In *The Twenty-Fourth Conference on Artificial Intelligence* (AAAI),
July 2010.

Ad Hoc Autonomous Agent Teams: Collaboration without Pre-Coordination

Peter Stone
U. of Texas at Austin
pstone@cs.utexas.edu

Gal A. Kaminka
Bar-Ilan U.
galk@cs.biu.ac.il

Sarit Kraus
Bar-Ilan U. & U. of Maryland
sarit@cs.biu.ac.il

Jeffrey S. Rosenschein
Hebrew U.
jeff@cs.huji.ac.il

Abstract

As autonomous agents proliferate in the real world, both in software and robotic settings, they will increasingly need to band together for cooperative activities with previously unfamiliar teammates. In such *ad hoc team* settings, team strategies cannot be developed a priori. Rather, an agent must be prepared to cooperate with many types of teammates: it must collaborate without pre-coordination. This paper challenges the AI community to develop theory and to implement prototypes of ad hoc team agents. It defines the concept of ad hoc team agents, specifies an evaluation paradigm, and provides examples of possible theoretical and empirical approaches to challenge. The goal is to encourage progress towards this ambitious, newly realistic, and increasingly important research goal.

1 Introduction

Imagine that you are in a foreign country where you do not speak the language, walking alone through a park. You see somebody fall off of his bicycle and injure himself badly; there are a few other people in the area, and all of you rush to help the victim. There are several things that need to be done. Somebody should call an ambulance, someone should check that the victim is still breathing, and someone should try to find a nearby doctor or policeman. However, none of you know one another, and thus you do not know who has a mobile phone, who is trained in first aid, who can run fast, and so forth. Furthermore, not all of you speak the same language. Nonetheless, it is essential that you quickly coordinate towards your common goal of maximizing the victim's chances of timely treatment and survival.

This scenario is an example of what we call an *ad hoc team* setting. Multiple agents (in this case humans) with different knowledge and capabilities find themselves in a situation such that their goals and utilities are perfectly aligned (effectively, everyone's sole interest is to help the victim), yet they have had no prior opportunity to coordinate. In addition to the emergency setting described above, ad hoc teams may arise among robots or software agents that have been programmed by different groups and/or at different times such that it was not known at development time that they would need to coordinate. For example, rescue robots may be brought to an earthquake site from different parts of the world, or an e-commerce agent may need to coordinate with other legacy agents that can no longer be altered. Note that in this latter example, the agents may need to coordinate repeatedly on the same or similar tasks.

In order to be a good "ad hoc team player," an agent must be adept at assessing the capabilities of other agents, especially in relation to its own capabilities. If you are trained in first aid, you may be the best person to examine the fallen bicyclist. But if one of the other people is a doctor, you should take on a different role. Similarly, a good team player must also be adept at assessing the other agents' knowledge states (does anybody know the right phone number?). Furthermore, it must be proficient at estimating the effects of its actions on the other agents. How will the others react if you immediately run away, if you pull out a mobile phone, if you start screaming, or if you calmly measure the victim's pulse?

Ad Hoc Human Teams

The concept of ad hoc *human* teams has arisen recently in military and industrial settings, especially with the rise of outsourcing. There have also been autonomous agents developed to help support human ad hoc team formation (Just, Cornwell, & Huhns 2004; Kildare 2004). But ad hoc *autonomous agent* teams have not been relevant in the past because autonomous agents, especially robots, have tended to be deployed only for short times, and teams have been developed by cohesive development groups. As a result, it has typically been possible (and usually necessary) to adjust and tune the agents' behaviors so that they interact well with one another.

