Flexible Approximation of Structured Interactions in Decentralized Markov Decision Processes

(Extended Abstract)

Stefan J. Witwicki and Edmund H. Durfee
Computer Science and Engineering
University of Michigan
Ann Arbor, MI 48109
{witwicki, durfee}@umich.edu

Categories and Subject Descriptors

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

General Terms

Algorithms, Performance

Keywords

Multiagent Systems, Decentralized Markov Decision Processes, Event-Driven Interactions, Commitments

1. INTRODUCTION

Our work is motivated by cooperative planning problems where agents can affect each others' transitions and rewards, and so benefit from coordinating their actions, but in doing so must account for durational uncertainty in these actions. To reason about this uncertainty efficiently, agents can employ temporal decoupling (a paradigm that has been explored in a variety of restricted contexts [2, 3]) to constrain interactions to occur by selected time points, representing the uncertain occurrence for each time point with a probabilistic promise [5]. Here we summarize a reformulation of Becker's Event-driven DEC-MDP problems [1] that uses commitment models to exploit temporal structure. We argue that, in addition to representing optimal solutions, our approach enables more efficient, scalable computation of approximate solutions and a natural flexibility by which interactions can be modeled with more or less detail.

2. EVENT-DRIVEN DEC-MDP PROBLEMS

Figure 1 illustrates an example problem represented in the TAEMS description language (as discussed in [1]) involving two autonomous vehicle agents that must coordinate their execution of hierarchical tasks with uncertain durations so as to maximize expected quality within mission deadlines. We can model this problem using a *factored*, *locally fully observable*, *reward independent* flavor of Decentralized Markov Decision Processes with Event-Driven interactions [1] (or Event-Driven DEC-MDP for short). An Event-Driven interaction occurs through the satisfaction of a structured *dependency* d_{ij}, whereby the outcomes of actions taken by agent i change some of agent j's action choices or outcomes.

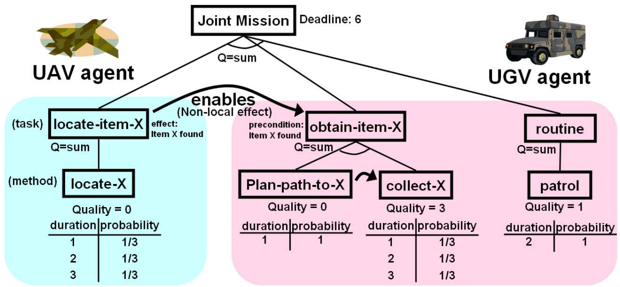

Figure 1: Autonomous Vehicle Example Problem

In our example problem from Figure 1, we can model the *enablement* effect between *locate-item-X* and *obtain-item-X* by a set of dependencies of the form d_{12}^{x,t_k}. For example, $d_{12}^{x,2}$ represents the dependency satisfied by *locate-item-X* completing at or before time 2. There is one such dependency for each time at which *obtain-item-X* could be enabled.

3. THE CURSE OF DIMENSIONALITY

Becker gains traction by exploiting structure to compute solutions more efficiently [1]. However, while the complexity of DEC-MDPs with Event-Driven Interactions is reduced from that of the general class of DEC-MDPs, it is still doubly exponential in the number of dependencies. Because his Coverage Set Algorithm (CSA) is an optimal solution method, its performance is critically affected by this exponential relationship. To combat the curse of dimensionality, we begin by establishing common structural properties that allow for reductions in complexity.

Definition 1. A **temporally uncertain interdependency** is a set of dependencies $X_{ij}^{e} = \left\{d_{ij}^{e,t_1}, d_{ij}^{e,t_2}, ..., d_{ij}^{e,t_k}, ...\right\}$, each of which has the same effect e at a different time t_k, with the property that when one dependency d_{ij}^{e,t_k} is fulfilled, all dependencies in X_{ij}^{e} with times $> t_k$ are also fulfilled.

This is an intuitive property of *enablement* interactions: it may be uncertain when a task will get enabled, but the task can be successfully executed at all time steps thereafter (subject to deadline constraints). With this additional

Cite as: Flexible Approximation of Structured Interactions in Decentralized Markov Decision Processes, (Extended Abstract), Stefan J. Witwicki, Edmund H. Durfee, *Proc. of 8th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2009)*, Decker, Sichman, Sierra and Castelfranchi (eds.), May, 10–15, 2009, Budapest, Hungary, pp. 1251–1252

structure, no matter how many dependencies are contained in X^e_{ij}, their histories can be represented compactly with a single Boolean satisfaction variable. Subsequently, Becker's augmented local state space only needs to be exponential in the number of temporally uncertain interactions (whereas before it was exponential in the number of dependencies).

