Commitment Abstraction for Efficient Planning and Coordination in Stochastic Multi-agent Systems

(Extended Abstract of Doctoral Thesis)

Stefan J. Witwicki
Computer Science and Engineering
University of Michigan
Ann Arbor, MI 48109
witwicki@umich.edu

Categories and Subject Descriptors

I.2.11 [**Artificial Intelligence**]: Distributed Artificial Intelligence—*Multiagent Systems*

General Terms

Algorithms, Design, Performance, Theory

Keywords

Multiagent Systems, Planning Under Uncertainty, Coordination, Commitments, Decentralized Markov Decision Processes, Agent Models, Policy Abstraction

1. INTRODUCTION

This thesis is motivated by the vast complexity of cooperative stochastic multi-agent planning, where agents can affect the transitions and rewards of one another, but in order to coordinate their interactions effectively, must account for the uncertainty in these actions. To combat this complexity, I exploit interaction structure in *weakly-coupled* problems to compute coordinated agent policies. I contend that when the degree of inter-agent dependence is sufficiently limited, the multi-agent problem can be solved more efficiently if it is broken up into (partially) decoupled subproblems: formulation of individual agent policies and coordination of abstract interactions. In support of this thesis, I develop an approach by which individual agents plan with local behavioral models that incorporate only those portions of negotiated nonlocal behavior that are needed for effective coordination.

2. PROBLEM DESCRIPTION

Figure 1 illustrates a multi-agent planning problem represented in the TAEMS language (as described in [1]). The objective is to plan policies for two autonomous vehicle agents that coordinate their execution of hierarchical tasks with uncertain durations so as to maximize expected quality within mission deadlines. We can model this example as a Decentralized Markov Decision Process (DEC-MDP) as discussed by Becker, Zilberstein, and Lesser [1]. With the characteristics that follow, I outline a class of weakly-coupled DEC-MDPs that is the focus of thesis.

Temporal Grounding. Agents perform activities with well-defined (but often probabilistically uncertain) durations. The goals of the system are temporally constrained with strict deadlines. This is all modeled by a *finite-horizon* DEC-MDP for which *time* is a necessary state feature.

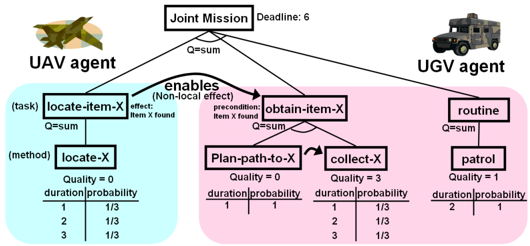

Figure 1: Autonomous Vehicle Example Problem

Decentralized Awareness. The agents do not have complete views of the world. Instead, each is only aware of a subset of information related to its individual activities. Technically, this corresponds to a *factored*, *locally fully observable* DEC-MDP where each agent's *local state* is composed of features related to the execution of its own tasks.

Structured Interactions. Agents have a limited influence on the outcomes of each others' activities. In particular, one agent may affect the local state transitions of another (sequentially, but not concurrently) through the *event-driven* manipulation of shared state features. In Figure 1, structured interaction occurs when the UAV *locates item X*, thereby *enabling* the UGV to successfully *obtain X*. I further assume an agent's non-local dependencies to be substantially less abundant than those within its local transitions.

Limited Planning Time. In many domains, it is important to plan coordinated behavior as soon as possible so as not to delay mission execution. Here, quickly-planned effective agent policies may be preferable to optimal policies that take longer to compute (or for large problems, are simply intractable). Thus, a solution to this class of problems is a method of policy computation that can (1) produce effective, coordinated (approximately-optimal) policies for problems both large and small, and (2) depending on problem difficulty, allow for trade-offs to be made between computation time and expected quality of computed behavior.

Previous work has solved related problems in restricted contexts [1, 4, 5], but no planning method (to date) constitutes a full solution to the class that I have outlined.

3. SOLUTION APPROACH

I propose an approach for coordinating interdependent agent activities through behavioral promises: *commitments*. A commitment encodes an agent's intention (and ability) to interact with other agents (by altering their local transitions). Because there is uncertainty in the system dynamics, a commitment also encodes time and probability information corresponding to when and with what likelihood other agents can expect the interaction to occur. Through negotiation of commitment values, the agents can plan their interactions and coordinate their individual behaviors around these planned interactions. The text that follows describes the components of my approach, cites results to date, and discusses planned research steps that I will take in completing my dissertation.

