

Probabilistic Planning for Decentralized Multi-Robot Systems

Christopher Amato
Dept. of Computer Science
University of New Hampshire
Durham, NH 03824

George Konidaris
Depts. of CS and ECE
Duke University
Durham, NC 27708

Shayegan Omidshafiei
LIDS
Massachusetts Institute of Technology
Cambridge, MA 02139

Ali-akbar Agha-mohammadi
Qualcomm Research
San Diego, CA 92121

Jonathan P. How
LIDS
Massachusetts Institute of Technology
Cambridge, MA 02139

Leslie P. Kaelbling
CSAIL
Massachusetts Institute of Technology
Cambridge, MA 02139

Introduction

Multi-robot systems offer a compelling application domain for multi-agent AI research: they pose challenging questions, they offer an immediate bridge to actual and tangible systems, and they have vast potential for real-world applications. However, most multi-robot systems are controlled by hand-built special-purpose algorithms that are difficult to design, implement and verify. For single robots, automatic planning systems provide a flexible general-purpose strategy for constructing plans given high-level declarative domain specifications, even in the presence of substantial stochasticity and partial observability (Thrun, Burgard, and Fox 2005). Our recent research has shown that this same strategy—finding the appropriate generic model and writing general-purpose planners for it—can be applied to multi-robot systems. Our methods allow automatic off-line construction of robust multi-robot policies that support coordinated actions. As a natural consequence, our methods can generate control and communication strategies that optimize the group’s overall objective.

Specifically, we are interested in problems where robots share the same objective function and each individual robot can only make noisy, partial observations of the environment. The decentralized partially observable Markov decision process (Dec-POMDP) is a general framework for representing multi-agent coordination problems. Dec-POMDPs have been widely studied in artificial intelligence as a way to address the fundamental differences in decision-making in decentralized settings (Amato et al. 2013; Bernstein et al. 2002; Oliehoek 2012). Like the POMDP (Kaelbling, Littman, and Cassandra 1998) model that it extends, Dec-POMDPs consider general dynamics, cost and sensor models. Any problem where multiple robots share a single overall reward or cost function can be formalized as a Dec-POMDP. Unfortunately, this generality comes at a cost: Dec-POMDPs are typically infeasible to solve except for small problems (Bernstein et al. 2002; Amato, Konidaris, and Kaelbling. 2014).

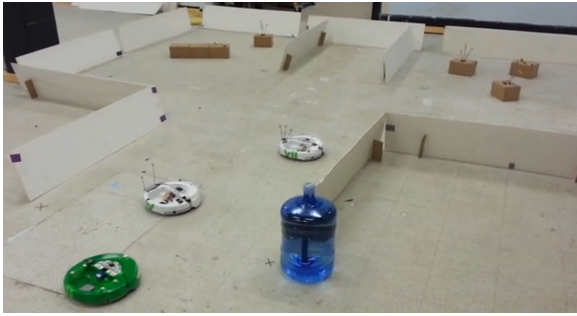
One reason for the intractability of solving large Dec-POMDPs is that current approaches model problems at a low level of granularity, where each robot’s actions are primitive operations lasting exactly one time step. Our recent research has addressed the more realistic *MacDec-POMDP* case where each robot has *macro-actions*: temporally extended actions which may require different amounts of time to execute (Amato, Konidaris, and Kaelbling. 2014). An alternative formulation is the Dec-POSMDP, which operates directly in belief space (Omidshafiei et al. 2015). These models allow coordination decisions to only occur at the level of deciding which macro-actions to execute. Macro-actions are a natural model for the modular controllers (e.g., navigating to a waypoint or grasping an object) sequenced to obtain robot behavior, bridging the gap between robotics research and Dec-POMDPs.

MacDec-POMDPs and Dec-POSMDPs

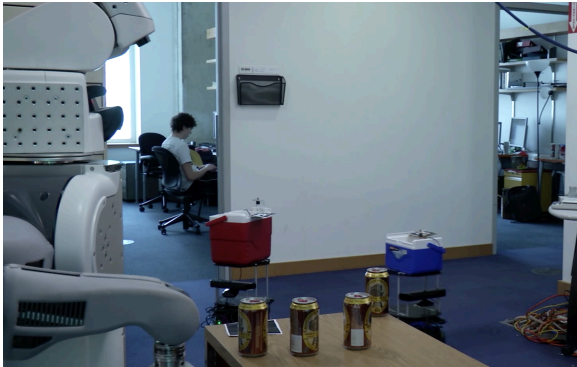
In Dec-POMDPs, multiple robots operate with partial and local views of the world. At each step, every robot chooses an action based on locally observable information, resulting in an observation for each individual robot. The robots share a single reward function, making the problem cooperative, but their local views mean that execution is decentralized.

MacDec-POMDPs incorporate macro-actions into the Dec-POMDP framework, where macro-actions have defined initial conditions where they can be executed and this execution continues until some terminal condition is reached. In the MacDec-POMDP framework, it is assumed that either a low-level (Dec-POMDP) model or a simulator is available in order to evaluate solutions. As a result, MacDec-POMDPs do not explicitly model the time until completion. In contrast, Dec-POSMDPs explicitly model the distribution of time until completion. Solutions in this semi-Markov model can then be evaluated using a higher-level model (including completion time) or in a simulator.

