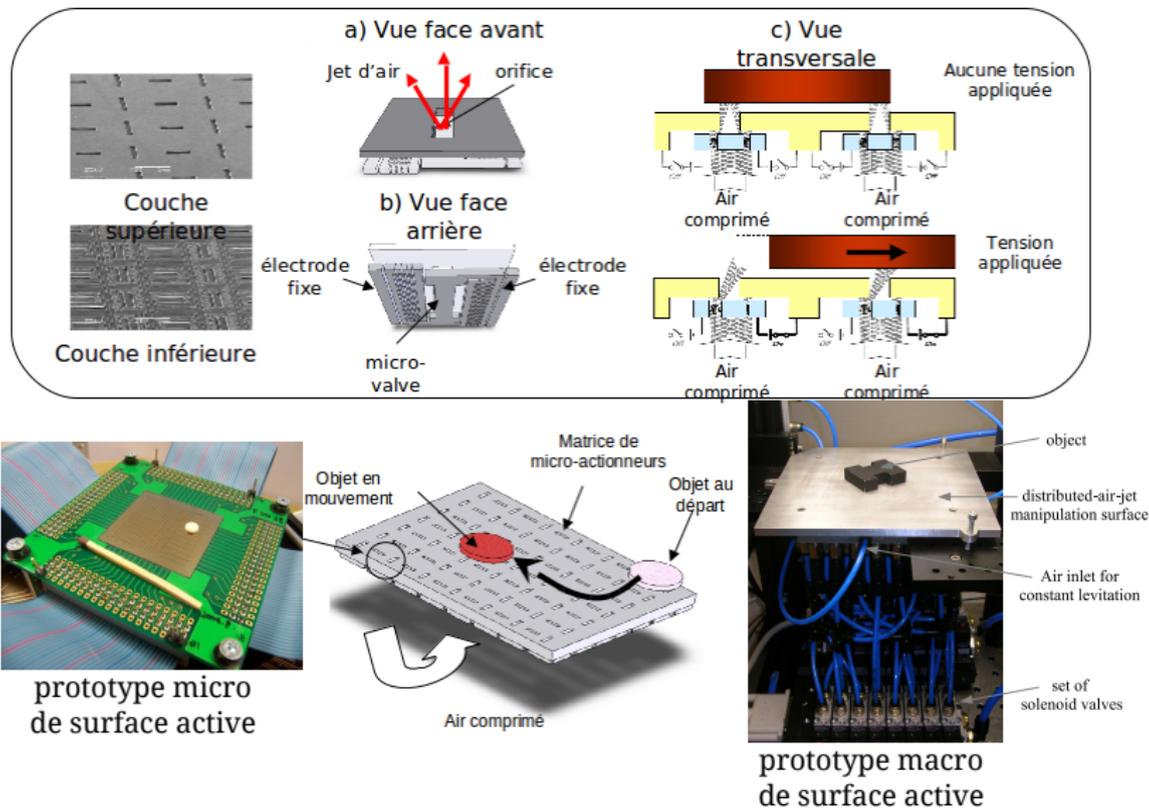


Contexte : projet ANR Smart surface

- Conception, développement et contrôle d'un système microrobotique distribué pour le convoyage, le positionnement et le tri de micropièces à l'échelle mésoscopique (μm au mm)
- Système distribué de micromanipulation totalement intégré



