Distributed Value Functions for the Coordination of Decentralized Decision Makers

(Extended Abstract)

Laëtitia Matignon, Laurent Jeanpierre and Abdel-Illah Mouaddib
Université de Caen Basse-Normandie, CNRS - UMR6072 GREYC - F-14032 Caen, France
{laetitia.matignon,laurent.jeanpierre,abdel-illah.mouaddib}@unicaen.fr

ABSTRACT

In this paper, we propose an approach based on an interaction-oriented resolution of decentralized Markov decision processes (Dec-MDPs) primary motivated by a real-world application of decentralized decision makers to explore and map an unknown environment. This interaction-oriented resolution is based on distributed value functions (DVF) techniques that decouple the multi-agent problem into a set of individual agent problems and consider possible interactions among agents as a separate layer. This leads to a significant reduction of the computational complexity by solving Dec-MDPs as a collection of MDPs. Using this model in multi-robot exploration scenarios, we show that each robot computes locally a strategy that minimizes the interactions between the robots and maximizes the space coverage of the team. Our technique has been implemented and evaluated in simulation and in real-world scenarios during a robotic challenge for the exploration and mapping of an unknown environment by mobile robots. Experimental results from real-world scenarios and from the challenge are given where our system was vice-champion.

Categories and Subject Descriptors

I.2.9 [**Artificial Intelligence**]: Robotics—*Autonomous vehicles*; I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent systems*

General Terms

Algorithms, Theory, Experimentation

Keywords

Cooperative multi-robot systems, robot coordination, robot planning, multi-robot exploration, distributed problem solving

1. INTRODUCTION

The approach developed in this paper is primary motivated by a real-world application of decentralized decision makers for an exploration and mapping multi-robot system. Our system has been developed and applied successfully in real-world scenarios during a DGA[1]/ANR[2] robotic challenge[3]. We focus only on the decision model allowing robots to cooperatively explore an unknown area and efficiently cover the space by reducing the overlap between explored areas of each robot. Such multi-robot exploration strategies have been proposed. They adopt either central agents or negotiation protocols with complicated processes [9, 1]. Besides existing works do not address the local coordination problem. So we took an interest in decentralized partially observable Markov decision processes (Dec-POMDPs) and their recent interaction-oriented (IO) resolution [8, 5, 2]. It takes advantage of local interactions and coordination by relaxing the most restrictive and complex assumption consisting in considering that agents are permanently in interaction. It is based on a set of interactive individual decision making problems and reduces the complexity of solving Dec-POMDPs thereby becoming a promising direction concerning real-world applications of decentralized decision makers. Consequently we propose in this paper an IO approach using distributed value functions so as to compute multi-robot exploration strategies in a decentralized way.

2. DISTRIBUTED VALUE FUNCTIONS

We propose an IO resolution of decentralized decision models using distributed value functions (DVFs) introduced in [6]. Our robots are independent and can share information by communication leading to some kind of observability completion. We assume full local observability for each robot and limited share of information. So our approach takes place in the Dec-MDP framework. DVF describes the Dec-MDP with a set of individual agent problems (MDPs) and considers possible interactions among robots as a separate interaction class where some information between robots are shared. This leads to a significant reduction of the computational complexity by solving Dec-MDP as a collection of MDPs. This could be represented as an IO resolution with two classes (no interaction class and interaction class) where each robot computes locally a strategy that minimizes conflicts, *i.e.* that avoids being in the interaction class. The interaction class is a separate layer solved independently by computing joint policies for these specific joint states.

DVF technique allows each robot to choose a goal which should not be considered by the others. The value of a goal depends on the expected rewards at this goal and on the fact that it is unlikely selected by other robot. Our DVF

Appears in: *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, Conitzer, Winikoff, Padgham, and van der Hoek (eds.), 4-8 June 2012, Valencia, Spain.

[1]French Defense Procurement Agency.
[2]French National Research Agency.
[3]http://www.defi-carotte.fr/

is defined solely by each robot in an MDP $< S, A, T, R >$. In case of permanent communication, robot i knows at each step t the state $s_j \in S$ of each other robot j and computes its DVF V_i according to :

$$\forall s \in S \quad V_i(s) = \max_{a \in A} \left(R(s,a) + \gamma \sum_{s' \in S} T(s,a,s') \right. \\ \left. [V_i(s') - \sum_{j \neq i} f_{ij} P_r(s'|s_j) V_j(s')] \right) \tag{1}$$

where $P_r(s'|s_j)$ is the probability for robot j of transitioning from its current state s_j to state s' and f_{ij} is a weighting factor that determines how strongly the value function of robot j reduces the one of robot i. Considering our communication restrictions, the robots cannot exchange information about their value functions. So we relax the assumptions concerning unlimited communication in DVF technique: each robot i can compute all V_j by empathy. Thus each robot computes strategies with DVF so as to minimize interactions. However when situations of interaction occur, DVF does not handle those situations and the local coordination must be resolved with another technique. For instance joint policies could be computed off-line for the specific joint states of close interactions.