Dependencies have a severe impact on the size of the CSA parameter space, the dimensionality of which is no less than the number of dependencies. Computing the coverage set entails computing the best response to every parameter setting, treating the space as if each parameter is independent of all others. But temporally uncertain interdependencies provide relationships that can be exploited. For example, if the probability of Agent 1 locating item X (from Figure 1) by time 3 is 0.5, the probability of it locating item X by time 4 must be at least 0.5. The approach that we overview in the next section takes advantage of these relationships to re-organize the parameter search space, mapping multiple related dependencies onto the same dimension.

4. COMMITMENTS

Commitment abstraction enables the problem of joint policy search to be reformulated as one of commitment-space search, whereby potential interactions are considered and then local policies computed and evaluated around these interactions. Here we extend our previous work [4, 5] to define event-related commitments.

Definition 2. A **commitment** $C(d_{ij}) = \rho$ is a guarantee that agent i will adopt a policy that in expectation will satisfy dependency d_{ij} for agent j with probability no less than ρ. Likewise, a **temporal commitment** $C\left(X^e_{ij}\right) = \langle \rho, t \rangle$ corresponds to the probabilistic satisfaction of a temporally uncertain interdependency X^e_{ij} by time t.

Commitments represent an agent's promises to fulfill other agents' nonlocal dependencies. For example, the UAV in Figure 1 can promise to enable the UGV's *obtain-item-X* task at time 2 with probability 2/3 by forming a commitment to the corresponding dependency. *Temporal commitments* represent more sweeping promises to satisfy all interaction dependencies whose times are greater than or equal to t. The UAV can promise, with a single temporal commitment of $\left\langle \frac{2}{3}, 2 \right\rangle$, to enable the UGV's *obtain-item-X* task at times 2, 3, 4, and 5.

A Commitment compactly conveys relevant nonlocal policy information that the dependent agent can, in turn, incorporate into its local state with a single variable $b \in \{T, F\}$ (where $b(s_i) = T$ implies that, in state s_i, the interdependency has been satisfied). A single temporal commitment accurately conveys $Prob(b|s_i)$ for states at time t but not at times before or after t. We can complete the model by expressing commitments at all potential satisfaction times.

Definition 3. A **complete temporal commitment set** $\bar{C}^{complete}\left(X^e_{ij}\right) = \left\{ C\left(d^{e,t_k}_{ij}\right) | d^{e,t_k}_{ij} \in X^e_{ij} \right\} = \{\langle \rho_k, t_k \rangle\}$ is a guarantee that agent i will perform actions so as to satisfy the temporally uncertain interdependency X^e_{ij} by each and every dependency time $t_k | d^{e,t_k}_{ij} \in X^e_{ij}$ with probability at least ρ_k respectively.

With *complete temporal commitment sets* come a sense of representational completeness in that agents can accurately model any dependency satisfaction scenario. Completeness has the consequence that a commitment search methodology can, in theory, be used to optimally solve Event-Driven problems with temporal structure. Unfortunately, for large problems, it will be infeasible to search the space of *temporal commitment sets* completely. Fortunately, commitments provide a framework for informed search techniques that find useful approximate solutions efficiently [5].

5. FLEXIBLE APPROXIMATION

So far, we have summarized two extremes of commitment modeling: *single commitments* and *complete commitment sets*. Agents can use a single commitment model to concisely capture the most critical time and probability for a particular interaction. But this compactness of representation comes at the cost of potential miscoordination (if the interaction occurs earlier or later than modeled) and hence suboptimal quality. The complete commitment set, on the other hand, will increase the nonlocal information contained in the augmented local models, and thereby increase the computational complexity of computing local policies, finding the right set of commitments, and solving the problem. But if the right set of commitments is found, the agents will coordinate optimally.

Which extreme is most suitable depends on the difficulty of the problem, the amount of planning time available, and the importance of high solution quality. The strength of using our commitment approach is that one need not pick either extreme. Our methodology is capable of representing a single commitment, all commitments, or a partial commitment set that balances quality and computation cost.

6. CONCLUSION

In this extended abstract, we have identified temporal structure common to many Event-Driven problems, and summarized a commitment-oriented reformulation that can take advantage of this structure. Commitments can compactly approximate nonlocal dependency information, suggesting a computational efficiency that will scale well with uncertainty. Moreover, by modeling interactions with more or less detail, they can also represent a range of approximate solutions, thereby enabling a flexible trade-off of computation time and solution quality.

7. ACKNOWLEDGMENTS

This material is based upon work supported by AFOSR under Contract No. FA9550-07-1-0262.

8. REFERENCES

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

[2] L. Hunsberger. Algorithms for a temporal decoupling problem in multi-agent planning. In *AI (Conference)*, pages 468–475, 2002.

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

[4] S. Witwicki and E. Durfee. Commitment-driven distributed joint policy search. *AAMAS*, 2007.

[5] S. Witwicki and E. Durfee. Commitment-based service coordination. *Int. Journal of Agent-Oriented Software Engineering*, 3(1):59–87, 2009.