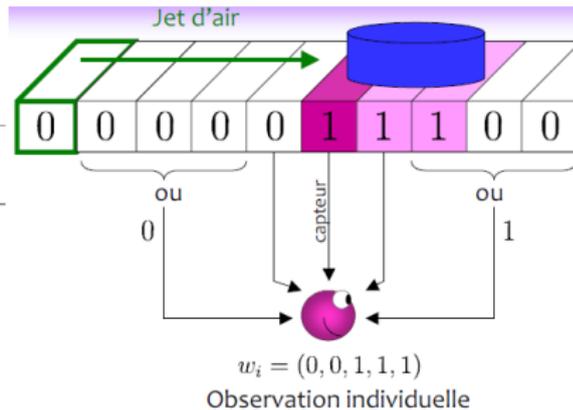
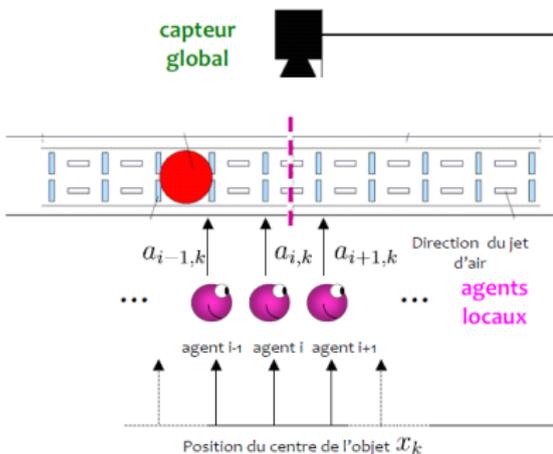
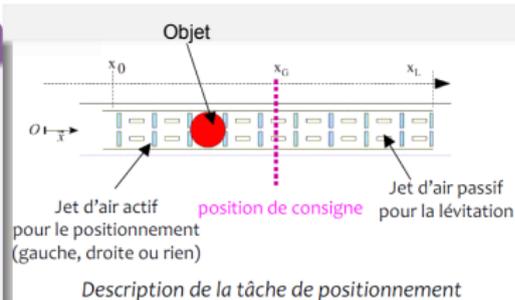
Contexte : projet ANR Smart surface

Surface active pneumatique [Fukuta *et al.* 06] [Chapuis *et al.* 08]

Objectifs : Contrôle adaptatif décentralisé par apprentissage par renforcement dans les jeux de Markov d'équipe

Hypothèses

- dynamique du système inconnue
- agents indépendants
- agents coopératifs
- observabilité individuelle totale / partielle (voisinage)



Résultats

Etude des enjeux de la coordination

- identification de facteurs de non-coordination : sélection d'équilibre, équilibres cachés, observations bruitées
- analyse des algorithmes existants/synthèse de l'état de l'art face à ces facteurs

Nouvelles méthodes

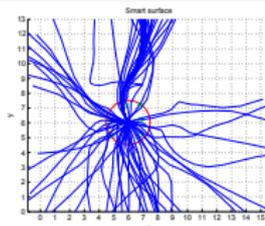
- Q-learning hystérique fondé sur des agents à *tendance optimiste réglable*
- SOaN fondé sur des agents s'adaptant automatiquement à la stochasticité du système (robustesse face à l'exploration des autres agents, réglage simple des paramètres)

$$Q_{i,SOaN}(s, a_i) \leftarrow [1 - G_i(s, a_i)]Q_{i,moy}(s, a_i) + G_i(s, a_i)Q_{i,max}(s, a_i)$$

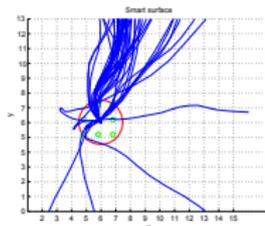
Résultats

Contrôle adaptatif décentralisé d'un système distribué de micromanipulation

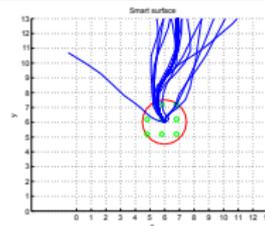
- Jeux de Markov étendus aux macro-actions (observations contrôlées et non-contrôlées)
- Méthode générique d'extension de nos algorithmes dans le cas de systèmes à actions localisées
- Application du SOaN à une tâche de convoyage par commande décentralisée



