Multi-Agent Machine Learning for Mobile Robots:
A Reinforcement Learning Approach

Howard M. Schwartz
Department of Systems and Computer Engineering
Carleton University
Canada

Mobile robots move around their environment executing various tasks. The robots perform actions and these actions cause the robot to have an effect on its. If the environment is not known in advance, then the robots will have to learn what actions to take and the robots may also have to adapt to a changing environment. The relationship between learning and action makes the reinforcement learning approach an excellent method for mobile robot learning.

The key aspect of learning for mobile robots is that they need to learn actions that affect the external environment which includes other cooperating and opposing robots. Imagine small children playing a game of tag. Some are running from base to base while others chase them. How did they all learn to play this game? What strategies are they playing?

The applications of these games are widespread in areas such as autonomous vehicles, border security, infrastructure security and defence applications. Two games in the field of game theory apply to these applications: the pursuer evader game and the defending a territory game. This talk will focus on our recent research investigating how the mobile robots can learn how to play these games. We propose new reinforcement learning methods and algorithms. The pursuer evader game and the defending a territory game are used as the learning platforms. We have proposed the use of adaptive fuzzy inference systems as the primary learning engine and the use of the fuzzy actor critic learning (FACL) method to adapt the parameters. Our work has evolved from studying the single pursuer evader problem and the single defender invader problem to the more complex and challenging multi pursuer chasing a high speed evader game and similarly a multi defender against a high speed intruder game. Both simulation and experimental results are shown.

We have modified the originally proposed FACL algorithm in a number of key ways such that the learning algorithms are more robust, relatively fast in terms of learning the policy and more tractable. We present a new method of conceptualizing the FACL algorithm. We deconstruct the original method and reconstruct it as an object oriented system of learning objects. Each object represents a particular set of learning rules that fire simultaneously and are defined as the rule set. Each rule set represents a node of a fuzzy neural structure. The overall system is structured as an object oriented neural fuzzy actor critic system. We define rule firing sets as the individual low level neural learners. Each rule firing set has its own set of adaptive parameters. Furthermore, statistical analysis of the temporal difference allows the robot to determine if convergence of the adaptive parameters has been achieved and to identify when the environment has changed and learning should be reinitialized. Simulation results show the accuracy and robustness of the learned policies and the sensitivity to adapt environmental changes.