3. EXPERIMENTS AND RESULTS

In these section is given an overview of our experiments. More details concerning our MDP model, our experimental platforms and our results can be found in [4]. First we made simulations with Stage[4] using different number of robots and various simulated environments. In fig. 1, we plot the time it takes to cover various percentages of the environment while varying the number of robots. During the beginning stage, robots spread out to different areas and covered the space efficiently. However, there is a number of robots beyond which there is not much gain in the coverage and that depends on the structure of the environment [7]. In the hospital environment, four robots can explore separate zones but the gain of having more robots is low compared to the overlap in trajectories and the risk of local interactions. Second we performed real experiments with our two robots besides the ones made during the challenge. Videos, available at `http://lmatigno.perso.info.unicaen.fr/research`, show different explorations of the robots and some interesting situations are underlined as global task repartition or local coordination.

4. CONCLUSIONS AND PERSPECTIVES

Our approach addresses the problem of multi-robot exploration with an IO resolution of Dec-MDPs based on DVFs. Experimental results from real-world scenarios and our vice-champion rank at the robotic challenge show that this method is able to effectively coordinate a team of robots during exploration. Though our DVF technique still assumes permanent communication similarly to most multi-robot exploration approaches where robots maintain constant communication while exploring to share the information they gathered and their locations. For instance classical negotiation based techniques assume permanent communication. However, permanent communication is seldom the case in

[4]http://playerstage.sourceforge.net/

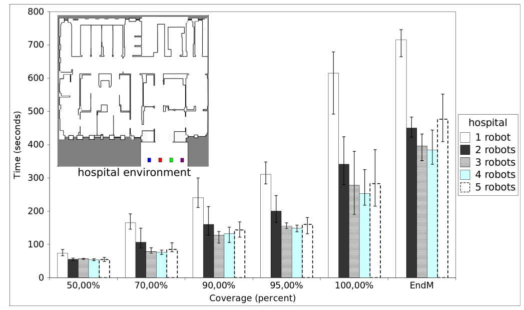

Figure 1: Hospital environment from Stage with starting positions and results averaged over 5 simulations.

practice and a significant difficulty is to account for potential communication drop-out and failures that can happen during the exploration leading to a loss of information that are shared between robots. Some recent multi-robot exploration approaches that consider communication constraints only cope with limited communication range issue and do not address the problem of failures as stochastic breaks in communication. So our short-term perspective is to extend our DVF to address stochastic communication breaks as it has been introduced in [3].

5. ACKNOWLEDGMENTS

This work has been supported by the ANR and the DGA (ANR-09-CORD-103) and jointly developed with the members of the Robots_Malins consortium.

6. REFERENCES

[1] W. Burgard, M. Moors, C. Stachniss, and F. Schneider. Coordinated multi-robot exploration. *IEEE Transactions on Robotics*, 21:376–386, 2005.

[2] A. Canu and A.-I. Mouaddib. Collective decision-theoretic planning for planet exploration. In *Proc. of ICTAI*, 2011.

[3] L. Matignon, L. Jeanpierre, and A.-I. Mouaddib. Distributed value functions for multi-robot exploration: a position paper. In *AAMAS Workshop: Multiagent Sequential Decision Making in Uncertain Domains (MDSM)*, 2011.

[4] L. Matignon, L. Jeanpierre, and A.-I. Mouaddib. Distributed value functions for multi-robot exploration. In *Proc. of ICRA*, 2012.

[5] F. Melo and M. Veloso. Decentralized mdps with sparse interactions. *Artif. Intell.*, 175(11):1757–1789, 2011.

[6] J. Schneider, W.-K. Wong, A. Moore, and M. Riedmiller. Distributed value functions. In *Proc. of ICML*, pages 371–378, 1999.

[7] R. Simmons, D. Apfelbaum, W. Burgard, D. an Moors M. Fox, S. Thrun, and H. Younes. Coordination for multi-robot exploration and mapping. In *Proc. of the AAAI National Conf. on Artificial Intelligence*, 2000.

[8] M. T. J. Spaan and F. S. Melo. Interaction-driven markov games for decentralized multiagent planning under uncertainty. In *AAMAS*, pages 525–532, 2008.

[9] R. Zlot, A. Stentz, M. Dias, and S. Thayer. Multi-robot exploration controlled by a market economy. In *Proc. of ICRA*, volume 3, pages 3016–3023, May 2002.