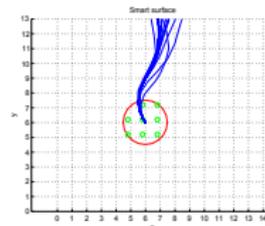
épisode 1 à 100



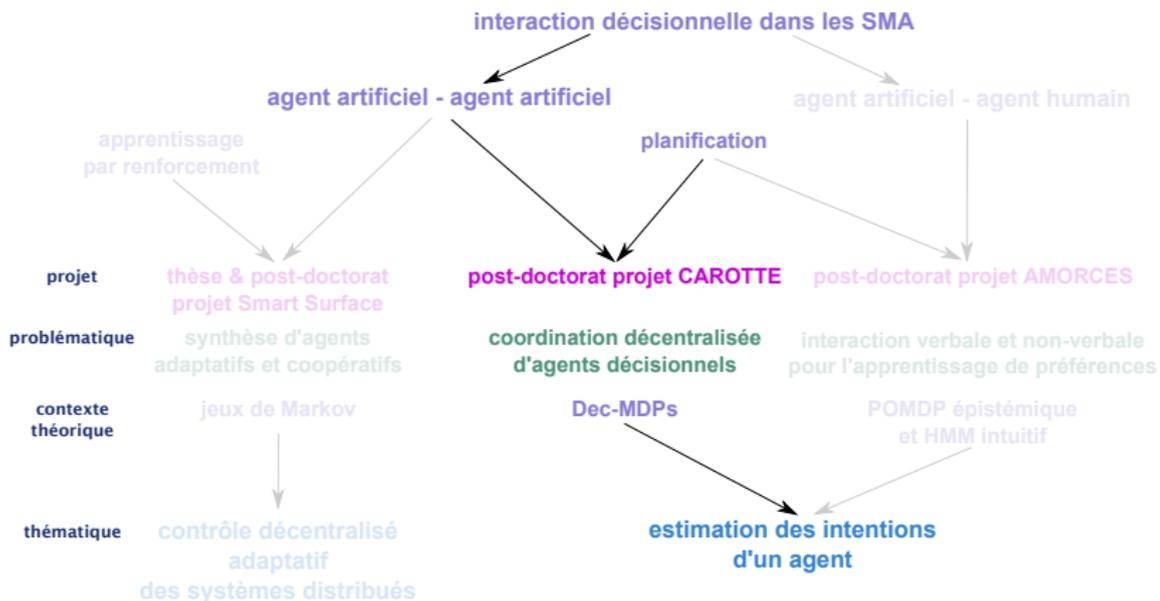
épisode 200 à 300



épisode 100 à 200



épisode 300 à 400



Contexte : défi robotique CAROTTE

5 équipes



PACOM



YOJI



CARTOMATIC



ROBOTS_MALINS



COREBOTS

Objectif

- système robotisé autonome pour la cartographie et l'exploration d'un environnement dynamique et inconnu

Problématiques du défi

- mobilité
- localisation et cartographie
- **décision** : coordination pour l'exploration
- détection d'objets
- **contraintes de communication** (limitée, coupures)

équipe ROBOTS_MALINS (GREYC-CNRS, THALES, INRIA groupe evolution)

- **système multi-robot décentralisé** pour l'exploration, la cartographie et la détection d'objets

Objectifs

Modèle de décision pour l'exploration multi-robot

Hypothèses

- robots indépendants
- pas de station centrale
- SLAM distribué (carte fusionnée et localisation) : observabilité locale totale
- communication entre robots limitée (partage des positions respectives)
- reconnaissance "au fil de l'eau"
- coupures de communication potentielles

Développer des stratégies d'exploration multi-robot sous contraintes de communication

Coordination décentralisée d'agents décisionnels :

- **coordination globale** : allocation des buts d'exploration (couverture efficace de la zone, minimiser les recouvrements).
- **interactions locales** : à minimiser car exploration non efficace (recouvrements) et conflits nécessitant une coordination locale.

