Building Collaborative Strategies via Imitation

Saleha Raza
Artificial Intelligence Lab, Institute of Business Administration, Pakistan
saleha.raza@khi.iba.edu.pk

Abstract

This research proposes the use of imitation based learning to build collaborative strategies for a team of agents. Imitation based learning involves learning from an expert by observing her demonstrating a task and then replicating it. This mechanism makes it extremely easy for a knowledge engineer to transfer knowledge to a software agent via human demonstrations. This research aims to apply imitation to learn not only the strategy of an individual agent but also the collaborative strategy of a team of agents to achieve a common goal. The effectiveness of the proposed methodology is being assessed in the domain of RoboCup Soccer Simulation 3D which is a promising platform to address many of the complex real world problems and offers a truly dynamic, stochastic, and partially observable environment.

Motivation

Collaboration and strategy building have always been a crucial part of goal-oriented multi-agent systems. The development of collaborative strategies in complex, dynamic systems, however, is a non-trivial task as it demands domain experts to thoroughly asses the problem and provide sufficient knowledge to autonomous agents to enable them to make coordinated decisions. Rule sets or state machines have conventionally be used to devise strategies but their formalization, in a comprehensive manner, is itself a great challenge. In such situations, what is required is an effective mechanism of knowledge transfer from human experts to software agents that facilitates the process of building strategies for an individual or a team of agents. Imitation based learning seems to be a promising choice in this case as this is far more natural for a human to demonstrate an action than to provide guidelines to perform that action. In the presence of multiple agents, imitation enables agents to learn how to collaborate with each other by looking at the way multiple human experts are collaborating with each other during demonstrations. This study proposes the use of imitation based learning to formulate collaborative strategies and demonstrates its effectiveness in the domain of RoboCup Soccer Simulation 3D. We have already established a RoboCup Soccer Simulation 3D team that participated in World RoboCup 2011.

Related Work

Various studies have been conducted, primarily in the field of robotics, which propose the idea of learning via demonstration to train robots. Some supervised learning based imitation include controlling a car within a simulated driving domain using Gaussian Mixture Models (GMMs) (Chernova and Veloso 2007), flying a simulated airplane using decision trees (Sammut et al., 1992) and learning obstacle avoidance and navigation behaviors using kNN classifiers (Saunders et al., 2006). Argal et al. (Argall et al. 2009) presents a survey of different imitation learning approaches. RoboCup Soccer has been a popular test bed for strategy building and optimization. Aler et al. (Aler et al. 2009) have applied imitation based learning and evolutionary algorithms to build and optimize RoboCup Soccer strategies while Wai A.C.L (Wai, A.C.L. 2011) proposed the use of Inductive Logic Programs to train RoboCup Soccer agents via imitation.

Most of the imitation based learning approaches proposed in the literature are focused on the problem of training a specific task to an individual agent and do not explicitly address coordination and teamwork among multiple agents. This research addresses coordination and aims to learn collaborative strategies from human experts via imitation.

Proposed Approach

The proposed approach of imitation learning, as shown in Figure 1, is divided into two phases: the demonstration phase and the learning phase.

