In *Artificial Intelligence* (AIJ),
171:402-05, January 2007.

Multiagent learning is not the answer. It is the question.

Peter Stone
Department of Computer Sciences
The University of Texas at Austin
1 University Station C0500, Austin, Texas 78712-1188
pstone@cs.utexas.edu
http://www.cs.utexas.edu/~pstone

Artificial Intelligence, 117:402–05, 2007.

Abstract

The article by Shoham, Powers, and Grenager called "If multi-agent learning is the answer, what is the question?" does a great job of laying out the current state of the art and open issues at the intersection of game theory and artificial intelligence (AI). However, from the AI perspective, the term "multiagent learning" applies more broadly than can be usefully framed in game theoretic terms. In this larger context, how (and perhaps whether) multiagent learning can be usefully applied in complex domains is still a large open question.

Shoham, Powers, and Grenager set for themselves the worthwhile goal of starting a discussion in the field regarding the definition, goals, and evaluation criteria of multiagent learning. I agree with them entirely that it is useful to step back and identify the existing and possible research agendas in the field, to try to classify existing research, to provide a vocabulary for classifying research to come, and to identify the challenging open questions. They provide an excellent starting point towards all of these ends. As evidenced by this special issue, their article has already accomplished their goal.

This response is mainly concerned with the emphasis of their article. Specifically, the authors' disclaimers notwithstanding, the article couches the area of MAL as addressable within the formal framework of game theory. In doing so, the article can be seen as portraying a potentially very broad AI research area in somewhat limited terms. Though the authors do acknowledge that not all MAL research falls within their specific focus on stochastic games, the exceptions they cite are still game theoretic in nature (specifically extensive-form games of incomplete and/or imperfect information).

While there is certainly a great deal of interesting and relevant MAL research that is indeed characterizable within the language of game theory, much of which is cited in the article, it is important to acknowledge that the tools and language of game theory only go so far. The authors do include caveats acknowledging this notion to some extent and are careful not to make any claims of being comprehensive in their survey of relevant research. However, if this discussion is truly intended to address all of MAL, it is important to give first class status in the agenda and taxonomy to work that is not usefully characterizable in game theoretic terms.

What makes a problem not usefully characterizable within game theoretic terms? In principle, every multiagent encounter can be characterized as a normal form or extensive form game. But in some cases, it is not only that the "convergence to an equilibrium is not a goal in and of itself," but that the very formulation of the encounter as a normal form or extensive form game, if even practical, does little to make progress towards a solution.

To draw an example from my own research, soccer is undoubtedly a multiagent encounter. Both in the real game and in the Soccer Server system [7] used at RoboCup,[1] every player has 10 teammates and 11 opponents, each acting independently. The decisions faced by the players, such as when and where to kick the ball, or where to move when not in possession of the ball are continuous in nature, are based on incomplete information, are highly stochastic, must be made in quick succession (10 times per second), have strong sequential dependencies, and may depend on the similarly complex and rapid decisions of 21 other teammates and adversaries. Though none of these properties is individually outside the realm of game theory, in practice, the scale (or *complexity* as it is called by Shoham et al.) of the problem is such that there's not much hope in trying to identify any sort of equilibrium or any other optimal solution concept for this interaction, at least given current methods. Indeed, considering robot soccer from a game theoretic perspective would be much like considering it from the perspective of POMDPs. Formally, yes, robot soccer is a partially observable Markov decision process. But the known algorithms for solving POMDPs fall short of scaling to such a problem by many orders of magnitude. Such a multiagent learning problem must be approached from a different perspective.

In that case, from what perspective should these more complex multiagent learning problems be approached? Indeed, that is the relevant question. There is no single correct multiagent learning algorithm — each problem must be considered individually. And in many cases, the question is still whether it is possible at all. Multiagent learning is the question — not the answer.

For example, in my book *Layered Learning in Multiagent Systems: A Winning Approach to Robotic Soccer* [10], the principal question addressed (as stated in Chapter 1) is "*Can agents learn to become individually skilled and to work together in the presence of both teammates and adversaries in a real-time, noisy environment with limited communication?*" The book proceeds to answer the question affirmatively, but the learning is fairly limited in scope. Indeed a main challenge addressed therein, and in any similarly complex problem domain where learning of a complete decision function is not feasible, is which aspects of the

[1] An international robot soccer initiative that hosts an annual competition. [3, 11, 6]

problem should be learned, and how they should be learned. In the book's case, the agents learn how to pass and where to pass in the presence of specific adversaries, but without building any explicit model of the effects of their own actions or the likely opponent actions.

That book is just one example of many multiagent learning problems that have been considered using non-game-theoretic approaches, and arguably that should not be considered game theoretically. A partial list of other examples includes collaborative multi-robot localization [2], distributed network routing [4], distributed factory optimization [9], in-city driving [8], tracking teams of enemy combatants [13], and bidding in auctions [12].

Bidding in auctions? That domain is often cited as one of the big successes of game theory, with many academics having advised the FCC on their design of the high-stakes spectrum auctions [1]. However, it can also be seen as a failure of game theory in the sense that the necessary simplification of the domain has repeatedly caused the deployment of mechanisms that can, in practice, be exploited by the bidders [14]. Similarly, the authors themselves cite the Trading Agent Competition (TAC), as a domain where "it is not reasonable to expect that players contemplate the entire strategy space... equilibria don't play here as great a predictive or prescriptive role." In some sense, this is an acknowledgement by the authors that game theory doesn't answer every question. But then in Section 4.3 they characterize most of multiagent learning results as focusing on self play and games with two agents. And their five agendas for multiagent learning are all characterized in game theoretic terms (except perhaps the fifth). Such a characterization risks marginalizing much of the multiagent learning work referenced above.