Résolution orientée-interaction des Dec-(PO)MDPs

Etat de l'art

Dec-(PO)MDP [Bernstein,2000] $\langle n, S, A, T, R, \Omega, O \rangle$

- n nombre d'agents, S ensemble d'états joints, $A = \{A_1, \dots, A_n\}$ ensemble d'actions jointes
- $T : S \times A \times S \rightarrow [0; 1]$ fonction de transition
- $R : S \times A \rightarrow \mathfrak{R}$ fonction de récompense
- Ω ensemble d'observations et $O : S \times A \times S \times \Omega \rightarrow [0; 1]$ fonction d'observation

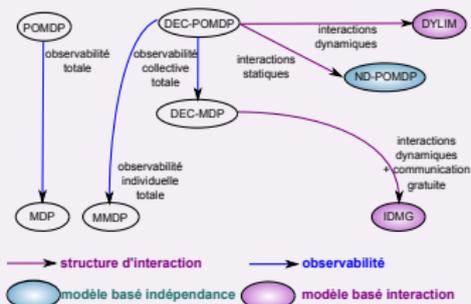
Résolution orientée-interaction des Dec-(PO)MDPs

Etat de l'art

Résolution optimale

- NEXP complete (Bernstein et al. 2002)
- **hypothèse de dépendances totales** : interaction permanente entre tous les agents

Modèles orientés interaction



- exploiter les interactions locales pour résoudre le modèle
- relacher l'hypothèse d'interactions permanentes entre tous les agents
- ND-POMDP (Nair et al. 2005), OC-DEC-MDP (Beynier & Mouaddib 2006), IDMG (Spaan & Melo 2008), DEC-SIMDP (Melo & Veloso 2011), DyLIM (Canu & Mouaddib 2011)

Résolution orientée-interaction des Dec-MDPs

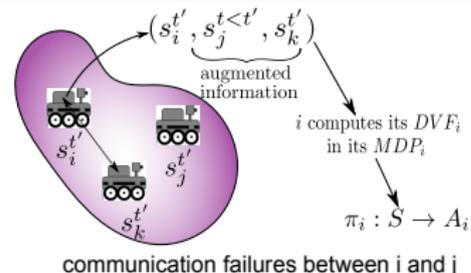
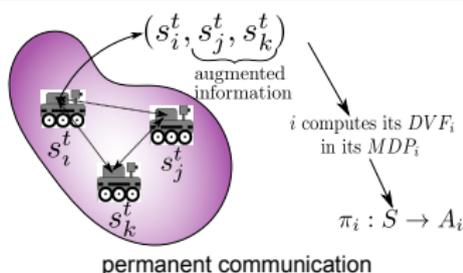
MDP augmenté avec DVF

Motivations

- 1 applications réelles
- 2 coordination globale en minimisant les interactions locales

Notre approche

- 1 MDP augmenté : décomposition du problème en un ensemble de MDPs locaux $\{MDP_1, \dots, MDP_n\}$, un par agent. **Résoudre une collection de MDPs**
- 2 Distributed Value Function (DVF) : chaque robot calcule sa DVF_i dans son MDP_i en utilisant l'information augmentée afin de minimiser les interactions



Résolution orientée-interaction des Dec-MDPs

MDP augmenté avec DVF

DVF pour un agent i dans $MDP_i = \langle S_i, A_i, T_i, R_i \rangle$

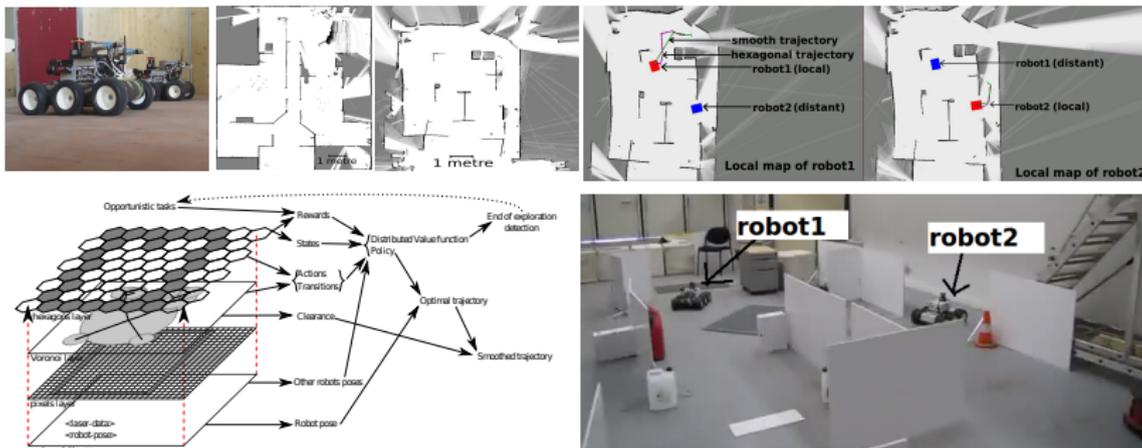
$$DVF_i(s_i) = \max_{a_i \in A_i} \left(R_i(s_i, a_i) + \gamma \sum_{s' \in S_i} T_i(s_i, a_i, s') [DVF_i(s') - \sum_{j \neq i} f_{ij} Pr(s' | s_j, \Delta t_j) V_j(s')] \right)$$

