International Journal of Innovative Computing, Information and Control
Volume **6**, Number **5**, May **2010**

 ISSN 1349-4198
pp. 2113–2126

EFFECTS OF COMMUNICATION IN COOPERATIVE Q-LEARNING

Paul Darbyshire[1] and Dianhui Wang[2]

[1]School of Management and Information Systems
Victoria University
Melbourne, Victoria 8001, Australia
Paul.Darbyshire@vu.edu.au

[2]Department of Computer Science and Computer Engineering
La Trobe University
Melbourne, Victoria 3086, Australia
Dianhui.Wang@latrobe.edu.au

Received November 2008; revised June 2009

Abstract. *Reinforcement learning has been utilized to investigate the distributed learning problem in many different multi-agent and team-based scenarios. Invariably, using a distributed learning approach by allowing agents to exchange what they have learned, by either direct communication or by episodic exchanges, the team does better in terms of achieving its goals. The cause-effect relationship between communication and improved performance has been studied in a number of different team based scenarios; however, it is difficult to find a detailed analysis of the effect the communication has as a function of time on the elements of the simulation. Of particular interest is the effect on the learning of agents when part of a communicating multi-agent system. Studies show the effects of cooperative learning in the form of accelerated group learning rates; but it is difficult to pinpoint through the literature exactly how this observed effect is realized within individual agents over time. This paper investigates the effect of communication as a function of time on a multi-agent simulation based on a military distillation, utilizing a modified Q-learning algorithm.*

Keywords: Agent communication, Reinforcement learning, Multi-agent systems, Complex system, Learning rates, Knowledge landscape

1. **Introduction.** In the last decade, agent-based simulations have been increasing in popularity as a research tool for studying a variety of problems from social and economic issues [1] to elements of complexity theory [2]. The application of learning techniques applied to multi-agent systems is a growing field [3], and an exciting, emerging area of research is Multi-agent Reinforcement Learning. Multi-agent reinforcement learning is concerned with how an agent can learn to act optimally in an unknown environment through trial and error interaction, and in the presence of other adaptable agents [4]. A multi-agent system, while generally Markov-like in nature, is essentially a non-Markov decision process, as the interaction between the agents themselves and concurrent learning affect the state transition probabilities [5]. Convergence is one of the measurements of performance often used in many applications utilizing traditional reinforcement learning techniques [6]. While learning is not guaranteed to converge in a non-Markovian environment, the main problem stems from the non-stationary nature of the task due to hidden variables caused by other agents affecting the environment. However, reinforcement learning, and in particular temporal difference (TD) algorithms such as Q-learning, can perform well in multi-agent systems [7].