Decentralized Multi-Robot Planning to Explore and Perceive

Laëtitia Matignon[1], Laurent Jeanpierre[2] and Abdel-Illah Mouaddib[2]

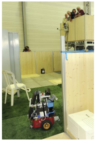
[1] LIRIS/CNRS UMR5205 - University of Lyon 1, France

[2] GREYC/CNRS UMR6072 - University of Caen Basse-Normandie, France

ECAI 2014 Workshop on Multi-Agent Coordination in Robotic Exploration

“CAROTTE” Robotic challenge

5 selected teams



PACOM

YOJI

CARTOMATIC

ROBOTS_MALINS

COREBOTS

Objective

- autonomous robotic system for mapping and exploration of an unknown and dynamic environment with unmanned ground vehicle(s) (UGV(s))

Issues of the challenge

- mobility
- localisation and mapping (SLAM)
- decision
- object detection and localisation
- communication constraints

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

- decentralized multi-robot system for exploration, mapping and object detection

Decentralized multi-robot planning to explore and perceive

Assumptions

- independant robots
- no central base station
- distributed SLAM (localization and shared map): full local observability
- limited communication between robots (share only their localization)

Outline

1. Decision model for multi-robot exploration based on decentralized Markov decision processes (Dec-MDPs)
2. Active perception combined with exploration

Markov decision processes

Decision-theoretic models based on Decentralized MDPs provide an expressive mean of modeling cooperative teams of decision makers.

MDP (Puterman 1994) $< S, A, T, R >$

- S set of states, A set of actions
- $T : S \times A \times S \rightarrow [0; 1]$ transition function giving the probability for the robot of transitioning from state s to s' after doing action a
- $R : S \rightarrow \Re$ reward function giving the robot's immediate reward for being in state s

Solving a MDP

- Find an optimal policy $\pi^* : S \rightarrow A$: sequence of actions maximising the long-term expected reward
- The value of π^* is defined by the optimal value function

$$V^*(s) = R(s) + \gamma \max_{a \in A} \sum_{s' \in S} T(s, a, s') V^*(s') \tag{1}$$

Decentralized Markov decision processes

Dec-(PO)MDP (Bernstein et al. 2002) $< n, S, A, T, R, \Omega, O >$

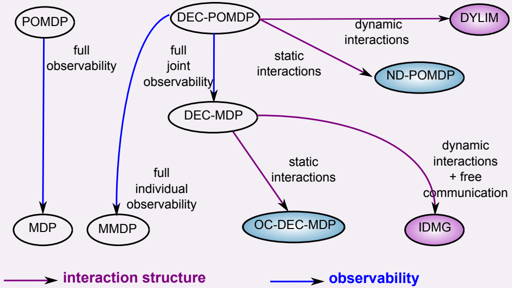
- n number of robots, S set of joint states, $A = \{A_1, ..., A_n\}$ set of joint actions
- $T : S \times A \times S \rightarrow [0; 1]$ transition function giving the probability for the n robots of transitioning from joint state s to s' after doing joint action a
- $R : S \rightarrow \Re$ reward function giving the robots' immediate reward for being in joint state s
- Ω set of observations et $O : S \times A \times S \times \Omega \rightarrow [0; 1]$ observation function can be left out if the states of agents are fully observable locally

Solving a Dec-(PO)MDP

- optimal resolution is NEXP complete (Bernstein et al. 2002)
- not realistic for large-scale problems/real-world applications
- assumption of total dependences: permanent interaction between all agents.

Resolution of Dec-(PO)MDPs

Interaction-oriented models



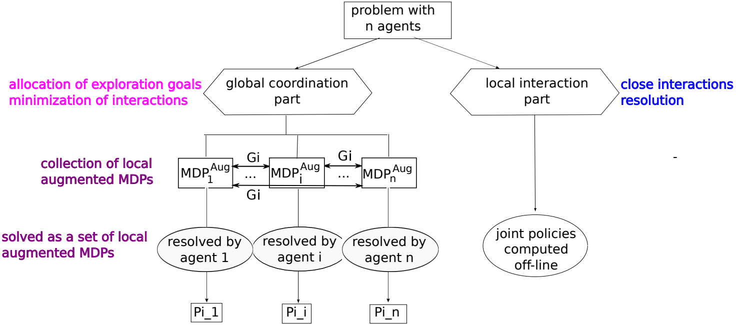
- exploit local interactions to solve the model
- relax the assumption of permanent interactions between all agents
  - static interactions: ND-POMDP (Nair et al. 2005)
  - dynamic interactions: DEC-SIMDP (Melo & Veloso 2011), DyLIM (Canu & Mouaddib 2011)