Ad Hoc Autonomous Agent Teams

As the field progresses towards long-lasting autonomy, however, reasoning on the fly about interactions with other agents will become increasingly essential. Unlike most team settings considered so far (e.g., (Stone & Veloso 1999; Rosenfeld *et al.* 2008)), agents in ad hoc team settings are not all programmed by the same people, and may not all have the same communication protocols or world models. Furthermore, they are likely to have heterogeneous sensing and acting capabilities that may not be fully known to one another. As a result, team strategies cannot be developed a priori. Rather, an agent that is to succeed in such an *ad hoc team* setting must be prepared to adjust its behavior to interact as well as possible with many types of teammates: those with which it can communicate and those with which it cannot; those that are more mobile and those that are less mobile; those with better sensing capabilities and those with worse capabilities. A good team player's best actions are likely to differ significantly depending on the characteristics of its teammates. The fact that humans are routinely called upon to coordinate in an ad hoc fashion strongly motivates the challenge of constructing autonomous agents of similar flexibility.

The Challenge

Our challenge to the community identifies a specific, novel, high-risk, but high-payoff research area. Specifically, we call for theoretical treatments and concrete implementations of robust, autonomous agents that are good ad hoc team players.

That is, we challenge the community:

To create an autonomous agent that is able to efficiently and robustly collaborate with previously unknown teammates on tasks to which they are all individually capable of contributing as team members.

The remainder of this paper is organized as follows. First, Section 2 gives further insight into the details of the challenge via a specification of how potential solutions can be evaluated. Then, Sections 3 and 4 discuss possible theoretical and empirical approaches to the challenge, respectively. Section 5 expands on the ways in which the challenge can be decomposed and/or the scope of the challenge can be gradually increased, Section 6 touches on prior research most related to the challenge, and Section 7 concludes.

2 Evaluation

Though there is plenty of room for theoretical treatment of the problem (see Section 3), the challenge introduced in Section 1 is ultimately an empirical challenge. Additionally, though it pertains to teamwork, it is fundamentally a challenge pertaining to building a *single* autonomous agent. In this section, we shed further light on the intention of the challenge by specifying a way in which potential solutions can be evaluated.

Let a_0 and a_1 be two ad hoc teammates whose performance is to be compared in domain D, from which tasks d can be sampled. For example, D may be a multiagent planning domain with each task having different initial conditions and/or goals; or D may be robot soccer, with each task being a match against a particular opponent team. It may also be the case that there is only one task in the domain that is repeatedly "sampled."

Assume that there is some quantitative performance measure, or "score" $s(B, d)$, that results when a set of agents B performs task d once, such as time to goal achievement, joint discounted reward over a fixed time, number of goals scored, or any other objective measure. Note that s may be a stochastic function.

Let there be a pool of potential teammates $A = \{a_2, \ldots, a_n\}$, each of which has some competency in domain D. Specifically, we say that agent $a \in A$ is competent if there is some subset $B \subset A$ such that $a \in B$, and the team comprised of the agents in B is able to achieve a minimal threshold expected performance, s_{min}, on all tasks in D:[1]

$$\forall a \in A, \forall d \in D, \exists B \subset A \text{ s.t. } a \in B \wedge E[s(B, d)] \geq s_{min}$$

Note that it may be possible for an *individual* agent to achieve s_{min}, for instance in a domain such as foraging where having teammates is helpful but not essential; or D may be a fundamentally multiagent domain in which agent *teams* are needed to perform the task at all, such as pushing heavy boxes. Note further that the agents in A need not be themselves aware that they are acting as teammates; indeed when they *are* "team aware" the interactions among the agents can become considerably more complex.

We propose comparing agents a_0 and a_1 as potential ad hoc teammates of the agents in set A in domain D according to the following procedure.

Evaluate(a_0, a_1, A, D)

- Initialize performance (reward) counters r_0 and r_1 for agents a_0 and a_1 respectively to $r_0 = r_1 = 0$.
- Repeat:
  - Sample a task d from D.
  - Randomly draw a subset of agents B, $|B| \geq 2$, from A such that $E[s(B, d)] \geq s_{min}$.
  - Randomly select one agent $b \in B$ to remove from the team to create the team B^-.
  - increment r_0 by $s(\{a_0\} \cup B^-, d)$
  - increment r_1 by $s(\{a_1\} \cup B^-, d)$
- If $r_0 > r_1$ then we conclude that a_0 is a better ad-hoc team player than a_1 in domain D over the set of possible teammates A.

