

Multi-agent Learning in Non-cooperative Domains

Mahendra Sekaran and Sandip Sen
Department of Mathematical & Computer Sciences
University of Tulsa
600 South College Avenue, Tulsa OK 74104-3189
mahend@euler.mcs.utulsa.edu

Motivation

Previous work in *coordination* (Bond & Gasser 1988) on multi-agent systems are specific either to cooperative or non-cooperative problem domains. Previous work in *learning* in multi-agent systems have considered agents operating to solve a cooperative task with explicit information sharing and negotiations (Weiß 1993, Tan 1993). In a companion paper (Sen, Sekaran, & Hale 1994) we describe a general purpose system which describes a cooperative domain in which two agents work together on a joint task without explicit sharing of knowledge or information. The focus of this poster is to extend this approach to a non-cooperative domain, where the agents have conflicting goals. The strength of this work lies in the fact that there is no explicit knowledge exchange between the agents and no dependencies on agent relationships.

Problem description

The problem we consider is the "Block pushing" problem in which two autonomous agents with different capabilities and no knowledge of the domain are to push a block from an initial position to individual goal positions in Euclidean space. At each step, each agent applies a force and angle and the block is transferred to the next state by a combination of the actions. A trial ends when the block reaches either one of the goal positions or when it is out of the playing field. Each agent through repeated trials learns an optimal policy to reach (as close as possible) its goal position. The feedback received after each action is an exponentially decreasing function of the current distance of the block from the optimal path to their respective goals. Based on the received feedback the agents update their individual policies using a well-known reinforcement learning scheme, the Q-learning algorithm (Watkins 1989). Convergence of policy matrices (there is no significant update in the individual policies over a set of trials) is used as the stopping criteria for experiments.

Experiments and Results

We found that when the force range for one agent is greater than the other agent, it is able to overpower the

other agent and drag the block to its goal. The number of trials to convergence increases as the force range for the weaker agent is increased (because it offers more resistance). When the two agents possess equal capabilities, neither of them are able to push the block to their goal and they manage to push the block to a position in between the goals. As the strength of the weaker agent increases the final block position moves away from the goal of the stronger agent. This phenomenon is accentuated when the number of discrete options available to the weaker agent is increased. This is because the stronger agent converges to a sub-optimal policy and can be avoided by choosing a probabilistic action scheme instead of a deterministic one (though policy convergence will be much slower).

We show that agents can use reinforcement learning schemes to achieve their goal even when working against a non-benevolent agent without requiring a model of the latter. Currently, we are investigating resource sharing between multiple agents using the same approach.

References

- A. H. Bond and L. Gasser. *Readings in Distributed Artificial Intelligence*. Morgan Kaufmann Publishers, San Mateo, CA, 1988.
- C. Watkins. *Learning from Delayed Rewards*. PhD thesis, King's College, Cambridge University, 1989.
- G. Weiß. Learning to coordinate actions in multi-agent systems. In *Proceedings of the International Joint Conference on Artificial Intelligence*, pages 311-316, August 1993.
- M. Tan. Multi-agent reinforcement learning: Independent vs. cooperative agents. In *Proceedings of the Tenth International conference on Machine Learning*, pages 330-337, June 1993.
- S. Sen, M. Sekaran and J. Hale. Learning to coordinate without sharing information. In *Proceedings of the twelfth national conference on Artificial Intelligence*, August 1994.