Introduction to the special issue on learning and computational game theory

Amy Greenwald · Michael L. Littman

Game theory is concerned with the decision making of utility-maximizing individuals in their interactions with one another and their environment. From its earliest days of study, researchers have recognized the important relationship between game theory and learning—using experience from past play to guide future decisions. Recently, there has been a surge in research that applies a computational perspective to learning in general-sum games. The editors have been involved with such projects: for example, deriving finite time bounds on learning algorithms known to converge to game-theoretic equilibria in the limit (Greenwald, Li, & Marks, 2006); and developing computationally efficient models of strategic interactions (Kearns, Littman, & Singh, 2001). Our purpose in editing this special issue on computational issues that pertain to learning in games was to provide an opportunity for sharing results in this rapidly growing area at the intersection of computer science and economics.

When facing an unknown opponent, learning plays a central role. As such, there are many ways of applying machine-learning techniques to games. As eloquently argued by Shoham, Powers, and Grenager (2007), it is critical for research in this area to be geared towards solving a precise problem, and that the criteria for judging the success of the work be clearly stated. We selected five articles that followed these important guidelines. Each of the contributions in the special issue spells out what information is available to the individual players (their own actions, the actions of others, utilities of all players, identity of the other players, etc.), how their performance is to be judged (utility of the player, social utility, convergence to equilibrium, stability of learned behavior, etc.), the model of uncertainty (randomized payoffs, noise in action perception, stochastic action effects, etc.), the relevant equilibrium concepts, and any computational concerns.

A. Greenwald (✉)
Department of Computer Science, Brown University, Providence, RI 02912
e-mail: amy@cs.brown.edu

M. L. Littman
Department of Computer Science, Rutgers, The State University of NJ, Piscataway, NJ 08854-8019

We believe the papers that appear here represent high-quality research in the area of machine learning and computational game theory. In this introduction to our special issue, we briefly survey these selected papers to highlight their contributions, breadth, and impact.

In "Slow Emergence of Cooperation for Win-Stay Lose-Shift on Trees," Elchanan Mossel and Sébastien Roch tackle a mainstay of research in game theory and learning, that of the iterated Prisoner's dilemma. With two players and two actions ("cooperate" or "defect") and a dominant strategy equilibrium, the Prisoner's dilemma is perhaps the simplest possible scenario in which to study learning and game theory. In addition, the authors focus on perhaps the simplest possible learning algorithm, *win-stay lose-shift*. A player that adopts this learning algorithm has an action choice—cooperate or defect—which it retains until it "loses" a round by its opponent selecting defect. At this point, the learner shifts its action to the only other one available.

In Mossel and Roch's study, the learners are assigned to nodes of a static network structure and play randomly selected neighbors. Prior work has shown that *win-stay lose-shift* learners will always eventually reach a global "all cooperate" state, at which point learning ceases. Computationally, it is important to understand how long it takes for a set of players to reach this state. It is generally believed that collections of agents and their interactions, say through markets, cannot solve computationally hard problems. As Daskalakis, Goldberg, and Papadimitriou (2006) said, "if your laptop can't find it, then neither can the market." This paper shows that even simple graphs like bounded-degree trees are "hard" for *win-stay lose-shift* learners—they can take exponential time to learn to cooperate.

Whereas the first paper selected a known learning strategy and analyzed its behavior, in "AWESOME: A General Multiagent Learning Algorithm that Converges in Self-Play and Learns a Best Response Against Stationary Opponents," Vincent Conitzer and Tuomas Sandholm take the opposite approach. They start with a desired behavior—a learner that converges to a Nash equilibrium in self-play and to a best response when facing a stationary opponent—and then design an algorithm to achieve this goal in a repeated game setting. Their stated criterion was originally proposed by Bowling and Veloso (2002), who provided an algorithm for the two-player two-action case. This paper provides a solution to games with arbitrary numbers of players and actions.

To achieve its desiderata, the AWESOME learning algorithm is complex, involving estimating opponent strategies, synchronizing behavior, recognizing "intent," and even working to avoid the possibility that its own intent will be misread. It can be viewed as multiple algorithms patched together to handle the different cases spelled out by the original learning goal. In the paper "A general criterion and an algorithmic framework for learning in multi-agent systems," Rob Powers, Yoav Shoham, and Thuc Vu go a step further. They offer three related contributions: (1) a broader and more robust set of criteria for learning, (2) an algorithm that achieves these criteria, and (3) extensive empirical comparisons of sets of learners in a wide variety of repeated games.

Their criteria seek to answer the seemingly simple question, "How can an agent learn to maximize its rewards in an environment containing other agents who may also be learning?" Like the previous paper, they adopt the perspective that the answer to this question should depend on the type of opponent the learner is facing. But, as the authors point out, we cannot hope to design an agent that could learn to maximize its rewards against an arbitrary set of opponents (an agent can only learn when its opponents' past play is correlated with their future play). Instead, the authors propose that we design agents that perform well against opponents in some prespecified target set. In fact, their approach is to selectively deal with three different target sets. In the two-player case, it informally reduces to: maximize against opponents in a pre-set known class, cooperate with others like you, and protect yourself

against the unknown. Like AWESOME, the resulting algorithm has a quilt-like character—different strategies are stitched together into a single algorithm that attempts to identify the opponent's class and then behaves accordingly.

