Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)

Adapting Plans through Communication with Unknown Teammates

Trevor Sarratt
University of California Santa Cruz
tsarratt@soe.ucsc.edu

Introduction

Coordinating a team of autonomous agents is a challenging problem. Agents must act in such a way that progresses toward the achievement of a goal while avoiding conflict with their teammates. In information asymmetric domains, it is often necessary to share crucial observations in order to collaborate effectively. In traditional multiagent systems literature, these teams of agents share an identical design for reasoning, planning, and executing actions, allowing perfect modeling of teammates. Ad hoc teamwork (Stone et al. 2010) further complicates this problem by introducing a variety of teammates with which an agent must coordinate. In these scenarios, one or more agents within a team can be unfamiliar, having unknown planning capabilities guiding their behavior.

Much of the existing ad hoc teamwork research focuses on reinforcement learning and decision-theoretic planning. Agents use models of known behavior to predict an ad hoc agent's actions, using decision theory to maximize expected utility in instances where the predicted actions are uncertain. Online learning refines these models with observations of behaviors during execution, increasing the accuracy of the models' predictions, permitting the team to coordinate more effectively (Barrett et al. 2013).

My thesis addresses the problem of planning under teammate behavior uncertainty by introducing the concept of intentional multiagent communication within ad hoc teams. In partially observable multiagent domains, agents much share information regarding aspects of the environment such that uncertainty is reduced across the team, permitting better coordination. Similarly, we consider how communication may be utilized within ad hoc teams to resolve behavioral uncertainty. Transmitting intentional messages allows agents to adjust predictions of a teammate's individual course of action. In short, an ad hoc agent coordinating with an unknown teammate can identify uncertainties within its own predictive model of teammate behavior then request the appropriate policy information, allowing the agent to adapt its personal plan. The main contribution of this work is the characterization of the interaction between learning, communication, and planning in ad hoc teams.

Prior Work: Inference from Known Behaviors

Prior experience with a variety of agents can provide a set of known behavior models for future use. Observations of an unknown teammate can then be used to identify a similar existing model as a predictor of its future actions (Barrett et al. 2013). Our initial work in ad hoc teams considered the case where an agent coordinates with a teammate whose behavior is not represented well by a single model but rather by a composition of modeled behaviors. In this scenario, the teammate may adopt new strategies on the fly, adapting its behavior or pursuing new goals unexpectedly. Traditional probabilistic approaches assume static behavior and do not correct the choice of model in a responsive manner (Sarratt and Jhala 2015a; 2015c). Furthermore, perfect Bayesian inference would require either known transition probabilities, which are not provided in ad hoc domains, or a mechanism for learning the transition likelihoods. By adopting a technique from on-line learning (Bousquet and Warmuth 2003), we demonstrated a computationally lightweight technique for responsively identifying changes in a teammate's behavior, correctly selecting an alternate model that most accurately expresses the post-change observed actions. This approach outperforms the standard weight revision method, showing promise for accurately modeling agents of inconsistent behavior as well as agents whose high level strategy can be expressed by a sequence of more simple behaviors.

Such model inference approaches, while suitable for a wide range of problems, are not without their limitations, however. The possession of a set of known models requires either prior experience or hand-authored knowledge. Even when complex behavior can be approximated with the composition of known, simple models, the low level behaviors must be available prior to coordination. Initial work in learning a model during coordination has relied on generalization of information across known models to speed up the process (Barrett et al. 2013), and as a result, such transfer learning techniques may not apply to novel conditions in which no prior knowledge is available. Furthermore, relying on observations necessitates an act be performed, perhaps irreversibly, before the information can be utilized to infer or construct an accurate model, a constraint unsuitable for domains with strict limits on potential attempts of completing a task. For this reason, we motivate communicating *in advance* of a point of uncertainty.

Current Work: Communication

Across many communicative multiagent frameworks, such as the COM-MTDP model (Pynadath and Tambe 2002) and STEAM (Tambe 1997), communicative actions are often limited to sharing observations. As agents in such systems have complete information regarding the planning capacities of their teammates, they can simply use the shared information to compute precisely how all agents will act. Since the policies of teammates is the source of uncertainty in ad hoc teams, it follows that policy information is a promising target for communicative acts. We introduced this idea in (Sarratt and Jhala 2015b).

