Approximate Solutions for Factored Dec-POMDPs with Many Agents

Oliehoek, F.A.; Whiteson, S.; Spaan, M.T.J.

Approximate Solutions for Factored Dec-POMDPs with Many Agents[1]

Frans A. Oliehoek [a]   Shimon Whiteson [b]   Matthijs T.J. Spaan [c]

[a] *Maastricht University. P.O.Box 616, 6200 MD Maastricht*
[b] *University of Amsterdam. Science Park 904, 1098 XH Amsterdam*
[c] *Delft University of Technology. Mekelweg 4, 2628 CD Delft*

Abstract

Dec-POMDPs are a powerful framework for planning in multiagent systems, but are provably intractable to solve. This paper proposes a factored forward-sweep policy computation method that tackles the stages of the problem one by one, exploiting weakly coupled structure at each of these stages. An empirical evaluation shows that the loss in solution quality due to these approximations is small and that the proposed method achieves unprecedented scalability, solving Dec-POMDPs with hundreds of agents.

1 Introduction

Cooperative multiagent systems are of significant scientific interest, not only because they can tackle distributed problems, but also because they facilitate the decomposition of problems too complex to be tackled by a single agent. As a result, a fundamental question in artificial intelligence is: how should teams of agents act so as to most effectively achieve common goals? When uncertainty and many agents are involved, this question is challenging, and has not yet been answered in a satisfactory way. Even single-agent decison making is complicated by incomplete knowledge of the environment's state (e.g., due to noisy sensors). However, multiagent settings exacerbate this problem, as agents have access only to their own sensors.

This paper focuses on the finite-horizon *decentralized partially observable Markov decision process* (Dec-POMDP) [1], a model that can represent such problems under uncertainty. We consider factored Dec-POMDPs, which—by explicitly representing the structure present in the problem—open the door to methods that exploit this structure and thus scale to many more agents. While there have been positive results for specific subclasses that restrict the type of structure, e.g., [7], so far only moderate scalability has been achieved for general Dec-POMDPs. Since even finding an ϵ-approximate solution is NEXP-complete, any method that is guaranteed to be computationally efficient cannot guarantee an absolute bound on the error. Therefore, in this paper, we abandon optimality guarantees and aim for scalability in the number of agents.

2 Approach

This paper proposes an approximate planning method for factored Dec-POMDPs. These are problems such as FireFightingGraph, illustrated in Fig. 1(a), where the state consists of multiple variables and each agent can only affect a subset of these variables directly. Additionally each agent only gets local observations that depend on a local subset of variables. The goal is to maximize the sum of local rewards over time.

---

[1]Full version appeared in the *Proc. of the Twelfth International Conference on Autonomous Agents and Multiagent Systems* [6].

(a) The FireFightingGraph benchmark. (b) Decompostion into smaller source problems. (c) Scalability w.r.t. agents.
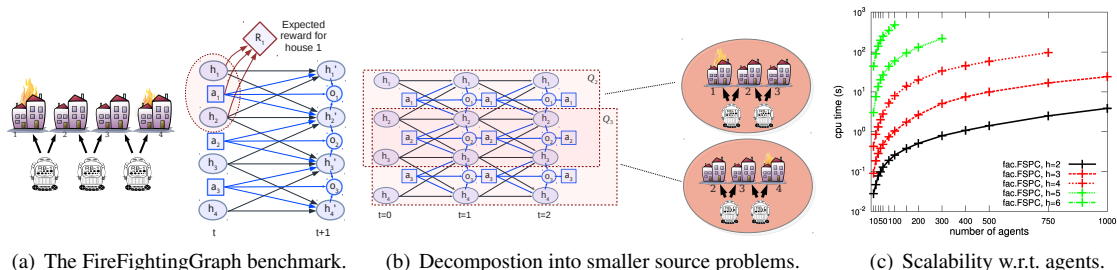
We propose an approach based on *forward-sweep policy computation* (FSPC), which approximates the Dec-POMDP by a series of one-shot *Bayesian games* (BGs), one for each stage [3]. Our method, called *factored FSPC* (FFSPC), exploits weakly coupled structure at each stage by replacing the BGs with *collaborative graphical BGs* (CGBGs) [4]. The main algorithmic contribution of the paper is a set of approximations necessary to make FFSPC feasible for problems with many agents:

1. First, we approximate the interaction structure between agents by constructing CGBGs with a predefined factorization. E.g., as illustrated by the rectangles in the left part of Fig. 1(b).
2. Second, instead of following the common practice of solving the underlying (PO)MDP, we employ a new class of heuristics based on *transfer planning* that directly approximate the factored Dec-POMDP value function. This works by defining source problems for each component and transferring the value functions to the original target problem (also shown in Fig. 1(b)).
3. Third, we use approximate inference [2] to efficiently construct the CGBGs.
4. Finally, we approximately solve the CGBGs by applying *Max-Sum* to their *agent and type independence factor graphs*, an approach that has shown state-of-the-art results [5].

An extensive empirical evaluation that shows that FFSPC is highly scalable with respect to the number of agents (see Fig. 1(c)), while attaining (near-) optimal values. In particular, our method shows scalability to hundreds of agents which is unprecedented for the class of factored Dec-POMDPs that we consider; previous approaches applicable to these problems have not been demonstrated beyond 20 agents. Moreover, a detailed analysis of our approximation indicates that there is no significant decrease in value due to sparse factorization and approximate CGBG construction, and that the transfer planning heuristics significantly outperform two baselines. While this works successfully scales to large number of agents, future work should concentrate on combining this scalability with the ability to tackle longer planning horizons [3].

References

[1] Daniel S. Bernstein, Robert Givan, Neil Immerman, and Shlomo Zilberstein. The complexity of decentralized control of Markov decision processes. *Math. of OR*, 27(4):819–840, 2002.

[2] Kevin P. Murphy and Yair Weiss. The factored frontier algorithm for approximate inference in DBNs. In *UAI*, 2001.

[3] Frans A. Oliehoek, Matthijs T. J. Spaan, Christopher Amato, and Shimon Whiteson. Incremental clustering and expansion for faster optimal planning in decentralized POMDPs. *JAIR*, 46:449–509, 2013.

[4] Frans A. Oliehoek, Matthijs T. J. Spaan, Shimon Whiteson, and Nikos Vlassis. Exploiting locality of interaction in factored Dec-POMDPs. In *AAMAS*, 2008.

[5] Frans A. Oliehoek, Shimon Whiteson, and Matthijs T. J. Spaan. Exploiting structure in cooperative Bayesian games. In *UAI*, 2012.

[6] Frans A. Oliehoek, Shimon Whiteson, and Matthijs T. J. Spaan. Approximate solutions for factored Dec-POMDPs with many agents. In *AAMAS*, pages 563–570, 2013.

[7] Stefan J. Witwicki and Edmund H. Durfee. Influence-based policy abstraction for weakly-coupled Dec-POMDPs. In *ICAPS*, 2010.