The authors make a compelling case for these criteria and their algorithm and also share a number of insights about the subject matter such as that learning is inherent in repeated games and that multiagent learning is inseparable from teaching. In fact, if learning is about balancing exploiting with exploring, then multiagent learning is about balancing exploiting and exploring with "explaining"—taking actions that convey information to the other players.

In "Online Calibrated Forecasts: Memory Efficiency versus Universality for Learning in Games," Shie Mannor, Jeff S. Shamma, and Gürdal Arslan focus on a critical aspect of learning, that of modeling the environment for use in effective decision making. They provide an approach to learning to make calibrated forecasts—event probability predictions that are accurate in hindsight—and apply it to two-player repeated games. Unlike typical "universal" approaches (Foster & Vohra, 1997), their approach is memory efficient.

In multiagent environments, predicting the behavior of one's opponents facilitates the sensible strategy of acting so as to maximize rewards assuming correct predictions. The authors show that an agent using their prediction algorithm when playing against a slow learner will make predictions that accurately track the behavior of the opponent and, consequently, it will execute a best response to its actual play. Moreover, if two players, both using their prediction approach, face off, their behavior can converge to a Nash equilibrium.

Learning for prediction also plays a central role in the final paper, "Bidding Agents for Online Auctions with Hidden Bids," by Albert Xin Jiang and Kevin Leyton-Brown. Unlike the other papers, which treat variations of the repeated game setting, this paper studies bidding in auctions. In auctions, bidders can often benefit from information about the valuations of their competitors and the likely clearing prices. Learning can be used to acquire this information.

Given the right data set, learning valuations can be posed as a classic machine-learning estimation problem. In the auction setting specifically, however, the technical challenge is: How can we learn bidder valuations given that only limited information is revealed through bidding? In particular, if a bidder's valuation is below the reserve price or above the clearing price, it is not made available to the learning algorithm. Simply ignoring this missing data leads to unacceptable bias and bad predictions—a less extreme example of the dangerous reasoning that led to the assumption the Space Shuttle *Challenger* could launch in cold weather although cold-weather instances were left out of the dataset (Dalal, Fowlkes, & Hoadley, 1989).

The authors show that Expectation-Maximization can be used to help predict the unobserved data values, and they validate their approach using synthetic examples. They also report tests performed with empirical data derived from 60 eBay auctions for Sony Playstation Consoles. The predictions were useful for both game-theoretic and decision-theoretic decision making, demonstrating that, indeed, the application of learning to games can yield practical algorithms.

The final paper, "Learning Payoff Functions in Infinite Games," by Vorobeychik, Wellman, and Singh, also examines learning in the auction setting. It tackles the problem of learning in games of incomplete information with infinite strategy sets. Here, the primary task is not to model the behavior of other players, but rather to learn payoff functions from data. A systematic exploration of strategy profiles might suffice in games with finite strategy sets, but for games with infinite strategy sets, the problem is more difficult. The authors propose applying off-the-shelf supervised-learning techniques (regression, support vector machines, and the like) to enable players to infer the general form of the payoff functions from a given

set of observations. Though the suggested algorithms are familiar, the proposed metric for evaluating the quality of learned payoff functions is novel. Rather than judging the quality of some learned payoff function $\hat{u}$ relative to the true payoff function u by measuring the difference between $\hat{u}$ and u, the authors measure the difference between a Nash equilibrium of the game with payoffs $\hat{u}$ and a Nash equilibrium of the game with payoffs u. They find that different ways of setting up the learning problem lead to differences in efficacy in their benchmark auction tasks.

But, you did not come to the special issue to hear us talk about it. You came to read the excellent papers themselves. We hope you find the work as stimulating as we have.

References

Bowling, M., & Veloso, M. (2002). Multiagent learning using a variable learning rate. *Artificial Intelligence*, *136*, 215–250.

Dalal, S. R., Fowlkes, E. B., & Hoadley, B. (1989). Risk analysis of the space shuttle: Pre-Challenger prediction of failure. *Journal of the American Statistical Association, 84*.

Daskalakis, C., Goldberg, P. W., & Papadimitriou, C. H. (2006). The complexity of computing a nash equilibrium. In *Proceedings of the Thirty-Eighth Annual ACM Symposium on Theory of Computing* (pp. 71–78).

Foster, D., & Vohra, R. (1997). Calibrated learning and correlated equilibrium. *Games and Economic Behavior, 21*, 40–55.

Greenwald, A., Li, Z., & Marks, C. (2006). Bounds for regret-matching algorithms. In *Proceedings of the Ninth International Symposium on Artificial Intelligence and Mathematics*.

Kearns, M., Littman, M. L., & Singh, S. (2001). Graphical models for game theory. In *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence (UAI)* (pp. 253–260).

Shoham, Y., Powers, R., & Grenager, T. (2007). If multi-agent learning is the answer, what is the question? *Artificial Intelligence*. Special issue on the foundations of research in multi-agent learning.