Note that there is a lot of potential variability both in the breadth of the domain D (how different the tasks can be) and especially in the breadth of teammate capabilities in A. We address these issues in more detail in Section 5. Note that we assume that agents a_0 and a_1 are aware of the domain D and the set of potential teammates A. But A may have infinite cardinality, in effect just placing bounds on teammate characteristics, such as the teammate will not be able to move faster than 2 m/s. In addition, even if A is finite, on each iteration the set B^- is initially unknown to the ad hoc team agents being evaluated.

This evaluation paradigm serves to emphasize that an ad hoc team agent is fundamentally an *individual* agent. To be successful, it must perform well with any set of teammates with which it is presented.

[1] For notational convenience, we assume that larger performance values indicate better task performance.

3 Example Theoretical Approach

Although our challenge is ultimately empirical, there is also ample room for theoretical analysis of components of the full problem. For example, aspects of ad hoc teamwork can be usefully studied within the framework of game theory (Leyton-Brown & Shoham 2008). Specifically, a good ad hoc team agent should be able to learn to interact with a previously unknown teammate in a fully cooperative (common payoff) iterative normal form game. If the teammate plays a fixed (possibly stochastic) strategy, the ad hoc team agent should simply learn what that strategy is and play the best response. But even in this simplest of scenarios, the problem can become quite intricate to analyze if the teammate may itself be adaptive.

Collaborative Multi-Armed Bandits

As an initial theoretical analysis of an aspect of ad hoc teamwork, Stone and Kraus (2010) consider a situation in which the ad hoc team player interacts repeatedly in a stochastic environment with a teammate that is both less capable and less knowledgeable than itself. Specifically, the teammate can only execute a subset of the actions that the ad hoc team agent can execute, and, unlike the ad hoc team agent, it is unaware of the relative utilities of the various actions. It is also unaware that it is acting as a part of a team. They formalize this situation as an instance of the well-studied k-armed bandit problem (Robbins 1952).

The basic setting of the k-armed bandit problem is as follows. At each time step, a learning agent selects one of the k arms to pull. The arm returns a payoff according to a fixed, but generally unknown, distribution. The agent's goal is to maximize the sum of the payoffs it receives over time. The setting is well-suited for studying exploration vs. exploitation: at any given time, the agent could greedily select the arm that has paid off the best so far, or it could select a different arm in order to gather more information about its distribution. Though k-armed bandits are often used for this purpose, the authors were the first to consider a multiagent cooperative setting in which the agents have different knowledge states and action capabilities.

In order to study the ad hoc team problem, the authors extend the standard setting to include two distinct agents, known as the *teacher* and the *learner*, who select arms alternately, starting with the teacher. They initially consider a bandit with just three arms such that the teacher is able to select from any of the three arms, while the learner is only able to select from among the two arms with the lower expected payoffs. The authors consider the fully cooperative case such that the teacher's goal is to maximize the expected sum of the payoffs received by the two agents over time (the teacher is risk neutral). Specifically, the authors make the following assumptions:

- The payoff distributions of all arms are fully known to the teacher, but unknown to the learner.
- The learner can only select from among the two arms with the lower expected payoffs.
- The results of all actions are fully observable (to both agents).
- The number of rounds (actions per agent) remaining is finite and known to the teacher.
- The learner's behavior is fixed and known: it acts greedily, always selecting the arm with the highest observed sample average so far. If there are any previously unseen arms, the learner selects one of them randomly (optimistic initialization).

The teacher must then decide whether to do what is best in the short term, namely pull the arm with the highest expected payoff; or whether to increase the information available to its teammate, the learner, by pulling a different arm. Note that if the teacher were acting alone, trivially its optimal action would be to always pull the arm with highest expected payoff.