Commitment Modeling

Weakly-coupled problems involve highly-independent agents that interact with one another only in a limited capacity. Instead of considering all nonlocal dynamics, why not abstract only that which is relevant for planning an agent's limited interactions? Commitment models provide effective, compact approximations of external behavior. Although commitments have been studied in various classical planning domains (by Durfee and Lesser [2], for example), my problems call for the application of commitment theory [3] to Markov Decision Processes. I conjecture that planning with compact commitment-augmented local MDPs will allow weakly-coupled agents to coordinate complex, uncertain behavior efficiently. To test this theory, I have developed a method of augmenting MDPs with commitment models for *enablements* (as are present in Figure 1) [6], and plan to extend my models to represent other structured interactions.

Commitment Enforcement

Agents can compute policies by applying standard MDP solution techniques to their commitment-augmented local models. But because they are modeling behavioral expectations, these local policies need to satisfy the commitments that agents have promised. I have developed a method of commitment enforcement that, unlike previous work that injects artificial rewards and penalties to bias agents' actions, constrains policies directly to probabilistically adhere to committed interactions [6]. My linear programming approach automatically determines whether or not a given commitment selection is feasible and, if it is, computes optimal local behavior with respect to the commitment selection.

Commitment Negotiation

My commitment infrastructure transforms the problem of computing coordinated behavior into a search over the space of possible commitments. In fact, I have proven that, for an interesting subset of those problems, there exist commitments that (when enforced) yield globally-optimal joint policies [8]. For difficult problems, searching the commitment space exhaustively will be intractable. But I have developed an effective approximate algorithm that iteratively selects a set of commitments, builds local policies that enforce those commitments, estimates global quality, and repeats until greedily converging [7]. I have also demonstrated the scalability of my approach, and made analytical arguments [8] about the advantages that it has over existing methods, but I plan to verify these arguments with further empirical comparisons to demonstrate its robustness compared to other approaches (e.g. [1]).

Flexible Temporal Abstraction

In addition to approximating behavior, commitments provide a natural abstraction of the timing of uncertain interactions. A complete commitment model of the UAV agent (from Figure 1) would model every possible time (1,2, and 3) that *locate X* could occur. Consider instead representing this interaction with only a single time (time 3, for example) and the probability of *locate X* finishing *by* that time. As I have shown [7], these temporally-abstract commitment models maintain compactness as we scale to problems with increased complexity. Commitments of this sort are capable of encoding a single interaction time, all possible times, or any number of times in between [8], allowing for a flexibility of approximation that I am in the process of evaluating.

Policies Over Commitments

The last component of my approach is motivated by the fact that there may be dependencies between interactions that my present commitment models do not consider. For example, if there is a chain of enablement interactions (whereby Agent 1 enables Agent 2, allowing Agent 2 to enable Agent 3, etc.), I may be able to take advantage of this dependency structure by explicitly accounting for changes in expected behavior. If Agent 1 fails to enable Agent 2, Agent 3 should change its expectation of getting enabled by Agent 2. I envision an extension to my approach that allows such changes to be incorporated into a meta-level policy over commitments.

4. CONTRIBUTIONS

I expect that my completed dissertation will contribute:

- a principled framework for nonlocal abstraction in MDPs,
- an arsenal of LP-based policy formulation techniques for constraining agent behavior,
- a scalable, efficient, flexibly-approximate solution methodology for a relevant class of DEC-MDP problem, and
- a novel system of dynamic commitments.

5. REFERENCES

[1] R. Becker, S. Zilberstein, and V. Lesser. Decentralized Markov decision processes with event-driven interactions. *AAMAS*, pages 302–309, 2004.

[2] E. Durfee and V. Lesser. Partial global planning: A coordination framework for distributed hypothesis formation. *IEEE Trans. SMC*, 21(5):1167–1183, 1991.

[3] N. Jennings. Commitments and conventions: The foundation of coordination in multi-agent systems. *The Knowledge Engineering Review*, 8(3):223–250, 1993.

[4] J. Marecki and M. Tambe. On opportunistic techniques for solving decentralized Markov decision processes with temporal constraints. *AAMAS*, pages 830–837, 2007.

[5] R. Nair, P. Varakantham, M. Tambe, and M. Yokoo. Networked distributed POMDPs: A synthesis of DCOP and POMDPs. *AAAI*, pages 133–139, 2005.

[6] S. Witwicki and E. Durfee. Commitment-driven distributed joint policy search. *AAMAS*, 480–487, 2007.

[7] S. Witwicki and E. Durfee. Commitment-based service coordination. *Int. Journal of AOSE*, 3(1):59–87, 2009. [http://www.umich.edu/~witwicki/pub/IJAOSE.pdf].

[8] S. Witwicki and E. Durfee. Flexible approximation of structured interactions in decentralized Markov decision processes (extended abstract). *AAMAS*, 2009.