UaiSoccer2D: Team Description Paper Robocup 2012

Andre Luiz C. Ottoni, Rubisson D. Lamperti, and Erivelton G. Nepomuceno

Federal University of Sao Joao del-Rei, MG, Brazil
andreottoni@ymail.com,duartelamperti@yahoo.com.br,
nepomuceno@ufsj.edu.br
http://www.ufsj.edu.br

Abstract. This article presents the UaiSoccer2D team, a simulation team of robots soccer simulated at UFSJ - Federal University of So Joo del-Rei, MG, Brazil. Will be present a modeling strategy for use of reinforcement learning algorithm Q-learning.

Key words: reiforcement learning, Q-learning, UaiSoccer2D.

1 Introduction

The UaiSoccer2D Soccer Team is made for students, teachers of undergraduate and postgraduate of UFSJ - Federal University of Sao Joao del-Rei, MG, Brazil. The team has goal of develop research in area of Robotic, Artificial Intelligence, Statistic and Software Engineering applied Robotss Soccer. For this, UaiSoccer2D Team has sent some works [9], [10], [11], [12] and [16] and has participated of several events in country.

In 2011, our team organized and participated in a tournament called 1^a Mineira Cup of Soccer Simulation 2D , in which was winner. Also, competed in Robocup Brazil Open (2011) [1] where got the fifth position. In this competitions, the team used Base UVA Trilearn 2003 [4]. From now on, we will use Helios Base Agent2D [1]; [2].

In this paper, we will show a model from strategic used for team, called Reinforcement Learning [17]; [21]. This technique will be using to understand and automate the decision taking. This way, will not be necessary a management and not even complete models of surroundings of simulation. Q-Learning is the algorithm used in this paper when agent has the ball [17]; [21].

2 Reinforcement Learning

Reinforcement Learning (RL) has allowed the learning of agents via of life in the enviroment, in which he lives [8]; [17]; [21]. The RL has been frequently cited in several groups of simulation in Robots's Soccer [5]; [22]. Some works use the Q-learning algorithm in specifics cases: when only the agent has the ball [6]; [8]. In [3] was used technique in order to accelerate the learning.

[1] Robocup Brazil Open (2011): http://www.cbr2011.org.

2.1 Q-learning

The Q-learning algorithm allow to establish a politic of actions interactively [17]; [5]; [8]. 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:

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

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 [7].

3 Modelling UaiSoccer2D Strategy

The method for modelling UaiSoccer2D strategy is split in three parts like follows:

- Definition and discretization of the actions of agents;
- Definition and discretization of the states of the environment;.
- Definition of values of the reinforced table "R" for each pair state (S) X Action (A);
- Implementing in the simulator.

3.1 Modelling of Agent's Actions

At this moment we will show actions possible of agent into simulated Robots Soccer. The table 1 shows actions of agent with the ball.

3.2 Modelling the Enviroment

The interaction of agents with the virtual world is interpreted through the states of the environment. These states are defined the characteristics of the environment a game of robot soccer. The characteristics are taken into account the positioning of the robot's own team with the ball and the distance of opponents players.

3.3 Modelling the Reinforcement Matrix

The reinforcement matrixes have the fixed reward in which the agent will receive to each State (S) X Action (A) available. The lines shown the states of environment and the columns illustrate the actions which the agents can take. The positive reinforcements (5, 10 e 20) illustrate interesting situations to win the game. The null reinforcement (zero) shows that can exist better actions. The negative reinforcement (-1) indicates that the action is inappropriate to the current state of robot.

Table 1. Actions of agent with the ball.

Id	Action
1	Dribble Fast
2	Dribble Slow
3	Dribble Normal
4	Pass/Kick
5	Hold Ball
6	Advance Ball

Table 2. State of the environment to the agent with the ball.

Id	State	Description
1	Opp B	The opponent agent is positioned behind of player in reference to axis X. The distance between the two agents is more than 4.5 meters.
2	Opp Close B	The opponent agent is positioned behind of player in reference to axis X. The distance between the two agents is more than 2.5 meters.
3	Opp Very Close B	The opponent agent is positioned behind of player in reference to axis X. The distance between the two agents is less than 2.5 meters.
4	Opp F	The opponent agent is positioned forward of player in reference to axis X. The distance between the two agents is more than 4.5 meters.
5	Opp Close F	The opponent agent is positioned forward of player in reference to axis X. The distance between the two agents is more than 2.5 meters.
6	Opp Very Close F	The opponent agent is positioned forward of player in reference to axis X. The distance between the two agents is less than 2.5 meters.

Table 3. Reinforcement matrixes

State/Action	A1	A2	A3	A4	A5	A6
S1	-1	-1	-1	20	-1	-1
S2	0	-1	0	-1	-1	0
S3	5	-1	-1	-1	-1	-1
S4	-1	-1	-1	20	-1	-1
S5	-1	5	0	0	-1	0
S6	-1	-1	-1	10	10	-1

3.4 Implementing in the Simulator

The implementation stage of Reinforcement Learning strategy was performed in simulator of Robots's football "RcSoccerSim" in two dimensions of Robocup. For this, we adopt like base team the Agent2D (Helios Base). In the Q-learning algorithm, the parameters used were the learning rate (α) and discount factor (γ) fixed in 0.95.

To save the learning information for robots was made a file called *q.txt*. In this file, the Q matrix (6x6) of learning was started with the value zero to each state versus action pair what shown the privation of intelligence before simulation first.

