2009 International Conference on Computational Intelligence and Security

Reinforcement Learning for Soccer Multi-agents System

Fahimeh Farahnakian
School of Computer Engineering
Iran University of Science &Technology
Tehran, Iran
Farahnakian@comp.iust.ac.ir

Nasser Mozayani
School of Computer Engineering
Iran University of Science &Technology
Tehran, Iran
mozayani@iust.ac.ir

***Abstract*— Recently the reinforcement learning method is actively used in multi-agent systems. Because of this method played a significant role by handling the inherent complexity of such systems.**

Robotic soccer is a multi-agent system in which agents play in real-time, dynamic, complex and unknown environment. Since the main purpose of a soccer game is to score goals, it is important for a robotic soccer agent to have a clear policy about whether it should attempt to score in a given situation.

Therefore we use reinforcement learning for optimizing policy. In the proposed method, the state spaces include two important parameters for shooting toward the goal; the distance between the ball and the goalkeeper and the probability which is obtained from the research of the UvA team. Of course, we select these parameters for effective features of scoring. Because they are more effective learning algorithm in real-time simulated soccer agent [1].

Experimental results have shown that policy achieved from reinforcement learning lead to more effective shoots toward the goal in simulated soccer agent.

Keywords-Reinforcement learning; Multi-agent system; Q-learning; Robocup

I. INTRODUCTION

Reinforcement learning [2] is known as a good machine learning method to construct optimal policies in unknown environments. Also, RoboCup is a standard problem so that various theories, algorithms and architectures can be evaluated [3]. Behavior learning for complex tasks is also an important research area in RoboCup.

In our previous describes the use of decision tree to kick and catch the ball for two simulated soccer agents [4].

In this paper, we propose a reinforcement learning method called Q-learning [5]. We apply the proposed method for learning the kicking skill of shooter player acting in the RoboCup 2D simulator. Many parameters affect the result of shooting toward the goal. Simulation teams such as UvA Trilearn [6], TsinghuAeolus [7], etc. have good techniques for scoring behavior.

In the proposed method, we focus two effective parameters:

-The distance between the ball and the goalkeeper

- The best point of the goal and the probability of scoring at this point are calculated by UvA Trilearn simulation team [8].

In section 2, we introduce the used reinforcement learning algorithm. Section3 presents Implementation and practical results and finally, section 4 is the conclusion.

II. Q-LEARNING

The standard Q-learning algorithm has several stages [9]:

1. for each s and a, initialize the table entry Q(s, a) to zero

2. Observe the current state (s)

3. Do forever:

- Select an action (a) through one of these methods and execute it:

- *Exploration or random*
- *Exploitation or based on Q-table*
- *Receive reward (r)*
- *Observe the new state (s')*
- *Update the table entry for Q(s, a) as follow:*

$$Q(s,a) = r + \gamma \max_{a'} Q(s',a'))$$

- $s \leftarrow s'$

There are two methods for selecting one action from the possible actions in every state [7 and 2]:

Exploration: or random action selection. So, optimal actions, which are not chosen yet, are added to the table.

Exploitation: action selection is according to the learned Q-table.

It is clear that action selection is more exploration at the beginning of learning, and is more exploitation towards the end of learning.

III. SIMULATION RESULT

In order to apply the Q-learning for scoring problem, we consider two steps. One step is train step that shooter player interactions with the world and Q-table is gathered. The other step is test step, shooter player that is using the

978-0-7695-3931-7/09 $26.00 © 2009 IEEE
DOI 10.1109/CIS.2009.275