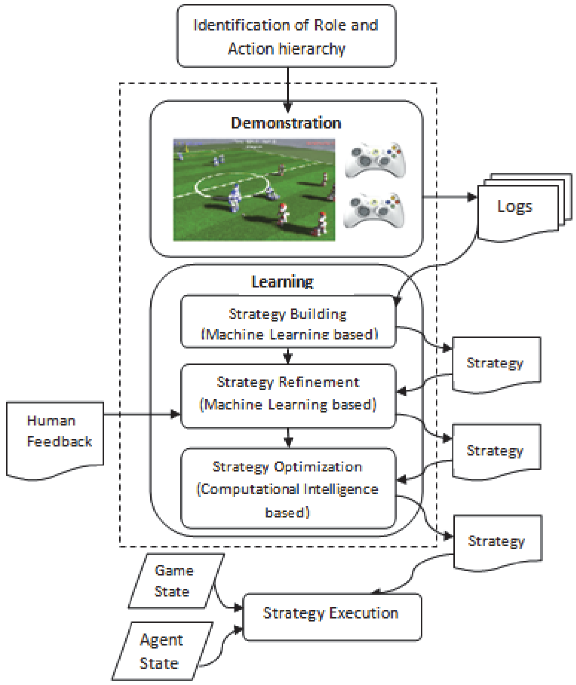

Figure 1 Imitation based learning in RoboCup Soccer

The demonstration phase allows human expert(s) to teach soccer agents by controlling them via game controllers. In the presence of multiple agents, the demonstration is performed by multiple human experts, each expert controlling a different agent. The state-action data gathered during this phase is then fed to the learning phase. The learning phase is divided into three sub phases:

Phase 1 - Strategy Building via demonstrations: This step involves combining and processing all of the data gathered by multiple experts during demonstrations and uses it to learn collaborative strategy. Different classification approaches are applied in a manner that facilitate coordination and knowledge sharing and allows agents to learn from the knowledge/experiences of each other. The demonstration made to one agent will also be available to the other agents during learning phase and all of them will learn and take decision keeping in mind what their fellow agents are doing at the moment.

Phase 2 - Strategy Refinement via human feedback: Once a basic strategy has been learned, the process of refinement involves executing that strategy and providing instantaneous human feedbacks for further refinements. The human can intervene whenever he finds the agent doing something wrong or the agent itself can request assistance when it encounters a new situation and feels indecisive about it.

Phase 3 - Strategy Optimization: The process of optimization deals with the application of the techniques of reinforcement learning and computational intelligence to further optimize the strategy learned via imitation. The real challenge in this step is to come up with a policy that performs better than the one demonstrated by the expert

Experiment

So far, we have experimented with some preliminary scenarios which are based on a single agent learning different tasks via imitation. These tasks include:

- A single player scoring goals in an empty soccer field
- A single player avoiding collision with obstacles

We gathered data by demonstration and built classification model on this data to imitate human behavior. The results have been very encouraging as our imitated agent, on average, scored four goals per match in task 1. For task 2, the collision avoidance agent successfully avoided collisions 90% of the time. Our next target is to conduct experiments in multi-player scenarios and exploit imitation to learn collaboration.

Acknowledgement

The author is greatly thankful to Prof. Mary-Anne Williams and Dr. Sajjad Haider for their invaluable feedback and guidance.

References

Aler, Ricardo, Jose M. Valls, David Camacho, and Albert Lopez. 2009. “Programming Robosoccer Agents by Modeling Human Behavior.” Expert Syst. Appl. 1850 1859. doi:10.1016/j.eswa.2007.12.033.

Argall Brenna D., Sonia Chernova, Manuela Veloso, and Brett Browning. 2009. “A Survey of Robot Learning from Demonstration.” Robotics and Autonomous Systems. 469 483. doi:10.1016/j.robot.2008.10.024.

Chernova, Sonia, and Manuela Veloso. 2007. “Confidence based Policy Learning from Demonstration Using Gaussian Mixture Models.” In Joint Conference on Autonomous Agents and Multi agent Systems.

C. Sammut, S. Hurst, D. Kedzier, and D. Michie. Learning to fly. 1992. In Proceedings of the Ninth International Workshop on Machine Learning.

A. Alissandrakis, C. L. Nehaniv, K. Dautenhahn, and J. Saunders. 2006. Evaluation of robot imitation attempts: comparison of the system's and the human's perspectives. In Proceedings of the 1st ACM/IEEE Interna tional Conference on Human Robot Interactions (HRI'06).

Wai, A.C.L. 2011. Learning by Imitation using Inductive Logic Program. MS Thesis. Department of Systems and Computer Engineering, Carleton University, Ontario, Canada.