

IJCAI-95 Workshop on Adaptation and Learning in Multiagent Systems

Sandip Sen

■ The goal of the Workshop on Adaptation and Learning in Multiagent Systems was to focus on research that addresses unique requirements for agents learning and adapting to work in the presence of other agents. Recognizing the applicability and limitations of current machine-learning research as applied to multiagent problems and developing new learning and adaptation mechanisms particularly targeted to this class of problems were the primary research issues that we wanted the authors to address. This article outlines the presentations that were made at the workshop and the success of the workshop in meeting the established goals. Issues that need to be better understood are also presented.

Research in multiagent systems has produced techniques that allow multiple agents sharing common resources to coordinate their actions so that individually rational actions do not adversely affect overall system efficiency (Bond and Gasser 1988). Whereas previous research efforts looked at offline design of agent organizations, behavioral rules, negotiation protocols, and so on, it was recognized that agents operating in open, dynamic environments must be able to flexibly adapt to changing demands and opportunities (Lesser 1995). In particular, individual agents are forced to engage with other agents with varying goals, abilities, composition, and lifespan. To effectively use opportunities presented and avoid pitfalls, agents need to learn about other agents and adapt local behavior based on group composition and dynamics. Although standard supervised, unsupervised, and reinforcement learning techniques can be

used as starting points for exploring effective learning techniques in multiagent situations, one needs to augment these techniques to match environmental demands and agent characteristics. For example, multiple agents learning at the same time present unique challenges for learning and adaptation techniques.

Motivation

The goal of this workshop, held during the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI-95), was to focus on research that addresses unique requirements for agents learning and adapting to work in the presence of other agents. Recognizing the applicability and limitations of current machine-learning research as applied to multiagent problems and developing new learning and adaptation mechanisms particularly targeted to this class of problems were the primary research issues that organizers wanted the authors to address. The workshop call for papers particularly welcomed new insights into this class of problems from other related disciplines and emphasized the interdisciplinary nature of the workshop.

Papers of the following topics, among others, were sought and received: (1) the benefits of adaptive-learning agents over agents with fixed behavior in multiagent problems; (2) the exploration of the applicability of case-based, explanation-based, and inductive learning systems in novel multiagent problems; (3) the characterization of learning and adaptation methods in terms of modeling power, communication abilities, knowledge requirement, and processing abilities of

individual agents; (4) the development of learning and adaptation strategies for environments with cooperative agents, selfish agents, and partially cooperative agents (they cooperate only if individual goals are not sacrificed) and environments that contain a mixture of these types of agent; (5) the analysis and construction of algorithms that guarantee the convergence and the stability of group behavior; and (6) the study of adaptive behavior in team games, where one group of cooperative agents is pitted against another group of cooperative agents.

The need for learning in multiagent environments has recently been observed in both the distributed AI and machine-learning communities. As mentioned previously, there is ample motivation to study coordination mechanisms that allow agents to incrementally build models of the environment and of other agents and that enable agents to perform more effectively in a dynamic environment than when using static coordination knowledge. The organizers, contributors, and participants of the workshop felt that the timing of the workshop was appropriate to focus discussion on research issues that will enable multiagent-system researchers to develop useful applications over the next few years.

The workshop was motivated by these concerns, and about 45 attendees of the workshop brought their own unique perspectives to bear on these engaging and critical issues. The workshop schedule consisted of 10 oral and 6 poster presentations. The oral presentation sections consisted, with one exception, of three presentations on a common theme, followed by a panel discussion where core issues were raised, and relations between the different approaches were analyzed. The audience participation in the workshop was exemplary and contributed to the overall success of the workshop, as measured by participant satisfaction.

Workshop Sessions

In the first session, David Carmel and Shaul Markovitch (both of Technion,

Israel) discussed an approach to modeling opponents with an eye to developing optimal interaction strategies. They assume that agents' strategies can be modeled by finite automata models. The paper presented both a heuristic algorithm for inferring agent models from input-output behavior and a method for finding an optimal interaction strategy with the inferred model. This work builds on that of Angluin (1978) on inferring automata models with an oracle that answers membership and equivalence queries. Claudia Goldman (The Hebrew University, Jerusalem) and Rosenschein discussed a more cooperative scenario, where two agents mutually supervise each other to evolve better coordination. In their work, agents interchange labeled samples with or without the intervention of a mediator and try to infer an approximately correct description of the behavior of the other agent from these samples. The probabilistic concept learning schemes used in the paper are motivated by the model presented by Kearns and Shapire (1990). The paper by Thomas Haynes and Sandip Sen (both of University of Tulsa) addressed the issue of evolving coordination strategies for a group of agents using the genetic programming paradigm. They experiment with the well-known *predator-prey domain* (Benda, Jagannathan, and Dodhiawalla 1985), where four predator agents are trying to surround and capture a prey moving in a toroidal grid world. They also address the interesting issue of coevolving both cooperative and antagonistic agents.

