Simulation of a Soccer Robot

Team Description Proposal for Brazil Open 2008

Amir Khermandar[1], Mohammad Mazinany[1], Ehsan Takbaz Tavakoli[1], Samira Chavoshi[1], Bahareh Jafary[1], Mehdi Torshani[1]

Robotic Research Laboratory, Iran

E-mail: *Amirkhermandar@gmail.com*

***Abstract*. This paper focus of the Iran' effort in the Brazil Open soccer simulation 2D domain is to develop and to apply machine learning techniques in complex domains. To utilize fuzzy algorithm, employing the developed idea in [3] and modifying it based on human behaviors in real world, the fuzzy rules are derived and simulated via MATLAB software.**

I. Introduction

The Iran' project was established in 2005, starting off with a 2D team.

The team description paper at hand focuses on the Iran 2D, our team competing in soccer simulation's 2D league. The encouraging research goal of the Iran' has always been to exploit AI and machine learning techniques wherever possible. Particularly, the successful employment of reinforcement learning (RL) methods for diverse elements of the Iran' decision making modules has been and is our main focus as shall be detailed in the subsequent sections.

II. Implementation

In this section the algorithm that has been implemented and simulated is described in more details.

Our designed fuzzy control scheme in MATLAB Where we have used over *100* rules, obtained based on [4] and modified for our agent, in fuzzy block to control the agents.

Because not completed yet, our simulation results are not shown here and we are going to do this in our future works. After simulating the controlled behavior of agents via MATAB, in our future works, we are going to employ the proposed scheme for our agents. Figure 1 shows first and second order in agents working.

During the past four years, our team could greatly benefit from facing a nearly stable simulation environment. This allowed us to concurrently redesign vast parts of our team play and enhance several of the machine learning approaches we employ in such a manner that the resulting behaviors are highly competitive.

Currently, a considerable number of changes is being introduced to the simulation environment (Soccer Server Ver.12.1.3). As a consequence, our main focus in 2008 is to adapt our team and coach to the changes introduced.

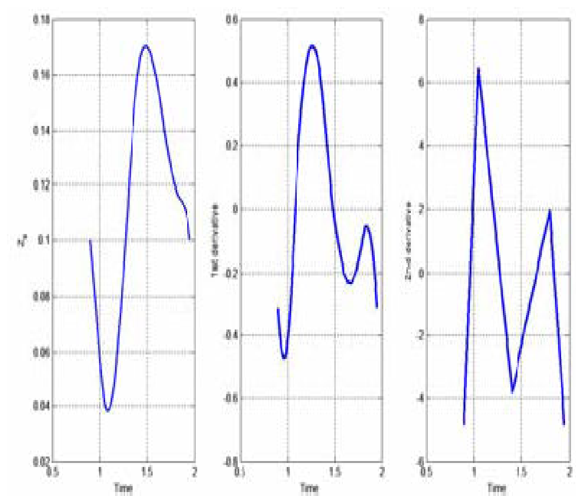

Figure1. Agents working

III. The Learning Algorithm

Learning to hassle, we update the value function's estimates according to the temporal difference learning TD(1) update rule [2], where the new estimate for $V(sk)$ is calculated as $V(sk) := (1 - \alpha) \cdot V(sk) + \alpha \cdot ret(sk)$ with ret $(sk) = \sum_j^N = k^{r(SK,\ \pi(SK))}$ indicating the summed rewards following state sk and α as a learning rate. Each time step incurs small negative rewards, a success goal state a large positive one and the final state of a failure episode a large negative one.

To approximate the value function [1], we employ multi-layer perceptron neural networks with one hidden layer. We perform neural network training in batch mode: Repeatedly a number of training episodes is simulated and in so doing a set of representative states $\acute{S} \subset S$ is incrementally built up where for each $s \in \acute{S}$ we have an estimated value $V(s)$ calculated as mentioned above.

IV. Balancing Deliberativity and Reactivity

There are two fundamental approaches to realize autonomous agents: the reactive and the deliberative approach [5]. A purely reactive agent receives some input via its external sensors, processes it and produces an output. On the contrary, an entirely deliberative agent has its own internal view of its environment and is thus able to build and follow its own plans, aiming to reach some specific goal state. In practice, it is, however, desirable to have a hybrid agent that represents a mixture of the former ones [6], being able to follow its own plans, but sometimes directly reacts to external events without deliberation.

The Iran' agent can be characterized as some hybrid agent with an accentuation on its reactive side. Although following that paradigm is suitable for the soccer simulation environment, it also has its drawbacks: Under certain circumstances, it can be advantageous to reduce reactivity, thus abandoning a short-term improvement of the current situation (e.g. getting into ball possession during the next simulation cycle), and to aim at a long-term profit instead.

A typical example in robotic soccer simulation is the problem of overcoming the opponents' offside line: For the attackers, who are assumed to be in ball possession, it is difficult to get ahead, if the defenders position themselves in a line and in so doing employ the offside rule in their interest. A reactive approach would most probably tend to repeated safe passes among teammates until, eventually, a "gap" in the opponents' defending line appears. This strategy, however, is deemed to last for a rather long time, if the opponent defense is competitive. Consequently, a goal-oriented behavior, which accepts losing the ball in case of failure but most likely improves our team's current situation significantly in case of success, seems promising.

V. Summery

In this team description paper we have outlined the characteristics of the Iran team participating in Brazil Open 2008 2D Soccer Simulation League. We have stressed that our main research focus lies on the development of reinforcement learning techniques and their integration into our team.

Currently, we are particularly interested in extending RL approaches to multi-agent scenarios (team play) as well as in advancing the quality and accuracy of learned behaviors to be equal or superior to hand-coded ones.

Acknowledgment

This project is being supported by the Robotic Research Laboratory of Iran' (Karaj).
The authors would like to thank Mrs. Hematpour, Mr.Mohseni Nezhad and Mr. Sarmadi for their suggestions and supports during all the phases of this project.

References

[1] M. Riedmiller, T. Gabel, F. Trost, T. Schwegmann, Brainstormers 2D – Team Description 2008.
[2] Sutton, R.S.: Learning to Predict by the Methods of Temporal Differences. Machine Learning 3 (1988) 9–44
[3] S. Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa. Resolved momentum control: Humanoid motion planning based on the linear and angular momentum. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, pages 1644–1650, Las Vegas, USA,October 2003
[4] Dynamic Fuzzy Q-Learning Control of Uncertain Systems with Applications to Humanoids "IEEE Conference on Control Applications Toronto, Canada, August 28-31, 2005"
[5] M. Riedmiller, T. Gabel, Brainstormers 2D – Team Description 2007
[6] Weiss, G.: Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence. Massachusetts Institute of Technology (1999)