Perhaps the authors do intend that research situated in complex domains could fall within their taxonomy. Much of the research in these settings can be characterized in similar terms to those put forth by the authors, such as learning a model of the game or opponent; learning one's own utility, etc. And there are indeed examples of successful abstractions of complex multiagent interactions to game theoretical terms, including in TAC [15]. But in the more complex settings, the issues are bound to differ, at least to the extent that the abstract analysis doesn't tell the whole story.

Before closing, I would just like to address a few more minor points in the article.

- In Section 3, the authors state that "in a multiagent setting one cannot separate learning from teaching." However it is important to remember that teaching assumes learning — on the part of the other agents. Learning, on the other hand, can take place without any such assumptions about the learning (or teaching) abilities of the other agents. For example, Littman and I consider a set of teaching strategies and analyze how they interact with various learning (but not teaching) strategies [5]. In that work we demonstrate that teaching and learning can be synergistic, but that having multiple teachers can lead to problems (consider 2 "bully" agents in the game of Chicken).

- Also in Section 3, the authors state that "there is no a priori reason to expect that machine learning techniques that have proved successful in AI for single-agent settings will also prove relevant in the multi-agent setting." While technically correct, I think this statement is somewhat misleading in that there's also no a priori reason that single agent methods *can't* apply. They may be more or less effective when assumptions, such as domain stationarity, are violated; but effective single-agent approaches may still be useful first-cut solutions in multiagent settings and in some settings may prove effective — for example in combination with teaching agents as suggested above.

- In Section 5, the authors give examples of learning algorithms being used to compute properties of the game. Another example that might be added is that Q-learning computes the best response policy to an opponent's stationary strategy. Note that this observation ties together the authors' discussions of model-based (e.g. best response) and model-free (e.g. Q-learning) approaches in Sections 4.1.1 and 4.1.2.

In summary, multiagent learning is definitely a good tie between game theory and AI: there is much work that falls in the intersection of these two areas, and the article by Shoham, Powers, and Grenager very effectively characterizes both its strengths and current limitations. But from an AI perspective, multiagent learning should be considered more broadly than game theory can address. In this context, how (and perhaps whether) multiagent learning can be usefully applied in complex domains is still a large open question.

Acknowledgements

This work was supported by NSF CAREER award IIS-0237699, ONR YIP award N00014-04-1-0545, and DARPA grant HR0011-04-1-0035.

References

[1] Peter C. Cramton. The FCC spectrum auctions: An early assessment. *Journal of Economics and Management Strategy*, 6(3):431–495, 1997.

[2] Dieter Fox, Wolfram Burgard, Hannes Kruppa, and Sebastian Thrun. A probabilistic approach to collaborative multi-robot localization. *Autonomous Robots*, 2000.

[3] Hiroaki Kitano, editor. *RoboCup-97: Robot Soccer World Cup I.* Springer Verlag, Berlin, 1998.

[4] Michael L. Littman and Justin A. Boyan. A distributed reinforcement learning scheme for network routing. In Joshua Alspector, Rodney Goodman, and Timothy X. Brown, editors, *Proceedings of the 1993 International Workshop on Applications of Neural Networks to Telecommunications*, pages 45–51. Lawrence Erlbaum Associates, Hillsdale NJ, 1993.

[5] Michael L. Littman and Peter Stone. Implicit negotiation in repeated games. In *Proceedings of The Eighth International Workshop on Agent Theories, Architectures, and Languages (ATAL-2001)*, pages 393–404, August 2001.

[6] Itsuki Noda, Adam Jacoff, Ansgar Bredenfeld, and Yasutake Takahashi, editors. *RoboCup-2005: Robot Soccer World Cup IX*. Springer Verlag, Berlin, 2006.

[7] Itsuki Noda, Hitoshi Matsubara, Kazuo Hiraki, and Ian Frank. Soccer server: A tool for research on multiagent systems. *Applied Artificial Intelligence*, 12:233–250, 1998.

[8] Seth Rogers, Claude-Nicholas Fiechter, and Pat Langley. An adaptive interactive agent for route advice. In *Proceedings of the Second Annual Conference on Autonomous Agents*, pages 198–205. ACM Press, May 1999.

[9] J. Schneider, W. Wong, A. Moore, and M. Riedmiller. Distributed value functions. In *International Conference on Machine Learning*, 1999.

[10] Peter Stone. *Layered Learning in Multiagent Systems: A Winning Approach to Robotic Soccer*. MIT Press, 2000.

[11] Peter Stone, Tucker Balch, and Gerhard Kraetzschmar, editors. *RoboCup-2000: Robot Soccer World Cup IV*, volume 2019 of *Lecture Notes in Artificial Intelligence*. Springer Verlag, Berlin, 2001.

[12] Peter Stone, Robert E. Schapire, Michael L. Littman, János A. Csirik, and David McAllester. Decision-theoretic bidding based on learned density models in simultaneous, interacting auctions. *Journal of Artificial Intelligence Research*, 19:209–242, 2003.

[13] Milind Tambe. Tracking dynamic team activity. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Menlo Park, California, 1996. AAAI Press.

[14] Robert J. Weber. Making more from less: Strategic demand reduction in the FCC spectrum auctions. *Journal of Economics and Management Strategy*, 6(3):529–548, 1997.

[15] Michael P. Wellman, Patrick R. Jordan, Christopher Kiekintveld, Jason Miller, and Daniel M. Reeves. Empirical game-theoretic analysis of the TAC market games. In *AAMAS-06 Workshop on Game-Theoretic and Decision-Theoretic Agents*, 2006.