The second session started with Ciara Byrne (King's College) and Peter Edwards's paper on refining the knowledge bases of individual group members to improve the effectiveness of the entire group. Their refinement facilitator agent uses KQML messages to coordinate refinements that benefit the group. M. V. Nagendra Prasad, Victor Lesser (both of University of Massachusetts), and Susan Lander's (Blackboard Technology Group) paper dealt with agents learning about their roles in an organization and about the local and joint search spaces in group decision making. They use dif-

ferent supervised learning schemes, including a form of instance-based learning (Aha, Kibler, and Albert 1991), in building a group of agents that learn to effectively design artifacts. They concluded that even though learning by itself does not allow the agents to produce the same solution quality that can be obtained by direct negotiation, it does provide for significant savings in communication cost. Sen and Mahendra Sekaran (University of Tulsa) addressed the dilemma of an agent deciding whether to help another agent in the environment. They showed that agents using a probabilistic reciprocity mechanism can form stable groups that perform at the optimum level. This research shows interesting possibilities for designing agent societies where optimal system performance can be obtained even though individual agents are self-motivated (this characteristic is more representative of open systems than the assumption that all agents are cooperative or benevolent by design).

In a one-of-a-kind paper, Larry Glicoes, Rich Staats (both of LMI), and Michael Huhns (MCC Corp.) described their design of an intelligent-agent-based distribution system for the U.S. Department of Defense to move personnel, equipment, and supplies. A system of static and mobile agents uses historical data and real-time data communicated by satellite to push shipments through to meet deadlines. The agents must learn to adjust their preference for other agents, as well as modes of transportation under different system and environmental conditions, so that efficient transportation of personnel and goods is achieved both for routine operations and unforeseen contingencies.

The last session of oral presentations involved multiagent systems using reinforcement learning techniques. The first paper in the group, by Sen and Sekaran, evaluated the classifier system approach based on genetic algorithms (Holland 1986) and found it to be at least as effective as the more popular Q-learning approach (Watkins and Dayan 1992) on domains with varying agent cou-

pling and feedback delays. In this work, the authors assume that agents learn from environmental feedback only and are not even aware of the presence of other agents. These assumptions, together with the fact that multiple agents are learning concurrently, make it difficult for individual agents to find optimal policies even after repeated interactions. Experiments presented, however, showed that close to optimal performance can be produced under certain assumptions of agent coupling and feedback delays. Tuomas Sandholm (University of Massachusetts at Amherst) and Robert Crites investigated the use of the Q-learning algorithm in the iterated prisoner's dilemma game. The learning agent was able to develop optimal strategies against opponents with static strategies, but when both players were learning concurrently, the learners were less effective. These two papers highlight the problem posed to traditional machine-learning approaches by the nonstationary environments created by concurrent learning by multiple agents. Maja Mataric (Brandeis University) also stressed the inadequacy of associated assumptions made in traditional reinforcement-learning literature when an agent tries to cope with a real world with noisy perception and action and inconsistent reinforcement, particularly in the presence of other agents. She argued for the effective use of existent domain knowledge to design heterogeneous reward functions and goal-specific progress estimators to speed the reinforcement-learning process in situated domains. Her presentation also included a video of groups of robots learning to solve cooperative tasks.

The poster session in the early afternoon was informal but informative and provided sufficient opportunities for attendees to discuss mutual interests and ideas. Pan Gu (North-eastern University) and Anthony Maddox's poster presented a distributed reinforcement-learning model (DRLM) where agents share experience and provide feedback to peers. The DRLM is used in a real-time environment by distributed agents to

process interrelated tasks. Anupam Joshi (Purdue University) presented a scientific computing scenario, with the PYTHIA Project, where agents use both supervised and unsupervised (using epistemic utility theory) learning mechanisms. Two noteworthy aspects of the paper are (1) a multiagent extension to the previously existent single-agent system and (2) the characterization of when agents in the PYTHIA system should or should not use learning mechanisms.

Britta Lenzmann (University of Bielefeld, Germany) and Ipke Wachsmuth presented an application of the VIENA (virtual environments and agents) system where agents learn user preferences for a three-dimensional environment from direct feedback. The overall behavior of the system is determined by the way agents, representing different perspectives of the environment, organize themselves based on feedback from the user. Yishay Mor, Claudia Goldman (both of Hebrew University), and Jeff Rosen-schein's poster analyzed the complexity of learning an opponent's model in game-theoretic negotiations. Even though learning the best response to a static strategy of an opponent, using a finite-automata model, can take exponential time, for a restricted class of simple automata, a polynomial-time learning algorithm was found.