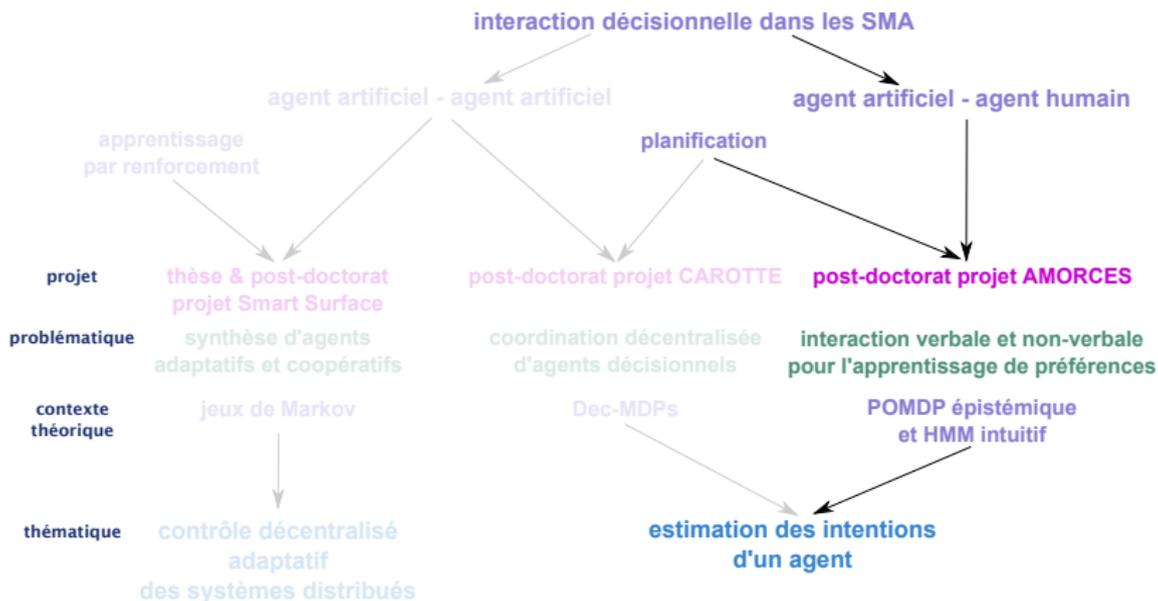
- minimise les interactions entre agents
- $Pr(* | s_j, \Delta t_j) V_j(*)$ prédiction des intentions d'un agent

DVF appliqué à l'exploration multi-robot

- prédiction des intentions d'un agent avec *wavefront propagation algorithm* et empathie
- coordination globale qui minimise les interactions entre agents
- coordination locale résolue avec politique jointe calculée *off-line*
- environnement dynamique : mise à jour du modèle/politique à chaque pas de décision, approche greedy

CAROTTE : Planification de stratégies d'exploration multi-robot





Contexte : projet ANR AMORCES

- un robot compagnon doit se rendre utile (aider ou assister) et avoir des compétences sociales.
- types d'interactions (Karami, 2011) :



(Cirillo et.al., 2009)

Coopération :

- Réalise une tâche à la place de l'humain
- Réalise la tâche seul
- Veille à ne pas gêner l'humain et **respecter ses préférences**

