

A Research Statement

for Bob Price

Technologies such as machine-learning, game-theory, and probabilistic modeling and planning are increasingly becoming a critical resource for understanding what is going on around us and controlling it in an intelligent way. While computational techniques have demonstrated their potential in many applications, applying generic techniques to general problems remains both theoretically and computationally challenging. In my own research I have noted that human learning and planning occur in a specific social context and I have drawn from this observation inspiration for accelerating well-known probabilistic learning and planning techniques through exploitation of social constraints. In the process, I have also become interested in how learning, planning and game-theoretic techniques can be used to model and accelerate human learning and planning in a social context.

Contributions

My work in the area of accelerating learning and planning in social environments has resulted in a new a paradigm I call *implicit imitation*. Like other contributions to computational models of imitation, implicit imitation accelerates learning and improves planning through observation of other knowledgeable agents. In my work, I have used the general frameworks of Markov decision processes, Bayesian learning and reinforcement learning to understand various approaches to imitation. The analysis revealed a niche for a new form of imitation based on the idea of implicit transfer of information about how an agent's action choices affect its environment. Implicit imitation is unique in that it can be applied to agents with different incompatible representations of their environments, agents in competitive situations, observations of groups of agents with varying expertise, agents with different goals and agents with partially-relevant behaviours.

The simplest implicit imitation model can be developed under the assumption that the observing agent has the same action capabilities as the agents it is observing. I developed a specific algorithm for this case, showed how it can be coupled to generic reinforcement learning techniques and then demonstrated how implicit imitation can dramatically accelerate learning and how properties of the mentor and problem domain affect transfer of knowledge between agents. In subsequent work, this model was extended through robust statistical tests to situations in which the observing agent has different action capabilities than the agent(s) it is observing. Empirically, the enhanced model was shown to be both necessary and sufficient for extending the gains from implicit imitation to situations in which agents have differing action capabilities. A revised version of the paradigm resulted in a new model called *Bayesian imitation* which uses an explicit Bayesian pooling mechanism to make maximum use of observations. Empirically the Bayesian technique has been shown to outperform its antecedents. We believe that the Bayesian technique will also permit agents to learn to avoid negative examples. As described below, the Bayesian technique also provides a useful foundation for more advanced work in partially-observable domains.

In the course of working with Markov-decision processes and reinforcement learning, I have also had the opportunity to contribute to extensions to problem representations in these areas. Traditionally, probabilistic techniques such as Markov-decision processes and reinforcement learning maintain distributions and expected utilities over propositional descriptions of the world. In co-operation with two colleagues I worked on extending these models to use probability distributions and expected utilities over situation calculus axioms written in first-order logical sentences. The representation not only permitted the consequences of actions to be succinctly stated, but also allowed computations to be performed directly on these representations.

Research Plans

My present research into the implicit imitation paradigm has raised a number of new and interesting questions for researchers investigating rational agents learning in a social context. The potency of implicit imitation, in itself, raises an important question: if observations of other agents can dramatically improve an agent's performance, how can competitive agents reason about the value of the information that their action choices communicate to other agents? In the course of my research I have also asked the question about whether imitation can be used to learn protocols for dealing with multi-agent co-ordination issues and to what extent imitation should vary when imitating solutions to single agent tasks versus imitating solutions to tasks that involve group interaction. The framework I recently derived for partially-observable imitation raises the question of how an agent should reason about the value of actions taken to improve the visibility of the agent it is observing. There are a number of outstanding fundamental problems in the field of imitation. Of particular interest is the problem of finding a mapping between perceptions of other agents in the environment and an agent's own representation of its state in its environment. In addition to questions specific to imitation, there are a number of generic problems to be worked on. Extending implicit imitation to continuous domains requires the ability to reason about the confidence or certainty an agent has about various sources of data derived from continuous approximations of sampled data.

In a broader context, I am also interested in how computational mechanisms can be used to support and accelerate human learning and interaction. Over the past year I have been looking at how Bayesian models and document retrieval techniques might be used to improve human learning using models that explicitly take time and cognitive capacity (such as ACTR) into account. Recently, I have spent some time examining how game theoretic and social choice models can be used to understand and support choices in groups of agents that are human or working for humans. When human values are relevant, utility functions of individual agents may become relative and interdependent creating difficulties for traditional techniques.