By these assumptions, the learner is both less capable and less knowledgeable than the teacher, and it does not understand direct communication from the teacher. It is tempting to think that we should begin by improving the learner. But in the ad hoc team setting, that is not an option. The learner "is what it is" either because it is a legacy agent, or because it has been programmed by others. Our task is to determine the teacher's best actions *given* such learner behavior.

This setting is only a limited representation of the full ad hoc team setting from Section 1. However it retains the most essential property, namely that a single agent in our control must interact with a teammate without the advance coordination. Nonetheless, even this initial scenario presents some interesting mathematical challenges. Specifically, Stone and Kraus prove several theoretical results pertaining to which arms the teacher should consider pulling, and under what conditions (including for the natural generalization with more than three arms). Furthermore, when the payoffs from the arms are discrete, they present a polynomial algorithm for the teacher to find the optimal arm to pull. (Stone & Kraus 2010).

The study of collaborative k-armed bandits described in this section serves as a starting point for the theoretical analysis of ad hoc teams. However it leaves open many directions for extensions: situations in which the teacher does not have full knowledge, the learner is "team aware," the number of iterations is not known, and/or the learner's behavior is not known a priori to the teacher, among others. We hope that this ad hoc teamwork challenge will inspire many new and challenging problems in the theoretical analysis of optimal control and optimal teamwork.

4 Example Empirical Approach

While theoretical analyses are likely to be able to identify some situations in which optimal decision-making is possible, they are also likely to identify aspects of ad hoc teamwork that are not tractably solvable. There are also likely to be aspects of ad hoc teamwork that are not easily analyzable at all. In such aspects of the problem, there will be plenty of room for useful empirical analyses. In this section, we illustrate a possible empirical approach using the domain of robot soccer.

Human Soccer

Human soccer is a perfect example of the ad hoc team setting. When given the opportunity, *teams* of human soccer players train together and hone their interactions so as to refine their ability to cooperate with one another. However *individual* soccer players are also able to join "pick-up" games with players that they've never met before, let alone played with. They can even do so in foreign countries where they don't know the language, leaving no way to communicate with their ad hoc teammates other than mutual observation and experimentation. A talented human player is able to make quick judgments about how she will best fit into her ad hoc team. When playing with worse players, she is likely to play in the center of the field; when playing with better players she may look for a supporting role that limits responsibility. Furthermore, at the beginning of the game, it may be useful in this situation to take actions that highlight one's particular strengths, such as kicking with the left foot, or passing the ball precisely, so as to *teach* one's teammates how best to incorporate the newcomer into the team.

Robot Soccer

Similarly, robot soccer teams at the international RoboCup competitions are typically developed as cohesive units with communication protocols and sophisticated methods for distributing the players on the field into complementary positions (Stone & Veloso 1999). However, it is also possible to consider a "pick-up" game in which the players are not able to pre-coordinate.

As an instantiation of the bandit example in Section 3, one could consider a center midfielder, who can pass to any of three forwards, teaching an outside midfielder, who can only pass to two of them. If the outside midfielder is new to the team and still learning about the forwards' capabilities, we find ourselves in an instantiation of exactly the abstract k-armed bandit scenario. Just as in the abstract setting, it is naturally extensible to model sophisticated learners, partial teacher knowledge, unknown learner behavior, and so on.

Eventually, one can imagine staging field tests of ad hoc team agents at the annual RoboCup competitions. Several participants with robots of varying strengths could be invited to participate in a robot soccer "pick-up game." That is, the robots would be placed on the field as a team without any prior coordination among the human programmers.

A successful ad hoc team player will be able to quickly evaluate whether it is playing with forwards or defenders, whether it is playing with more skillful players or with less skillful players, etc., and adjust its play accordingly. If it is placed on a team with no goalie, then it should notice and adopt that role; if it is placed on a team with worse players, it should actively go to the ball more often, and so on.

The essential aspect is that the ad hoc team player should be able to deal with whatever teammates it might come across, and without any foreknowledge of the teammates' actual behaviors on the part of the agents or the programmers.

