UFSJ2D (UaiSoccer2D + RoboCap): Team Description Paper Robocup 2014

Andre Luiz C. Ottoni, Heitor Magno R. Junior, Itallo G. Machado, Lara T. Cordeiro, Erivelton G. Nepomuceno, Eduardo B. Pereira, Rone I. da Silva, Marcos S. de Oliveira, Luiz O. R. Vasconcelos, Andre M. Lamounier, Fellipe Lobo, Felipe M. Nomiya, Francisco A. R. Neto, and Joao G. Rocha

Federal University of Sao Joao del-Rei, MG, Brazil
ufsj2d@gmail.com
http://www.ufsj.edu.br

Abstract. This article presents the UFSJ2D Team, a team of simulated robots soccer from UFSJ - Federal University of Sao Joao del-Rei, MG, Brazil. The main goals in this paper are to apply reinforcement learning in the optimization of decision taking and to model a strategy using fuzzy logic.

Key words: reinforcement learning, Q-learning, fuzzy logic, UFSJ2D.

1 Introduction

The UFSJ2D Team is a joining project of UaiSoccer2D and RoboCap Teams, both from Federal University of Sao Joao del-Rei (UFSJ), Brazil.

The UaiSoccer2D Team is part of UAIrobots Group, which operates in lots of researching, extension and teaching lines of robotics at UFSJ. The UaiSoccer2D Team has been participating in the 2D simulation category of robotics competitions since 2011. The first participation of the Team on RoboCup was in Mexico, 2012[1]. That same year, UaiSoccer2D placed fourth in Latin American Robotics Competition (LARC 2012) and placed second in Brazilian Robotics Competition (CBR 2013) in 2013[2].

The RoboCap Team was created in 2008, when a group of students of Mechatronics Engineering got together to participate of several robotics competitions. The 2D simulation team is more recent and was created in 2012. The first participation of RoboCap in Brazilian Robotics Competition was in 2013.

The UFSJ2D Team uses Helios Base (Agent2D 3.1.1) as base code [1]; [2]. Besides that, the formation of Marlik Team [20] was edited and adapted to UFSJ2D.

Some publications had positive results in Reinforcement Learning and Fuzzy Logic at simulation platform of RoboCup [5]; [16]; [12]; [13]; [9]; [11]; [4]; [3].

[1] RoboCup 2012: http://www.robocup2012.org.
[2] CBR2013: http://www.cbrobotica.org.

This paper is organized in sections: at section 2 and 3, Reinforcement Learning and Fuzzy Logic strategies are shown, respectively, and the conclusions are in section 4.

2 Reinforcement Learning

The RL Reinforcement Learning (RL) has been frequently cited in several groups of simulation in Robots's Soccer [7]; [19]. Some works use the Q-learning algorithm in specifics cases: when only the agent has the ball [9]; [11]. In [5] was used technique in order to accelerate the learning.

2.1 Q-learning

The Q-learning algorithm allow to establish a politic of actions interactively [21]; [7]; [11]. The main focus of Q-Learning is which the algorithm of learning learns a function optimal about all space of couple state-action (SxA). Provided that the split of state space and of actions space allow not entering new information. When the optimal function Q is learned by agent, he will know which action will give the greatest reward in a specific case. The function Q(s,a) of reward expected is learning through of errors and trial given by equation following:

$$Q_{t+1} = Q_t(s_t, a_t) + \alpha[r_t + \gamma V_t(s_{t+1}) - Q_{t+1}(s_t, a_t)] \tag{1}$$

where α is called learn rate, r is reward rate, γ is discount factor and $V_t(s_{t+1}) = max_a Q(s_{t+1}, a_t)$ is the utility of state "s" resulting from the action "a", it was got using the function Q learned at moment [10].

2.2 Modelling RL Strategy

Defining Actions: The actions above are for the agent with ball possession:

1. Action: Dribbling A (Carry the ball to the goal with dribbling A);
2. Action: Dribbling B (Carry the ball to the goal with dribbling B);
3. Action: Passing A (Pass the ball to some player from his team with pass A);
4. Action: Passing B (Pass the ball to some player from his team with pass B);
5. Action: Through Passing;
6. Action: Shooting (Shoot the ball toward the goal).

Defining Environment States: To feature the environment performance of the agents, the soccer field was divided into five zones. Each zone has three cells in a total of fifteen cells in the environment. The X and Y coordinates are used to define the sections. This structure is shown in Figure 1.

Another information that is considered in defining environment states of the agent with the ball is the distance from the closest opponent (*dist*). In this case, to *dist* less than four, the opponent is *close*. Otherwise, the opponent is *distant*. This distance was adopted considering the sum of diameters of two robots.

