Communication in learning multi-agent systems

Multi-agent systems [7] consist of a number of interacting agents, each trying to maximise some reward, or payoff, over time. This covers a variety of models from simple two player games such as the well-known prisoner's dilemma to complex economic systems. In such systems there is commonly interaction and communication between agents. Learning of strategies can be enhanced considerably using information received from other agents [1]. Conversely, an agent can provide misleading information in order to encourage other agents towards strategies beneficial to it.

There are a number of avenues to be explored here. First, a good representation of communication models is necessary; encoding information received as part of the state is clumsy. It is more appropriate to model communications explicitly, breaking the state model into parts concerning the information received. Relational models [5] of state representation are probably useful here.

Deciding what information to communicate is a further problem; when is it beneficial to make certain facts known to an adversary? Models of lying can be incorporated here: the choice of 'facts' under consideration need not only be those known to be true to the agent.

Given that other agents may be either lying or wrong about information, an agent must have incorporate into its model a means of determining how to update its state given communication from another agent (who may be lying about its own identity). Such models can also be used to determine how much trust it puts into the information coming in from its various sensors, although we can assume that these will not strategically lie. Ramchurn et. al [6] provide an overview of trust in multi-agent systems, including models of lying, and cite some open questions such as strategic lying and use of social networks.

I propose to explore the way in which communication evolves in interacting systems. Simple communication channels will be provided to a system of interacting heterogeneous agents. Existing work such as [4], [3] studies different aspects of the way in which simple communicative models can emerge in evolutionary systems. One possible way of doing this is to provide communication channels, and a small alphabet of signals, so that sending a signal down a communication channel is a possible action in a reinforcement learning problem. It will then be possible to examine how the agents learn to make use of these signals.

My model will include predators and prey. In such a model it is in the interests of the predator to communicate that it is not a predator, or that it is not stalking some particular prey, while it is in the interests of the prey to correctly identify predators. There are several precedents for this kind of work, such as [2], although the focus is primarily on co-operation. The novelty in the project will be the potential complexity of communication signals, coupled with no prior information about the signals, and the aim of emerging "lies" as well as co-operative behaviour.

Parallel with the experimental work, there is theoretical work which can be done on extending an agent's trust model based on both sensory input and communications, and on how the agent uses the trust model in its decisions. It should be possible for agents to use reputation mechanisms as well as its own communication channels, and for agents to "interrogate" one another to try and determine if they are lying. The inclusion of social networks here is relevant; one could consider an ecological model in which complete trust is afforded to one's kin [8].

References

[1] Social learning theory. `http://teachnet.edb.utexas.edu/~lynda_abbott/Social.html`.

[2] Cristobal Baray. Evolving cooperation via communication in homogeneous multi-agent systems. 1997.

[3] Patrick Grim and Trina Kokalis. Environmental variability favors the emergence of communication. 2004.

[4] Dimitar Kazakov and Mark Bartlett. Co-operative navigation and the faculty of language. 2004.

[5] E. Morales. Scaling up reinforcement learning with a relational representation. 2003.

[6] Sarvapali D. Ramchurn, Dong Hunyh, and Nicholas R. Jennings. Trust in multi-agent systems. 2004.

[7] Yoav Shoham, Rob Powers, and Trond Grenager. Multi-agent reinforcement learning: a critical survey. 2003.

[8] H. Turner and D. Kazakov. Stochastic simulation of inherited kinship-driven altruism. 2003.