Takuya Ohko (Keio University), Kazuo Hiraki, and Yuichiro Anzai's poster presented the LEMMING learning system that reduces communication cost in the contract-net protocol (Smith 1980) (used for task allocation in multiagent systems). Using case-based reasoning (Kolodner 1993), the LEMMING system can learn to send information selectively to relevant agents, thus reducing waste of communication cost involved in broadcast communication. Andrea Schaerf, Yoav Shoham (Stanford University), and Moshe Tennenholtz's (Israel Institute of Technology) poster investigated a loosely coupled system where agents concurrently adapt to each other and a changing environment. This paper analyzed the effects of adaptive behavior parameters and communication on system efficiency when a group of reinforcement learn-

ing agents try to balance the load in a distributed system.

The workshop concluded on a positive note, with the attendees voicing the need for similar workshops to be held in the future. A significant portion of the attendees expressed a desire to attend the 1996 AAAI Spring Symposium on Adaptation, Coevolution, and Learning in Multiagent Systems to be held at Stanford University on 25 to 27 March 1996.

The schedule of the workshop, as well as abstracts of the presented papers, can be accessed on the web at <http://euler.mcs.utulsa.edu/~sandip/wshop/schedule.html>.

What's Next?

The workshop helped focus on several key issues in multiagent-learning research. The following list presents some of the issues that we need to better understand before significant progress can be made in this nascent area of research:

Individual versus cooperative learning: Agents can either individually try to model others using their experience and perception and with the purpose of personal gain, or they can actively share and participate in constructing a group model and plan of activities that will benefit the entire group. Distinct forms of learning scheme will be suited for each of these two learning modes.

Concurrent versus staggered learning: The number of agents learning and adapting at the same time will influence the rate of convergence of the learning processes used by individual agents.

Agent interactions: Agents can interact frequently or infrequently; their interactions can be regulated and anticipated (as in a fixed organization) or completely unpredictable (as in open systems). The flux in agent groups and the length of the period over which agents interact determine how effectively agents can adapt to others.

Agent relationships: Some agents can have more or less control over group activities or shared resources and can force some situations that will aid in their learning process.

Agent modeling: Assumptions about the behavioral complexity of other agents or limitations in cognitive abilities will constrain the learning abilities or suggest learning schemes for agents.

Environmental feedback: The rate and nature of environmental feedback are key to the kind of learning mechanisms that can be used.

References

- Aha, D. W.; Kibler, D.; and Albert, M. K. 1991. Instance-Based Learning Algorithms. *Machine Learning* 6:37-66.
- Angluin, D. 1978. On the Complexity of Minimum Inference of Regular Sets. *Information and Control* 39:337-350.
- Benda, M.; Jagannathan, V.; and Dodhiawalla, R. 1985. On Optimal Cooperation of Knowledge Sources, Technical Report BCS-G2010-28, Boeing AI Center, Boeing Computer Services, Bellevue, Washington.
- Bond, A. H., and Gasser, L., eds. 1988. *Readings in Distributed Artificial Intelligence*. San Francisco, Calif.: Morgan Kaufmann.
- Holland, J. H. 1986. Escaping Brittleness: The Possibilities of General-Purpose Learning Algorithms Applied to Parallel Rule-Based Systems. In *Machine Learning: An Artificial Intelligence Approach, Volume 2*, eds. R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, 593-623. San Francisco, Calif.: Morgan Kaufmann.
- Kearns, M. J., and Shapire, R. E. 1990. Efficient Distribution-Free Learning of Probabilistic Concepts. In Proceedings of the Thirty-First Annual Symposium on Foundations of Computer Science, 382-391. Washington, D.C.: IEEE Computer Society.
- Kolodner, J. 1993. *Case-Based Reasoning*. San Francisco, Calif.: Morgan Kaufmann.
- Lesser, V., ed. 1995. *Proceedings of the First International Conference on Multi-Agent Systems*. Menlo Park, Calif.: AAAI Press.
- Smith, R. G. 1980. The Contract Net Protocol: High-Level Communication and Control in a Distributed Problem Solver. *IEEE Transactions on Computers* C-29(12): 1104-1113.
- Watkins, C. J. C. H., and Dayan, P. 1992. Q-Learning. *Machine Learning* 8:229-256.

Sandip Sen received his Ph.D. from the Electrical Engineering and Computer Science Department of the University of Michigan at Ann Arbor in 1993. Since then, he has been an assistant professor in the Department of Mathematical and Computer Sciences at the University of Tulsa.