Our interaction-oriented approach

Motivations

- real-world application of Dec-MDPs to multi-robot exploration
- global coordination to allocate exploration goals
- minimize local interactions (overlapping, conflicts, collisions)

Our interaction-oriented approach (Matignon et al. 2012)



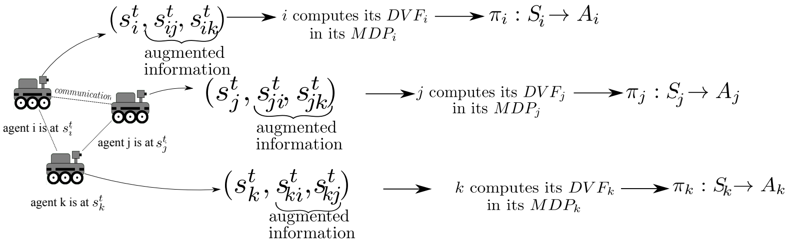
Global coordination class

- $MDP_i^{Aug} =< S_i, A_i, T_i, R_i, G_i >$
  - $< S_i, A_i, T_i, R_i >$ individually models agent i in the absence of other agents
  - G_i the augmented information enables interactions between MDPs

Reduction of the complexity to solve one MDP (polynomial) per agent

Augmented MDP resolution

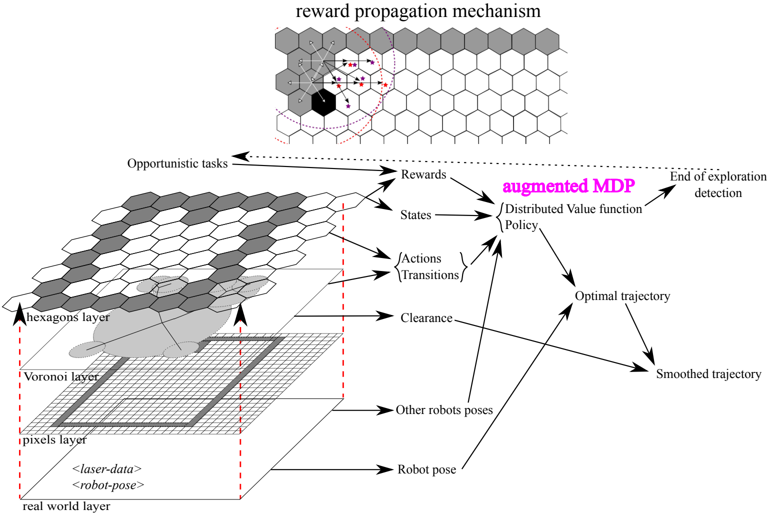
- G_i is limited to the states of other agents (exploration context)



$$\forall s_i \in S_i \quad DVF_i(s_i) = R_i(s_i) + \gamma \max_{a_i \in A_i}\Big(\sum_{s' \in S_i} T_i(s_i, a_i, s')[DVF_i(s') - \sum_{j \neq i} f_{ij} P_r(s'|s_{ij}) V_j(s')]\Big)$$

- $P_r(s'|s_{ij})$ probability for agent j of transitioning from s_{ij} to s'
- V_j value function of agent j
- f_{ij} weighting factor

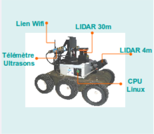
Augmented MDP model

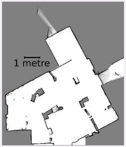


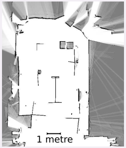
- One decision step: build the model, compute a policy from DVF, produce a smooth trajectory
- greedy approach

Experimental platforms

Robotic platform







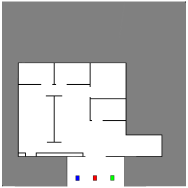
- Wifibot μ-troopers
- software based on a Data Distribution Service (publish-subscribe messaging pattern)
- various modules on each robot: laser acquisition, multi-robot SLAM (Xie et al. 2010), decentralized decision, mobility, object recognition (shape/template matching)
- the architecture: laser scans/states exchange, robust to communication failures

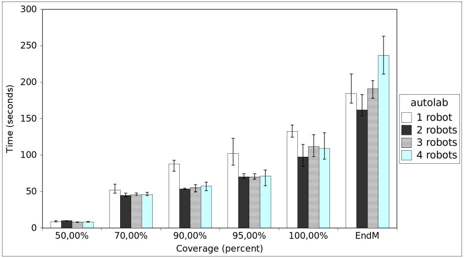
Simulations with Stage

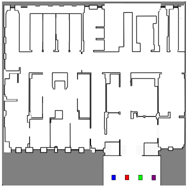
The simulator architecture mimics the robotic platform architecture:

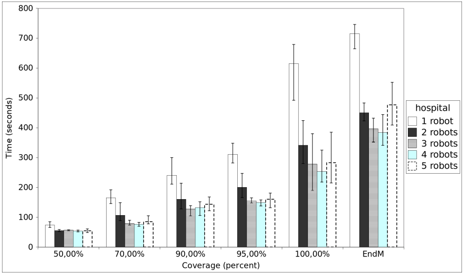
- DDS $\Leftrightarrow$ IPC shared memory segment
- laser acquisition module $\Leftrightarrow$ laser virtual sensor
- multi-robot SLAM module and mobility module $\Leftrightarrow$ position virtual device
- same decentralized decision module

Experimental results



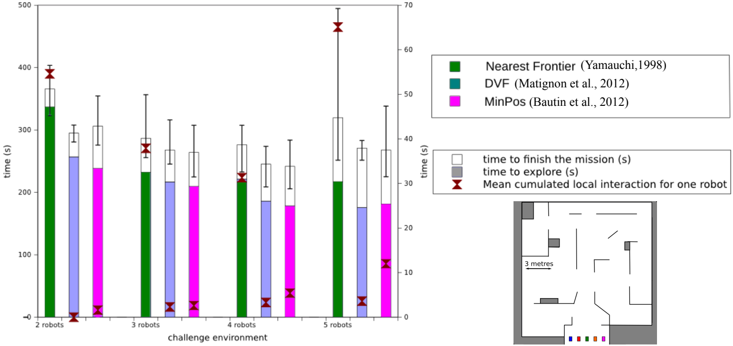






Experimental results

- Nearest frontier (Yamauchi,1998): greedy approach, no coordination
- MinPos (Bautin et al.,2012): the cost function is for each robot-frontier pair, the number of robots closer than it towards the considered frontier



Object recognition

First approach

- object detection and decision independant
- pictures to detect objects gathered along the way
- poor performance because some objects not photographed

Object recognition

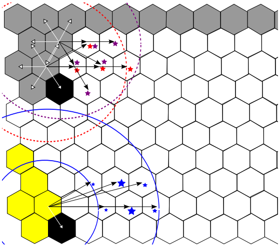
First approach

- object detection and decision independant
- pictures to detect objects gathered along the way
- poor performance because some objects not photographed

Second approach

- decision module extended to cover the explored space with photos
- two criteria: exploration and picture coverage

Active perception combined with exploration

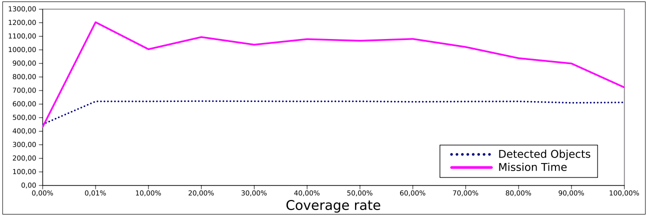


Reward propagation mechanisms

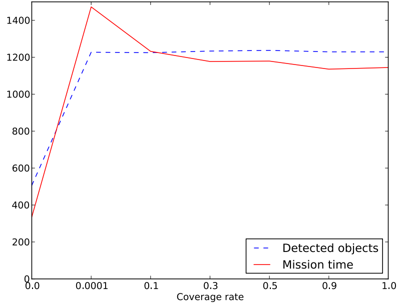
- $R_{i,exp}$ is the reward function to explore
- $R_{i,cov}$ is the reward function to take a photo
- $\alpha \in [0,1]$ is the picture coverage rate

$$\forall s_i \in S_i \quad DVF_i(s_i) = (1-\alpha)R_{i,exp}(s_i) + \alpha R_{i,cov}(s_i) +$$
$$\gamma \max_{a_i \in A_i} \Big(\sum_{s' \in S_i} T(s_i, a_i, s')[DVF_i(s') - \sum_{j \neq i} f_{ij} P_r(s'|s_{ij}) V_j(s')] \Big)$$

Experimental results in simulation

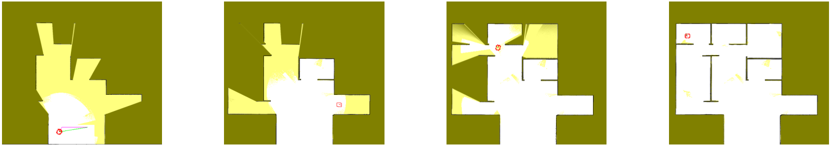


Results with autolab environment (over 340 simulations)