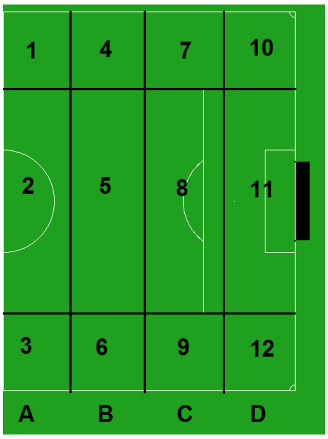

Fig. 1. Proposed division of the soccer field in zones and cells. This structure is valid when the team is attacking from the left to the right.

Defining the Reinforcement Matrix: The environment of simulated soccer robots involves a great complexity to the team reach the primary reward scoring a goal. A common method originally used in animal training, is called reinforcement modeling, which provides additional rewards for "progress" [14]. Thereby, the goal of scoring can be divided into "get the ball possession", "dribbling toward the goal" and "shoot towards goal". Intermediate reinforcements are important to accelerate the learning, however, these reinforcements must receive lower values when the robot does not reaches the target [15].

The Table 1 presents the penalties and reinforcements defined for each field zone. The increasing of the value of reinforcement as the agent reaches the attacking field must be noted.

Table 1. Values of Reinforcements and Penalties for each Zone.

Zone	**Penalty**	**Reinforcement**
A	-10	-1
B	-1	0
C	0	1
D	1	10
D (Cell 11)	10	40

The goal is to value each correct step done by a robot. In other words, in reinforcement modelling, learn some offensive game strategy with the ball possession is the goal. The rewards increase in value as the team advances zones in the playing field, seeking the Zone D, and Cell 14. In this stretch of the field , the agent will be closer to score. Therefore, for each zone, penalty and reinforcement values are set. The value of penalty is lower than the value of reinforcement, because the execution of an correct action corresponds to reinforcement. In Cell 11, the correct action chosen must be shooting.

3 Fuzzy Logic

3.1 Description

A Fuzzy Logic strategy is used to improve the intensity in marking. The soccer field was divided in three parts and according to the ball position, the agent keeps a secure distance to his opponent using Fuzzy Logic.

If the agent does not take the ball from his opponent, this secure distance will help the agent to recover his marking.

3.2 Fuzzy Logic Input

The Fuzzy Logic Input is described at Figure 2 in which the soccer field was divided in three areas: $Area_1$ (-52,5 a -35), $Area_2$ (-15 a 15) e $Area_3$ (35 a 52,5).

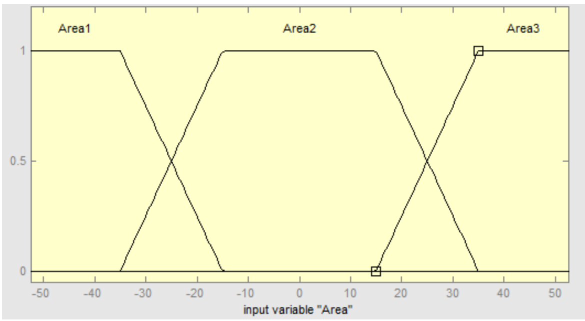

Fig. 2. Input variable Area.

3.3 Fuzzy Logic Output

The Fuzzy Logic Output is the security distance that the agent will be from his opponent. The variables are VeryClose, Close and Away, showing the distance which the agent is from his opponent. The output is described at Figure 3:

3.4 Rules

1. If Area = $Area_1$, than Mark = VeryClose;
2. If Area =$Area_2$, than Mark = Close;
3. If Area = $Area_3$, than Mark = Away;

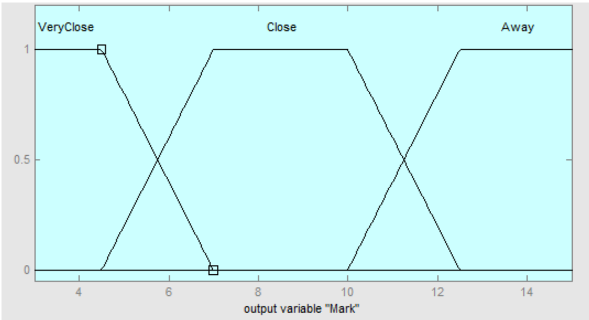

Fig. 3. Output variable Mark.

3.5 Response Curve

The graph obtained as response is shown at Figure 4 and is observed that when the ball is close of team's own goal, the marking is also closer. In otherwise, when the ball is in the field of attack, marking is not intense.

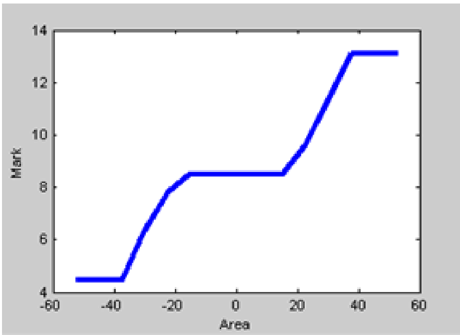

Fig. 4. Response Curve.