Assistance :

- Assiste l'humain en le guidant/l'informant
- Détecte le besoin d'assistance
- L'humain réalise la tâche



(Hoey et.al., 2010)



(Karami et.al., 2011)

Collaboration :

- Une participation conjointe
- Détecte le niveau d'engagement de l'humain

Objectifs

Hypothèses : mission coopérative homme-robot

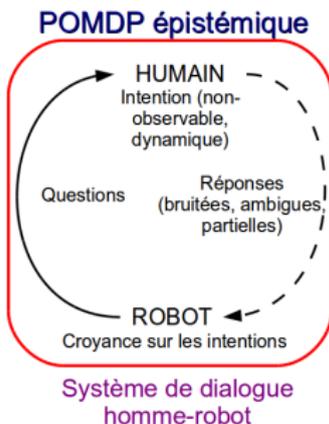
- Un ensemble de n tâches $K = \{k_1, k_2, \dots, k_n\}$.
- Le robot et l'humain forment une équipe, chacun pouvant réaliser une des tâches de la mission.
- Le robot doit réaliser ces tâches en accord avec les préférences de l'humain sur celles-ci.
- Les préférences de l'humain sur les tâches sont dynamiques.

Estimation des intentions d'un agent humain

- Inférence par le robot des préférences de l'humain pour améliorer sa coopération avec l'humain.
- Interaction verbale et non-verbale pour l'apprentissage des préférences dans le cadre de missions coopératives.

Reconnaissance de l'intention explicite - verbale - épistémique

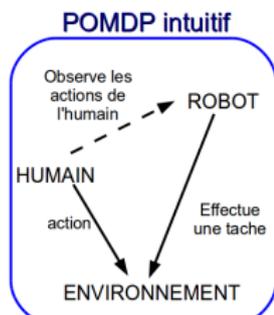
Systèmes de dialogue



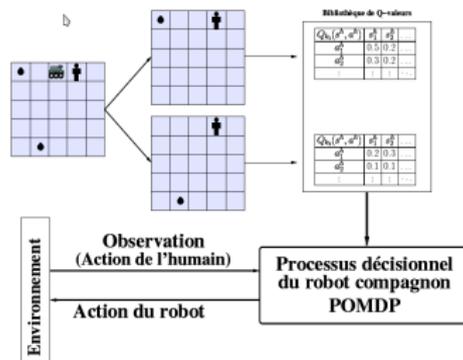
- décider la séquence de questions à poser pour désambiguïser au mieux l'intention de l'humain.
- interpréter les réponses conflictuelles (erreurs de reconnaissance ou changement d'intention ?).
- minimiser les questions (Rosenthal & Veloso, Ro-MAN, 2011).
- approches POMDP (Young et. al.,2010) ou réseaux bayésiens (Thomson et.al.,2008).
- détection d'un changement d'intention couteux et pénible pour l'humain.

Reconnaissance de l'intention implicite - non-verbale - intuitive

Observations des actions



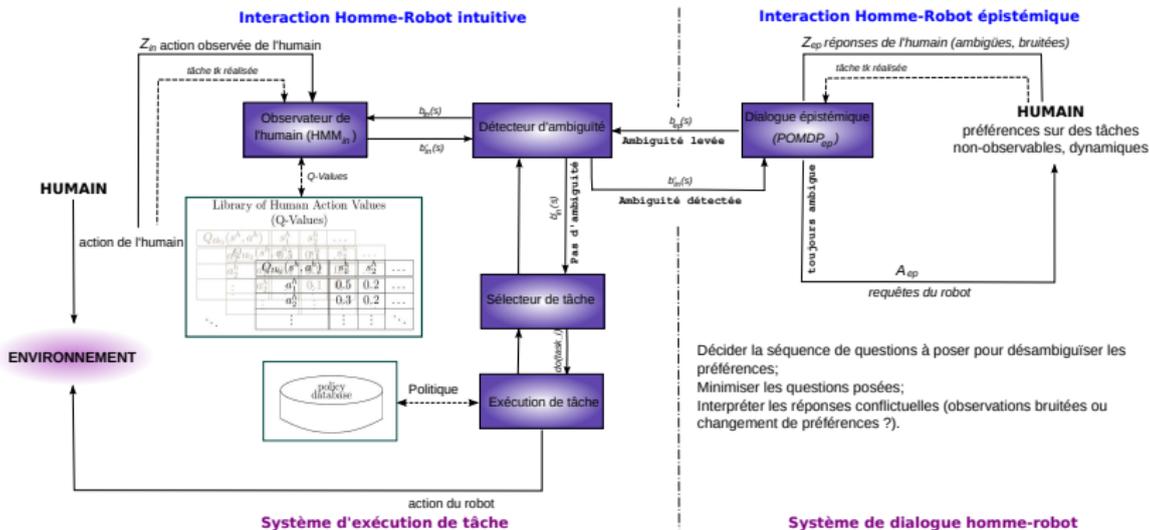
Interaction homme-robot intuitive
pour une mission collaborative