5 Controlling the Scope: Task and Teammate Breadth

An ad hoc team player must be prepared to collaborate with all types of teammates. Thus one possible view of the process of creating a fully capable ad hoc team player is that it is akin to equipping it with a toolbox, each tool being useful for interacting with a class of possible teammates, as well as with a method for identifying to which class the current teammates belong.

From this perspective, in order to create an ad hoc team player, one will need to address three high-level technical challenges.

1. Identify the full range of possible teamwork situations that a complete ad hoc team player needs to be capable of addressing.
2. For each such situation, find theoretically optimal and/or empirically effective algorithms for behavior.
3. Develop methods for identifying and classifying which type of teamwork situation the agent is currently in, in an online fashion.

Challenges 2 and 3 are the core technical aspects of the challenge. But the first can also be seen as a sort of knob, which can be used to incrementally increase the difficulty of the challenge.

For this purpose, we start from the literature on ad hoc human team formation (Kildare 2004) to organize teamwork situations along three dimensions:

Teammate characteristics: features of the individual teammates such as action capabilities, sensing capabilities, decision making and learning capabilities, whether they can communicate directly, and prior knowledge.

Team characteristics: features of the collection of team members such as whether they are homogeneous or heterogeneous, how many teammates are on the team, and whether they can observe each other's actions.

Task characteristics: features of the cooperative task to be performed such as the goal, the time horizon, whether it is turn-taking, and how closely coordinated the agents need to be in their actions. Can they divide the task at a high level and then act independently, or do they need to coordinate low-level actions?

By initially limiting teammate, team, and task characteristics (A and D from Section 2), we can render the challenge approachable even though the full-blown version is quite ambitious. For example, Stone and

Kraus's k-armed bandit scenario summarized in Section 3 is appropriate for situations in which the *teammates* have limited action capabilities, perfect sensing, greedy decision making, no direct communication, and prior knowledge limited to their own observations; the *team* is heterogeneous, consists of two agents, and can fully observe each other's actions; and the *task*'s goal is to maximize the sum of discrete action utilities over a finite horizon where the agents act individually in a turn-taking fashion. In that work, the authors found the theoretically optimal action for the ad hoc team player, thus taking a first step towards research challenge 2.

We expect that the initial responses to this challenge will address subproblems by similarly limiting the scopes of A and D. Indeed, most of the examples given in this paper consider sets A such that the potential teammates are not even necessarily aware that they are a part of an ad hoc team; this case is the simplest to consider. However, it will be an important aspect of the challenge to consider sets A that include other agents that are aware that they are on an ad hoc team. Eventually, there will be opportunities to generalize and/or combine these subproblems into a more complete ad hoc team agent.

6 Discussion and Related Work

This challenge is predicated on the assumption that software agents and robots will soon be able to be deployed for extended periods of time in the real world. That is, their usefulness will outlive our ability to easily change their behaviors. Such a phenomenon has already occurred with conventional computer programs, as became apparent when the world was worried about the "Y2K bug." COBOL programmers were called out of retirement to try to reprogram computers that were essential to business processes and vital infrastructure, but that were black boxes to everyone who used them.

Were autonomous agents to become similarly long-lived, it would be a huge landmark in their robustness and reliability. However it would also expose us to the problem that this challenge addresses. Namely, a new agent may need to collaborate with older agents whose behavior is already fixed and not easily changeable.

Ad hoc teams are also already needed for environments where agents with diverse capabilities and no common framework must quickly work as a team. As presented in Section 1, one example is when robots from different developers come together on a common rescue mission. A second example arises when software agents, programmed in isolation, must act within a team setting. These agents might need to analyze scheduling data from different people to help coordinate meetings on their behalf, or they might need to coordinate with legacy agents that can no longer be altered.

The main focus of this research challenge is ad hoc teams in which teammates need to work together without any prior coordination. This perspective is at odds with most prior treatments of teamwork, such as SharedPlans (Grosz & Kraus 1996), STEAM (Tambe 1997), and GPGP (Decker & Lesser 1995) which define explicit coordination protocols, languages, and/or shared assumptions about which the agents are mutually aware. In applications such as the annual RoboCup robot soccer competitions, entire teams of agents are designed in unison, enabling explicit pre-coordination via structures such as "locker room agreements" (Stone & Veloso 1999).