The model presented only aims the learning when the agent has the ball. This way only a robot can access the file *q.txt* to each time, but the learning of agents is accumulating in Q Matrix. Resulting in communication between players called blackboard. The blackboard is a mutual structure to agents in which he can perform the written and read of learning information of team.

4 Results

In this paper, we checked the behavior of robot's team after two stages training of Reinforcement Learning Algorithm. The training was performed against a Helios2011 Team [2].

- Stage 1: After 30 (thirty) training games;
- Stage 2: After 90 (ninety) training games.

The evaluation of learning stages was through six random teams. In this set, we can find the UaiSoccer2D team (2011) [10] and the competitions teams of Worldcup 2011 (Robocup): Helios2011 (2 place/Japan) [2], Hfutengine2011 (9 place/China) [24], Aua2d (11 place/China) [22] e Rio-ne (13place/Japan) [23]; and Agent2D (Helios Base team) [1].

They were tried in 5 (five) games against the learning stages (1 and 2) in total 60 games for evaluation. In these games there was no learning. In table 4 are presented the data of games for evaluating each stage, in which the cell represent the sum of five games.

Table 4. Results For The Evaluation By Stage Of Learning.

Team	**Balance**	**Balance**	**Conceded**	**Conceded**	**Scored**	**Scored**
	Stage 1	**Stage 2**	**Stage 1**	**Stage 2**	**Stage 1**	**Stage 2**
Agent2D	-1	-6	17	11	16	5
Aua2d	23	22	0	1	23	23
Helios2011	-24	-13	25	14	1	1
Hfutengine2011	26	39	7	3	33	42
Rio-ne	23	30	8	4	31	34
UaiSoccer2D2011	120	121	0	0	120	121
Average	27,83	32,17	9,5	5,5	37,33	37,67

5 Conclusion and Futures Works

This paper showed a modelling to reinforcement learning for soccer of robots simulated. The development a technique of learning allows adapting the different

situations of game. At this moment, we made only routines to the situation, in which our team has the ball. In future works, we will make routines to the situation, in which the adversary team has the ball. Also, we will use some statistics tools to understanding the behavior of team with the Q-learning and others techniques of game strategy will be measure with reinforcement learning.

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. 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).
4. 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).
5. Hessler, A., Berger M. and Endert, H.: DAInamite 2011 Team Description Paper. In Robocup 2011 (2011).
6. 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).
7. Monteiro, S. T. e 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).
8. Neri, J. R. F, Santos, C. H. F e Fabro, J. A.: Description Of Team GPR-2D 2011. In: Robocup Brazil Open, 2011 (2011).
9. Ottoni, A. L. C., Nepomuceno, E. G., Oliveira, F. F. e Oliveira, M. S.: Analyze the behavior of cooperative multi-agent systems by means of statistical tests. In X Statistical Meeting Miner. (2011).
10. Ottoni, A. L. C., Nepomuceno, E. G. e Oliveira, F. F.: UaiSoccer2D Team Description Paper CBR 2011. In Robocup Brazil Open (2011).
11. Ottoni, A. L. C. ; Nepomuceno, E. G.: Application of Software Engineering in a Multiagent System Modeling Cooperative. In: I Meeting Miner Computational Modeling (2011).
12. Ottoni, A. L. C. ; Nepomuceno, E. G ; Oliveira, F. F. ; Oliveira, M. S.: The link between research and extension in a Soccer Team of Simulated Robots. In: I Meeting Miner Computational Modeling (2011).
13. Hawasly, M.; Ramamoorthy, S.: EdInferno.2D Team Description Paper for RoboCup 2011 2D Soccer Simulation League. In Robocup 2011 (2011).
14. Russell, S. J. e Norving, P.: Artificial Intelligence. Campus, 2nd edtion (2004).
15. 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).
16. Viriato Neto, C. M. ; Oliveira, C. S. ; Ottoni, A. L. C. ; Nepomuceno, E. G.: Modeling Strategies for Cooperative Multiagent Systems through the SCILAB mathematical. In: I Meeting Miner Computational Modeling (2011).
17. Watkins, C. J. e Dayan, P.: Technical note Q-learning. Machine Learning, 8:279 (1992).

18. Bai, A.; Lu, G.; Zhang, H.; Chen, X.: WrightEagle 2D Soccer Simulation Team Description 2011. In Robocup 2011 (2011).
19. Lau, N.; Reis, L.P.; Mota, L.; Almeida, F.: FC Portugal 2D Simulation: Team Description Paper. In Robocup 2011 (2011).
20. Fraccaroli, E. S.; Carlson, P. M.: FC Portugal 2D Simulation: A Fuzzy approach for modeling the team WARTHOGSIM. In Robocup 2011 (2011).
21. Sutton, R. S.; Barto, A. G.: Reinforcement Learning: An Introduction. The MIT Press. Cambridge, Massachusetts; London, England (1998).
22. Tao, L.; Zhang, R.: AUA2D Soccer Simulation Team Description Paper for RoboCup 2011. In Robocup 2011 (2011).
23. Ikegami, T.; Kuwa, Y.; Takao, Y.; Okada, K.: 2D Soccer Simulation League Team Description Ri-one 2011. In Robocup 2011 (2011).
24. Chen, R.; Li, Y.; Wang, H.; Fang, B.: HfutEngine2011 Simulation 2D Team Description Paper. In Robocup 2011 (2011).