Scaling Up Optimal Heuristic Search in Dec-POMDPs via Incremental Expansion

Matthijs T.J. Spaan [a] Frans A. Oliehoek [b] Christopher Amato [c]

[a] *Delft University of Technology, Delft, The Netherlands*
[b] *Massachusetts Inst. of Technology, Cambridge, MA, USA*
[c] *Aptima, Inc., Woburn, MA, USA*

Abstract

We advance the state of the art in optimal solving of decentralized partially observable Markov decision processes (Dec-POMDPs), which provide a formal model for multiagent planning under uncertainty.

The full version of this paper has appeared in *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI-11)*.

1 Introduction

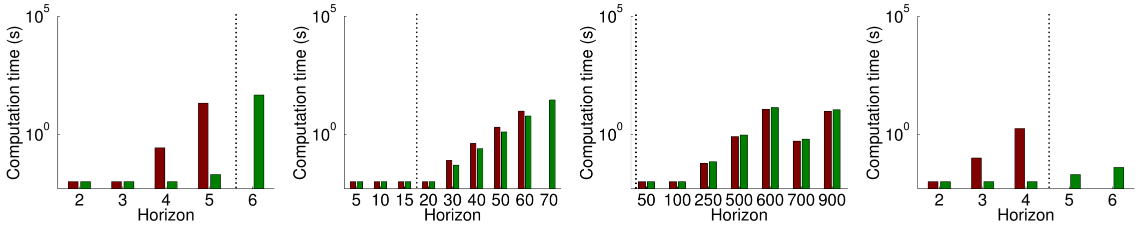
Planning under uncertainty for multiagent systems is an important problem in artificial intelligence, as agents may often possess uncertain information while sharing their environment with other agents. Due to stochastic actions and noisy sensors, agents must reason about many possible outcomes and the uncertainty surrounding them. In cooperative systems, finding optimal joint plans is especially challenging when each agent must choose actions based solely on local knowledge due to nonexistent or noisy communication. Possible application domains include multi-robot teams, communication networks, load balancing, and other problems in which agents need to coordinate under uncertain conditions.

The decentralized partially observable Markov decision process (Dec-POMDP) is a formal model for such planning problems [1]. In this paper we consider the optimal solution of Dec-POMDPs over a finite horizon. Unfortunately, optimal solution methods and even bounded approximations suffer from doubly-exponential complexity (NEXP-Complete); the search space for horizon $h + 1$ is exponentially larger than the one for horizon h. Even though the high worst-case complexity results preclude optimal methods from being applicable for some larger problems, as approximate algorithms come with no guarantees, optimal methods are necessary as a tool to analyze their performance.

2 Scaling up heuristic search in Dec-POMDPs

We provide significant advances to the state of the art in optimal Dec-POMDP solution methods by extending Multiagent A* (MAA*) [4] —which performs an A* search through the tree of possible partial joint policies— and derived methods with a new technique for incremental expansion of search tree nodes. Expanding a node in this context entails generating all possible children, which is a major source of intractability since the number of such children is doubly exponential in the depth of the node. In practice, however, only a small number of the generated nodes may actually be queried during the search. Our key insight is that if a method is able to *incrementally* generate children in order of their heuristic value, not all nodes need to be expanded at once. We integrate incremental expansion in Generalized MAA* with incremental clustering (GMAA*-IC), an MAA* extension that uses lossless history clustering for improved scalability [2]. We prove that the resulting algorithm, dubbed GMAA*-ICE, is correct and expands search nodes in the same order [3].

As with any A* method, our approach's performance depends on the tightness of the heuristic. In many problems the upper bound provided by the value function of the underlying MDP is not tight enough for



(a) Dec-Tiger (5) (b) Recycling Robots (15) (c) Broadcast Channel (25) (d) GridSmall (4)

Figure 1: Experimental results comparing the computation times of GMAA*-IC with full expansion (red bars, left) and incremental expansion (green bars, right), using the hybrid Q_{BG} representation. Between parentheses is the longest horizon solved before (indicated as well by the vertical dashed line).

heuristic search to be effective [1]. Other heuristics are tighter, such as those based on the underlying POMDP solution or the value function resulting from assuming 1-step-delayed communication. However, they require storing values for all joint action-observation histories or representing them as a potentially exponential number of vectors. A crucial insight is that the size of a tree-based representation grows exponentially when moving forward in time, while the vector-based representation grows in the opposite direction. We exploit this insight by introducing a hybrid representation that is more compact.

Figure 1 shows a subset of the experimental results on benchmark domains, showing that the effect of incremental expansion is complementary to clustering of histories. With just the new heuristic representation, optimal plans could be found for larger horizons than any previous work for four benchmarks. In one case, horizons that are over an order of magnitude larger could be reached. By exploiting incremental expansion, we achieve further improvements in scalability. The combination of the hybrid representation and incremental expansion provides a powerful method for optimally solving Dec-POMDP over longer horizons.

3 Conclusions

Decentralized POMDPs offer a rich model for multiagent coordination under uncertainty. Optimal solution methods for Dec-POMDPs are of great interest; they are of practical value for smaller or decomposable problems and lie at the basis for most successful approximate methods. We advanced the state of the art by introducing an effective method for *incremental expansion* of search nodes. We proved that the resulting algorithm, GMAA*-ICE, is search-equivalent to GMAA*-IC and therefore complete. A new bottleneck, the amount of space needed for representation of the heuristic, was addressed by introducing a representation that is a hybrid between tree-based and vector-based representations. We showed the efficacy of our methods on a suite of benchmark problems, demonstrating a significant speedup over the state of the art.

Acknowledgments

M.S. is funded by the FP7 Marie Curie Actions Individual Fellowship #275217 (FP7-PEOPLE-2010-IEF). Research supported in part by AFOSR MURI project #FA9550-09-1-0538.

References

[1] Frans A. Oliehoek, Matthijs T. J. Spaan, and Nikos Vlassis. Optimal and approximate Q-value functions for decentralized POMDPs. *Journal of Artificial Intelligence Research*, 32:289–353, 2008.

[2] Frans A. Oliehoek, Shimon Whiteson, and Matthijs T. J. Spaan. Lossless clustering of histories in decentralized POMDPs. In *Proc. of Int. Conf. on Autonomous Agents and Multi Agent Systems*, 2009.

[3] Matthijs T. J. Spaan, Frans A. Oliehoek, and Christopher Amato. Scaling up optimal heuristic search in Dec-POMDPs via incremental expansion. In *Proc. of Int. Joint Conf. on Artificial Intelligence*, 2011.

[4] D. Szer, F. Charpillet, and S. Zilberstein. MAA*: A heuristic search algorithm for solving decentralized POMDPs. In *Proc. of Uncertainty in Artificial Intelligence*, 2005.