Other than the multi-armed bandit work previously described, the work that we are aware of that takes a perspective most similar to ad hoc teams is that of Brafman and Tennenholtz (Brafman & Tennenholtz 1996) in which they consider a teacher agent and a learner agent repeatedly engaging in a joint activity. While the learner has no prior knowledge of this activity, the teacher understands its dynamics. However they mainly consider a situation in which teaching is not costly: the goal of their teacher is to maximize the number of times that the learner chooses the "right" action. Thus in some sense, the teacher is not "embedded" in the environment as a real *teammate*.

Although the complete challenge put forth in this paper is very ambitious and likely to take many years to meet in a fully satisfactory way, there are numerous existing techniques that may be useful starting points for certain aspects of the challenge (e.g., for certain properties of A and D). The remainder of this section provides a small sampling of such existing techniques.

- As mentioned in Section 3, game theory (Leyton-Brown & Shoham 2008) provides a useful theoretical foundation for multiagent interaction. Though originally intended as a model for human encounters (or those of human institutions), it has become much more broadly applied over the last several decades.

- A good ad hoc team player may need to make an explicit assumption that its teammates are observing and reacting to its actions (that they are "team aware"). In doing so, the agent is actually planning its actions *intending* for them to be observed and interpreted. *Intended plan recognition* (in contrast to *keyhole recognition*) is the term used when the observed agent knows that it is being observed, and is acting under the constraints imposed by this knowledge (Carrbery 2001). Much of the work on planning for intended recognition settings has focused on natural language dialogue systems (Sidner 1985; Lochbaum 1991).

- An important aspect of the ad hoc team challenge is recognizing, or "modeling" the capabilities of one's teammates. For this purpose, work from the *opponent modeling* literature (e.g., (Carmel & Markovitch 1995; Stone, Riley, & Veloso 2000; Oshrat, Lin, & Kraus 2009)) may be readily adaptable to similarly model teammates.

  In addition, a good ad hoc team agent must recognize the possibility that, while it is attempting to model its teammates, the teammates may be simultaneously modeling it. In such a case, the agent is engaged in a *recursive modeling* setting (Vidal & Durfee 1995).

- Claus and Boutilier (Claus & Boutilier 1998) show how reinforcement learning can provide a robust method for agents to learn how to coordinate their actions. This work is one of many approaches for cooperative multiagent learning (see surveys at (Stone & Veloso 2000; Panait & Luke 2005)).

In addition to existing AI methods, as mentioned in Section 1, previous work has examined the use of agents to support the formation of *human* ad hoc teams (Just, Cornwell, & Huhns 2004; Kildare 2004). This work relies on an analysis of the sources of team variability, including member characteristics, team characteristics, and task characteristics (Kildare 2004), which we borrow as a structure for classifying types of autonomous teammates in Section 5. In addition, software agents have been used to support the *operation* of human teams (Chalupsky *et al.* 2001), and for distributed information gathering from distinct, otherwise independent information sources (Sycara *et al.* 1996). But we are not aware of any such work that enables an autonomous agent to itself act as an ad hoc teammate with previously unknown teammates.

7 Conclusion

This paper presents the concept of ad hoc autonomous agent teams and challenges the community to develop novel theoretical and empirical approaches to creating effective ad hoc teammates. Though today most agent teams are developed as a unit, we believe that it will not be long before autonomous agents, both in software and robotic settings, will very often need to band together on the fly (possibly with humans on their teams as well!). We hope that this challenge will encourage the research necessary to ensure that agents will be able to do so effectively when that time arrives.