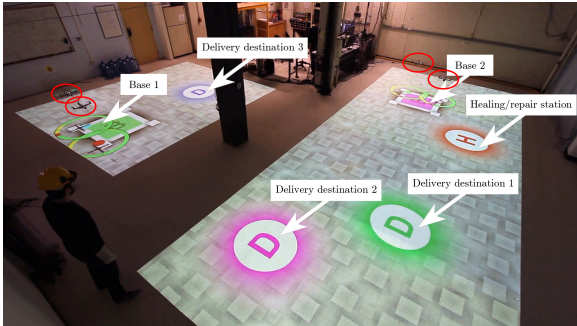
Two Dec-POMDP algorithms have been extended to the MacDec-POMDP case (Amato, Konidaris, and Kaelbling.



(a) The warehouse problem



(b) The bartender and waiters domain



(c) The package delivery problem

Figure 1: Experimental domains

2014), but other extensions are possible. The resulting solution is a policy for each agent, which can be represented as a set of trees. In the Dec-POSMDP, we represent the policy as a finite-state controller for each agent. Heuristic search (Amato et al. 2015a) and sample-based optimization methods (Omidshafiei et al. 2015) have been successfully used for generating the parameters for these finite-state controllers.

The MacDec-POMDP framework is a natural way to represent and generate behavior for general multi-robot systems. We assume an abstract model of the system in the form of macro-actions, which include the initiation and terminal conditions. These macro-actions are controllers operating with possibly continuous actions and feedback, but their operation is discretized for use with the planner. Given the macro-actions and simulator, the planner then automati-

cally generates a solution which optimizes the value function with respect to the uncertainty over outcomes, sensor information and other robots. This solution comes in the form of hierarchical state machines for use in a ROS-based robot environment.

Experiments

We performed comparisons with previous work on existing benchmark domains and demonstrated its effectiveness in different scenarios (Warehouse (Amato et al. 2015b), Bartender and waiters (Amato et al. 2015a), and Package delivery (Omidshafiei et al. 2015)). In the warehouse problem (Figure 1(a)), a team of robots is tasked with finding a set of large and small boxes in the environment and returning them to a shipping location. Here, coordination is needed not just for assigning robots to push specific boxes, but also requires that two robots push the larger box at the same time. In the bartender and waiters problem (Figure 1(b)), the waiters (Turtlebots) must find and deliver orders as quickly as possible, retrieving drinks from a bartender (PR2). In the package delivery problem (Figure 1(c)), the robots retrieve and deliver packages from base locations to delivery locations. In all problems there is stochasticity in the movements of robots and partial observability of the location of the other robots and the other objects (boxes, orders and packages).

These problems are very large (consisting of over a billion discrete states or having a continuous state space), and thus unsolvable by previous Dec-POMDP-based approaches. We also consider cases where the robots can send communication signals to each other, but we do not define the meaning of the messages. Therefore, our planner must determine where the robots should navigate, what boxes they should push and what communication messages should be sent (if at all) at each step of the problem to optimize the solution for the team. The robots must make these decisions based solely on the information they individually receive during execution (e.g., each robot’s location estimate as well as where and when boxes and other robots have been seen). Our methods outperform naive methods that do not consider uncertainty and generate optimized solutions for each problem based on the high-level domain description.

Summary

Multi-robot systems are an exciting application domain for AI research and Dec-POMDPs, specifically. MacDec-POMDP methods can produce high-quality general solutions for realistic heterogeneous multi-robot coordination problems by automatically generating control and communication policies, given a model. In contrast to most existing multi-robot methods that are specialized to a particular problem class, our approach can synthesize policies that exploit any opportunities for coordination that are present in the problem, while balancing uncertainty, sensor information, and information about other agents.

References

Amato, C.; Chowdhary, G.; Geramifard, A.; Ure, N. K.; and Kochenderfer, M. J. 2013. Decentralized control of partially

observable Markov decision processes. In *the IEEE Conference on Decision and Control*.

Amato, C.; Konidaris, G. D.; Anders, A.; Cruz, G.; How, J. P.; and Kaelbling, L. P. 2015a. Policy search for multi-robot coordination under uncertainty. In *the Robotics: Science and Systems Conference*.

Amato, C.; Konidaris, G. D.; Cruz, G.; Maynor, C. A.; How, J. P.; and Kaelbling, L. P. 2015b. Planning for decentralized control of multiple robots under uncertainty. In *the International Conference on Robotics and Automation*.

Amato, C.; Konidaris, G. D.; and Kaelbling, L. P. 2014. Planning with macro-actions in decentralized POMDPs. In *the International Conference on Autonomous Agents and Multiagent Systems*.

Bernstein, D. S.; Givan, R.; Immerman, N.; and Zilberstein, S. 2002. The complexity of decentralized control of Markov decision processes. *Mathematics of Operations Research* 27(4):819–840.

Kaelbling, L. P.; Littman, M. L.; and Cassandra, A. R. 1998. Planning and acting in partially observable stochastic domains. *Artificial Intelligence* 101:1–45.

Oliehoek, F. A. 2012. Decentralized POMDPs. In Wiering, M., and van Otterlo, M., eds., *Reinforcement Learning: State of the Art*, volume 12 of *Adaptation, Learning, and Optimization*. Springer Berlin Heidelberg. 471–503.

Omidshafiei, S.; Agha-mohammadi, A.; Amato, C.; and How, J. P. 2015. Decentralized control of partially observable Markov decision processes using belief space macro-actions. In *the International Conference on Robotics and Automation*.

Thrun, S.; Burgard, W.; and Fox, D. 2005. *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press.