Chapter XII
Three Perspectives on Multi-Agent Reinforcement Learning

Yang Gao
Nanjing University, China

Hao Wang
Nanjing University, China

Ruili Wang
Massey University, New Zealand

ABSTRACT

This chapter concludes three perspectives on multi-agent reinforcement learning (MARL): (1) cooperative MARL, which performs mutual interaction between cooperative agents; (2) equilibrium-based MARL, which focuses on equilibrium solutions among gaming agents; and (3) best-response MARL, which suggests a no-regret policy against other competitive agents. Then the authors present a general framework of MARL, which combines all the three perspectives in order to assist readers in understanding the intricate relationships between different perspectives. Furthermore, a negotiation-based MARL algorithm based on meta-equilibrium is presented, which can interact with cooperative agents, games with gaming agents, and provides the best response to other competitive agents.

INTRODUCTION

Reinforcement learning (RL) dates back to the early days of cybernetics and work in statistics, psychology, neuroscience, and computer science. Since the late 1980s, it has attracted increasing interests of researchers in the fields of machine learning (ML) and artificial intelligence (AI). Its promise is to find “a way of programming agents by rewards and punishment without needing to
specify how the task is to be achieved” (Kaelbling & Littman, 1996, p. 237).

Multi-agent reinforcement learning (MARL) can be considered an extension of RL to the multi-agent domain. Also, it is a new paradigm of distributed artificial intelligence applications. This domain has many research issues. Many researchers have been chasing their own interests since the early 1990s, and they can be divided into two major research groups. One group studies MARL in the view of machine learning, while the other studies MARL in the view of multi-agent systems.

In our opinion, MARL has three main perspectives. The first perspective is cooperative MARL. It studies how to speed a learning process via mutual interaction. Consequently, there have been many discussions of interaction methods. The second perspective is equilibrium-based MARL. It has a mathematical foundation of game theory and takes some equilibrium solutions as the optimal policy. From this perspective, developing an algorithm to solve the game, namely to achieve the specified equilibrium, is of the most importance. The third perspective, best-response MARL, is to achieve a no-regret policy against other competitive agents.

In practice, as one may see from the descriptions above, these three perspectives of MARL have different features and different applicable domains. We must carefully distinguish between these three perspectives to apply MARL correctly. However, this task can sometimes be difficult due to the complexity of multi-agent systems. Therefore, we intend to develop a general method by combining these three main perspectives of MARL to deal with almost all the learning problems in multi-agent domain. We achieve our goal by examining the concept of metagame and meta-equilibrium. We then propose a general framework of MARL and a negotiation-based MARL algorithm later in this chapter.

The rest of the chapter is organized as follows. In the second section, we discuss three perspectives of MARL technology and their state of arts. We compare these three perspectives of MARL and give the pseudo code for each of them. In the third section, we present a general framework of MARL and discuss negotiation-based MARL in detail. In the fourth section, we report some experiments conducted on the pursuit/prey grid world, and we investigate the performance of negotiation-based MARL. Finally, in the last section, we draw some conclusions and outline future work.

THREE PERSPECTIVES OF MARL

Cooperative MARL

One object of cooperative MARL is to solve the learning problem more effectively. Mutual interaction is the most important method in cooperative MARL. Tan (1993) gave three methods of exchanging information: exchanging states of environment, exchanging learning episodes (i.e., state, action, reward triplets), and exchanging policies or parameters. Nunes and Oliveira (2003) added the fourth method--exchanging advices. Later in 2005, this advice-exchange method was further used to cooperate between RL agents and agents using other learning methods such as evolutionary algorithm (Nunes & Oliveira, 2005).

The state-exchange method has been widely used in partially observable domains so that agents may have more complete information about the environment to improve their learning performance. Compared with state-exchange, episode-exchange often leads to a dilemma of how much information should be exchanged. Excessive exchange of information will result in poor exploration in search space. On the contrary, deficient exchange cannot help agents speed their learning. Another shortcoming of episode-exchange is that it needs high communication cost. Many researchers have indicated that the policy-exchange method is adequate for cooperative MARL. However, the best
Related Content

Enhancement of Conversational Agents By Means of Multimodal Interaction
www.igi-global.com/chapter/enhancement-conversational-agents-means-multimodal/54640?camid=4v1a

A Companionable Agent
www.igi-global.com/chapter/companionable-agent/54643?camid=4v1a

Conversational Agents in Language and Culture Training
www.igi-global.com/chapter/conversational-agents-language-culture-training/54646?camid=4v1a

Simulating UAV Surveillance for Analyzing Impact of Commitments in Multi-Agent Systems
www.igi-global.com/article/simulating-uav-surveillance-analyzing-impact/66066?camid=4v1a