Acknowledgments

Thanks to the UT Austin Learning Agents Research Group for useful comments and suggestions. This work was partially supported by grants from NSF (IIS-0917122, IIS-0705587), DARPA (FA8650-08-C-7812), ONR (N00014-09-1-0658), FHWA (DTFH61-07-H-00030), Army Research Lab (W911NF-08-1-0144), ISF (1357/07, 898/05), and the Fulbright and Guggenheim Foundations.

References

Brafman, R. I., and Tennenholtz, M. 1996. On partially controlled multi-agent systems. *JAIR* 4:477–507.

Carmel, D., and Markovitch, S. 1995. Opponent modeling in a multi-agent system. In Sen, S., ed., *IJCAI-95 Workshop on Adaptation and Learning in Multiagent Systems*, 8–13.

Carrbery, S. 2001. Techniques for plan recognition. *User Modeling and User-Adapted Interaction* 11:31–48.

Chalupsky, H.; Gil, Y.; Knoblock, C.; Lerman, K.; Oh, J.; Pynadath, D.; Russ, T.; and Tambe, M. 2001. Electric elves: Applying agent technology to support human organizations. In *IAAI-01*.

Claus, C., and Boutilier, C. 1998. The dynamics of reinforcement learning in cooperative multiagent systems. In *AAAI-98*, 746–752. AAAI Press.

Decker, K. S., and Lesser, V. R. 1995. Designing a family of coordination algorithms. In *ICMAS-95*, 73–80. Menlo Park, California: AAAI Press.

Grosz, B. J., and Kraus, S. 1996. Collaborative plans for complex group actions. *AIJ* 86:269–358.

Just, J.; Cornwell, M.; and Huhns, M. 2004. Agents for establishing ad hoc cross-organizational teams. In *IEEE/WIC/ACM International Conference on Intelligent Agent Technology*, 526–30.

Kildare, R. 2004. Ad-hoc online teams as complex systems: agents that cater for team interaction rules. In *Proceedings of the 7th Asia-Pacific Conference on Complex Systems*.

Leyton-Brown, K., and Shoham, Y. 2008. *Essentials of Game Theory: A Concise, Multidisciplinary Introduction*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan and Claypool Publishers.

Lochbaum, K. E. 1991. An algorithm for plan recognition in collaborative discourse. In *ACL*, 33–38.

Oshrat, Y.; Lin, R.; and Kraus, S. 2009. Facing the challenge of human-agent negotiations via effective general opponent modeling. In *AAMAS-09*, 377–384.

Panait, L., and Luke, S. 2005. Cooperative multi-agent learning: The state of the art. *AAMAS-05* 11:387–434.

Robbins, H. 1952. Some aspects of the sequential design of experiments. *Bulletin American Mathematical Society* 55:527–535.

Rosenfeld, A.; Kaminka, G. A.; Kraus, S.; and Shehory, O. 2008. A study of mechanisms for improving robotic group performance. *AIJ* 172(6-7):633–655.

Sidner, C. L. 1985. Plan parsing for intended response recognition in discourse. *Computational Intelligence* 1(1).

Stone, P., and Kraus, S. 2010. To teach or not to teach? decision making under uncertainty in ad hoc teams. In *AAMAS-10*. International Foundation for Autonomous Agents and Multiagent Systems.

Stone, P., and Veloso, M. 1999. Task decomposition, dynamic role assignment, and low-bandwidth communication for real-time strategic teamwork. *AIJ* 110(2):241–273.

Stone, P., and Veloso, M. 2000. Multiagent systems: A survey from a machine learning perspective. *Autonomous Robots* 8(3):345–383.

Stone, P.; Riley, P.; and Veloso, M. 2000. Defining and using ideal teammate and opponent models. In *IAAI-00*.

Sycara, K.; Decker, K.; Pannu, A.; Williamson, M.; and Zeng., D. 1996. Distributed intelligent agents. *IEEE Expert* 11(6).

Tambe, M. 1997. Towards flexible teamwork. *JAIR* 7:81–124.

Vidal, J. M., and Durfee, E. H. 1995. Recursive agent modeling using limited rationality. In *ICMAS-95*, 125–132. AAAI/MIT press.