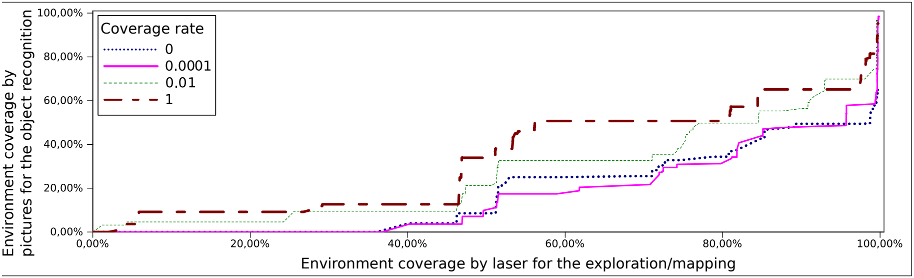


Results with challenge environment (over 700 simulations)

Experimental results in simulation



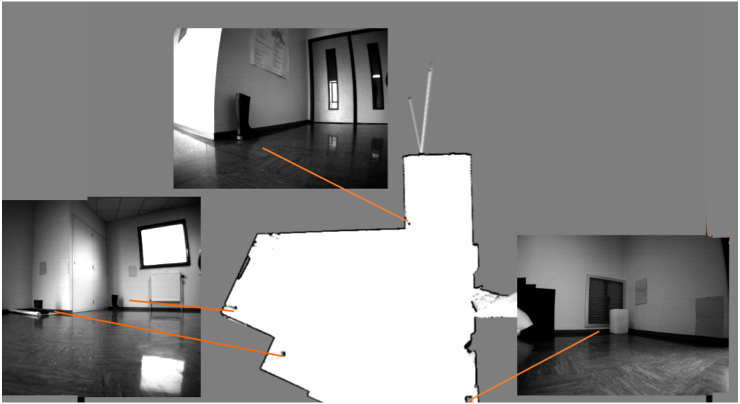
$\alpha = 1.0$: optimize photo coverage



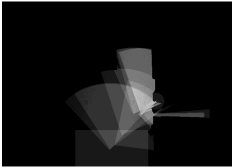
- α must be chosen to balance the priority of exploration and picture coverage

Experimental results with real robot

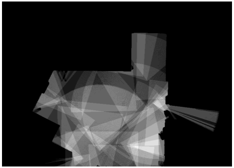
One robot, 4 objects:



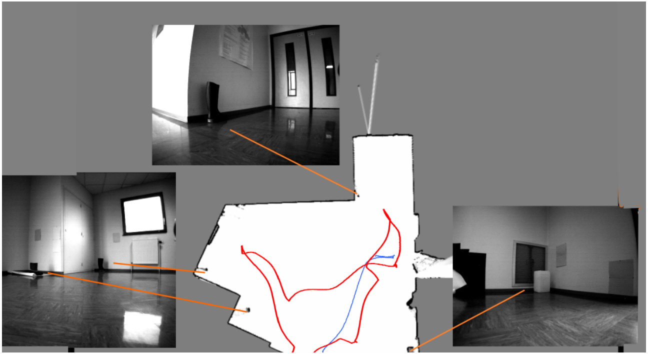
Experimental results with real robot



$\alpha = 0.0$
1 object taken in pictures



$\alpha = 1.0$
4 objects taken in pictures



Conclusion & Perspectives

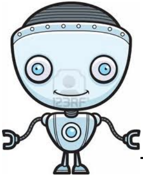
Conclusion

- Coordinated multi-robot exploration using Dec-MDPs
- Integration of the picture coverage criteria into the planning
- Balance the priority between exploration and picture coverage

Perspectives

- Compare object recognition results on real robots
- Active perception combined with exploration by planning viewpoints where the recognition process would be more reliable

Conclusion



Thank you !

laetitia.matignon@univ-lyon1.fr

http://liris.cnrs.fr/ lmatigno/

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.

Canu, A. & Mouaddib, A.-I. (2011).
Collective decision- theoretic planning for planet exploration, *Proc. of ICTAI*.

Matignon, L., Jeanpierre, L. & Mouaddib, A.-I. (2012).
Coordinated multi-robot exploration under communication constraints using decentralized markov decision processes., *Proc. of AAAI*.

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.

Puterman, M. L. (1994).
Markov decision processes, John Wiley and Sons.

Xie, J., Nashashibi, F., Parent, N. M. & Garcia-Favrot, O. (2010).
A real-time robust slam for large-scale outdoor environments, *17th ITS World Congress*.