In early decision-theoretic agent communication literature, various types of communication were theoretically validated in their effect on coordinating multiple agents. These included intentional messages, questions, proposals, threats, imperatives, and statements of propositional attitudes (Gmytrasiewicz, Durfee, and Rosenschein 1995). In each case, providing or requesting information adjusted one or more agents' policies through refining an agent's expectations of either its own policy's likelihood of success or the intentions of another agent acting within the same environment. Analogously, the refinement of predicted action probabilities and, consequently, an improved policy for a coordinating agent is desirable for ad hoc teams.

Whereas the broadcast of the intention of pursuing a goal addresses multiple state-action pairs within an agent's policy computation, we must consider that an unfamiliar teammate may not possess the capability of reasoning with high level abstractions such as joint plans or hierarchical goal decompositions. However, we put forth the observation that all agents involved are universally tasked with assigning actions to states, independent of the particular planning implementation details. From a general perspective, we consider how an ad hoc agent could benefit from obtaining a single state-action pair—the atomic component of a teammate model—from communication.

In our submission to *AAMAS-16* (Sarratt and Jhala Pending), we provided an efficient procedure for evaluating potential state-action pairs for communication by examining the uncertainty within a teammate model as well as empirically analyzed various aspects of the communicative capability. For our test domain, we used a variation of the multiagent pursuit domain where two agents attempt to capture a prey within a maze. We were able to show the tradeoff between collected information and communication rates and, likewise, the effect of communication costs on query rates and the resulting expected utility of the agent. Finally, our analysis determined that branch points in a maze are commonly communicated more frequently than neighboring cells, indicating that the tested agents determined resolving uncertainty at such states was associated with higher utility.

Future Work

While our previous work has focused on the ability to coordinate with computer agents, we intend to test our approach with human teammates, given the potential applicability of ad hoc team approaches to human-computer teams. Furthermore, over the ensuing year, we intend to transition to a more complex domain, using a slightly simplified form of the real-time strategy game, StarCraft.

An immediate extension to this work is the consideration of communicating multiple state-action pairs without independent evaluation. It is possible for two states to have no utility for communication individually but have non-zero utility when considered together. This opens up a combinatorial space of potential intentional information sets that could be communicated, similar to problem of picking a subset of observations to share within a team, as explored by Roth et al. 2006. Due to the intractable nature of the problem, the authors motivated the exploration of heuristics as approximate solutions. It is yet unclear how states in such collections will be related, though we hypothesize they could take any one of various forms, including sequences of successive states (forming a plan), areas of connected states (a local subspace of the domain), or groups of independently valuable states (such as branch points discussed earlier).

References

Barrett, S.; Stone, P.; Kraus, S.; and Rosenfeld, A. 2013. Teamwork with limited knowledge of teammates. In *AAAI*.

Bousquet, O., and Warmuth, M. K. 2003. Tracking a small set of experts by mixing past posteriors. *The Journal of Machine Learning Research* 3:363–396.

Gmytrasiewicz, P. J.; Durfee, E. H.; and Rosenschein, J. 1995. Toward rational communicative behavior. In *AAAI Fall Symposium on Embodied Language*, 35–43.

Pynadath, D. V., and Tambe, M. 2002. The communicative multiagent team decision problem: Analyzing teamwork theories and models. *Journal of Artificial Intelligence Research* 389–423.

Roth, M.; Simmons, R.; and Veloso, M. 2006. What to communicate? execution-time decision in multi-agent pomdps. In *Distributed Autonomous Robotic Systems 7*. Springer. 177–186.

Sarratt, T., and Jhala, A. 2015a. Rapid: A belief convergence strategy for collaborating with inconsistent agents. In *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Sarratt, T., and Jhala, A. 2015b. The role of models and communication in the ad hoc multiagent team decision problem. In *Proceedings of the Third Annual Conference on Advances in Cognitive Systems Poster Collection*, 27.

Sarratt, T., and Jhala, A. 2015c. Tuning belief revision for coordination with inconsistent teammates. In *Eleventh Artificial Intelligence and Interactive Digital Entertainment Conference*.

Sarratt, T., and Jhala, A. Pending. Communicating intentions for coordination with unknown teammates. In *Thirtieth AAAI Conference on Artificial Intelligence*.

Stone, P.; Kaminka, G. A.; Kraus, S.; Rosenschein, J. S.; et al. 2010. Ad hoc autonomous agent teams: Collaboration without pre-coordination. In *AAAI*.

Tambe, M. 1997. Agent architectures for flexible. In *Proc. of the 14th National Conf. on AI, USA: AAAI press*, 22–28.