4 Conclusion and Futures Works

In recent years, the competition results and publications in RoboCup Simulation 2D platform have been improved in UFSJ. The modeling of reinforcement learning strategy has been improved by researches since 2011. As for fuzzy logic applied in marking, is necessary do some adjusts due to being a recent study in UFSJ2D Team. In future works, the team will increase knowledge in reinforcement learning and fuzzy logic to improve performance and results of UFSJ2D Team.

References

1. Agent2D: http://sourceforge.jp/projects/rctools/.
2. Akiyama, H. , Shimora H., Nakashima, T., Narimoto, Y., Okayama T.: HELIOS2011 Team Description. In RoboCup 2011 (2011).
3. Alavi. M., Tarazkouhi, M. F., Azaran, A., Nouri, A., Zolfaghari, S., Boroujeni, H. R. S.: Robocup 2012- Soccer Simulation League 2D Soccer Simulation Riton. In RoboCup 2012 (2012).
4. Carlson, P.: Warthog Robotics Team Description Paper 2012. In RoboCup 2012 (2012).
5. Celiberto Jr, L. A. e Bianchi, R. A. C.: Reinforcement Learning Accelerated by a Heuristic for Multi-Agent System. In 3rd Workshop on MSc dissertations and PhD thesis in Artificial Intelligence (2006).
6. de Boer e J. R. Kok, R.: The incremental development of a synthetic multi-agent system: The uva trilearn 2001 robotic soccer simulation team. Masters thesis, University of Amsterdam, The Netherlands (2002).
7. Hessler, A., Berger M. and Endert, H.: DAInamite 2011 Team Description Paper. In Robocup 2011 (2011).
8. Ikegami, T.; Kuwa, Y.; Takao, Y.; Okada, K.: 2D Soccer Simulation League Team Description Ri-one 2011. In Robocup 2011 (2011).
9. Kerbage, S. E. H., Antunes, E. O., Almeida, D. F., Rosa, P. F. F.: Generalization of reinforcement learning: A strategy for cooperative autonomous robots. In Latin American Robotics Competition (2010).
10. Monteiro, S. T. and Ribeiro, C. H. C.: Performance of reinforcement learning algorithms under Conditions of Sensory Ambiguity in Mobile Robotics. In Control and Automation Magazine, Vol.15 No.3 - July, August and September 2004 (2004).
11. Neri, J. R. F, Zatelli, M. R., Santos, C. H. F. and Fabro, J. A.: Team Description Paper GPR-2D 2012. In: Robocup 2012 (2012).
12. Ottoni, A. L. C., Lamperti, R. D. and Nepomuceno, E. G.: UaiSoccer2D Team Description Paper RoboCup 2011. In Robocup 2012.
13. Ottoni, A. L. C., Lamperti, R. D., Nepomuceno, E. G. and Oliveira, M. S.: Desenvolvimento de um sistema de aprendizado por reforo para times de robos - Uma anlise de desempenho por meio de testes estatisticos. XIX Congresso Brasileiro de Automtica, ISBN 978-85-8001-069-5, pp. 3557 - 3564 (2012).
14. Russell, S. J. e Norving, P.: Artificial Intelligence. Campus, 2nd edtion (2004).
15. Selvatici, A. H. P. and Costa, A. H. R.: Aprendizado da coordencao de comportamentos primitivos para robos moveis. Revista Controle e Automacao, vol. 18, pp. 173 186, 06 (2007).
16. Stone, P., Sutton, R. S. and Kuhlmann, G.: Reinforcement Learning for RoboCup-Soccer Keepaway. Adaptive Behavior, vol. 13, no. 3, pp. 165188 (2005).
17. Silva, A. T. R., Silva, H. G., Santos, E. G., Ferreira, G. B., Santos, T. D., Silva, V. S.: iBots 2010: Description of the Team. In Robocup Brazil Open (2010).
18. Sutton, R. S.; Barto, A. G.: Reinforcement Learning: An Introduction. The MIT Press. Cambridge, Massachusetts; London, England (1998).
19. Tao, L.; Zhang, R.: AUA2D Soccer Simulation Team Description Paper for RoboCup 2011. In Robocup 2011 (2011).
20. Tavafi, A., Nozari, N., Vatani, R., Mani Rad Yousefi, M. R., Rahmatinia, S., Pirdir, P.: MarliK 2012 Soccer 2D Simulation Team Description Paper. In RoboCup 2012 (2012).
21. Watkins, C. J. e Dayan, P.: Technical note Q-learning. Machine Learning, 8:279 (1992).