POMDP Augmenté (Karami et.al., HRI, 2010)

- associer une action observée de l'humain à une intention possible de celui-ci.
- les actions des humains sont des conséquences directes de leurs intentions.
- approches POMDPs (Broz et. al., Ro-MAN, 2011)(Karami et.al., HRI, 2010)

Modèle hybride d'interactions verbales et non-verbales pour l'inférence des préférences cachées et dynamiques



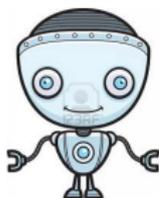
- Alternance entre le dialogue avec l'humain (*inférence épistémique*) et l'exécution/observation de l'humain (*inférence intuitive*).

- Switch de l'interaction épistémique à intuitive :
 $\forall s_{ep} \in \mathcal{S}_{ep}, \forall k \in K, b_{ep}(s_{ep}(k) = \text{unknown}) < \epsilon.$

- Détecteur d'ambiguïté : $s_{in}^* = \arg \max_{s_{in} \in \mathcal{S}_{in}}(b_{in}(s_{in}))$, $s_{ep}^* = \arg \max_{s_{ep} \in \mathcal{S}_{ep}}(b_{ep}(s_{ep}))$

Perspectives

- exploration multi-robot avec perception active, modèle hiérarchique leader/suiveur
- enrichir/améliorer le **modèle de préférences** du modèle hybride d'interaction verbale et non-verbale (apprendre un CP-net représentant le modèle cognitif de l'agent humain ?),
- développement d'un **modèle général pour l'interaction décisionnelle** qui considère l'observabilité partielle, des interactions dynamiques et un protocole général de communication, ainsi que sa résolution *via* DVF,
- intelligence collective : étude/comparaison de différentes méthodes issues des MDPs et de la robotique collective pour les mêmes cadres applicatifs (liens entre différentes stratégies d'exploration multi-robot (MinPosition, DVF, ...)),
- ...



Merci !

- Bernstein, D. S., Givan, R., Immerman, N. & Zilberstein, S. (2002). The complexity of decentralized control of markov decision processes, *Math. Oper. Res.* 27 : 819–840.
- Beynier, A. & Mouaddib, A.-I. (2006). An iterative algorithm for solving constrained decentralized markov decision processes, *Proc. of AAAI*, pp. 1089–1094.
- Canu, A. & Mouaddib, A.-I. (2011). Collective decision- theoretic planning for planet exploration, *Proc. of ICTAI*.
- Melo, F. S. & Veloso, M. M. (2011). Decentralized mdps with sparse interactions, *Artif. Intell.* 175(11) : 1757–1789.
- Nair, R., Varakantham, P., Tambe, M. & Yokoo, M. (2005). Networked distributed pomdps : A synthesis of distributed constraint optimization and pomdps, *Proc. of AAAI*, pp. 133–139.
- Spaan, M. T. J. & Melo, F. S. (2008). Interaction-driven markov games for decentralized multiagent planning under uncertainty, *Proc. of AAMAS*, pp. 525–532.