iBots 2012 – Team Description Paper

Alexandre T. R. Silva, Carlos A. S. P. Rodrigues, Vinícius P. Jardim,
Ruben A. Nascimento, Vitor S. Silva, Herinson B. Rodrigues and Nilo S. Marques Jr*

Abstract **— This paper presents solutions developed by the iBots team in the RoboCup Simulation 2D category. The Agent2D framework was chosen as support of the team development. The solutions implemented are: Reinforcement Learning to improve the direct pass, and passing into space (through pass and leading pass).**

I. Introduction

This paper describes solutions developed by the iBots team of the RoboCup Simulation 2D category. The team is based in Agent2D 3.1.1 [1], which depends on librcsc library – the version used was 4.1.0 [2]. Among the solutions implemented there are: Reinforcement Learning to improve the direct pass, and passing into space (through pass and leading pass).

II. Passes

In the technical fundamentals of soccer, the pass is one of the most important. The pass is performed several times during a match and its quality can define the final score of matches. Wrong passes give the opponent the opportunity to counter attack, which is important in offensive strategy of the teams. The counter attack tries to make use of the opponent's defense disorganization.

Barbosa [3] conducted a study about the attack in soccer. He analyzed the teams ranked in the top two of some of the major leagues in 2008/2009 season. The selected teams were: Real Madrid, F. C. Barcelona (Spain), Internazionale and A. S. Rome (Italy), Manchester United and Chelsea (England) and F. C. Porto and Sporting C. P. (Portugal).

Barbosa [3] found that the counter attack was the method of attack with the most positive offensive sequences for all the teams observed, except for the Sporting team. Positive Offensive Sequence is every offensive sequence which results in a shot on target. For this study Barbosa considered "shots on target" the following situations: the ball completely crossed the goal line, was defended by the goalkeeper, or hit the crossbar.

If wrong passes result in counter attack, which is presented as the most effective offensive form, then it implies that individual technical skills (or the lack of quality of skills) influence on the collective performance of a team. In other words, teams with low quality pass tend to be more counter attacked, in proportion to the number of shots on target, and also tend to take more goals.

The pass aims to: advance the team, explore empty spaces and reduce the physical stress of the players. Passes can be classified into two groups:

- direct pass;
- passing into space.

A. Direct pass

Direct passes are defined by sending the ball to the exact position of a teammate. Therefore, the percentage of success of this type of pass is high and it is very used during a soccer match. These passes are implemented in Agend2D. However, it was found through empirical observations that the team lost the ball due to wrong passes, resulting in counter attack to the opposing team.

For this reason, we adopted a Reinforcement Learning algorithm to enable players to improve their passes. The choice of an appropriate technique requires compatibility between the way the information is processed and how it becomes available in the environment to be applied. The most popular algorithm of Reinforcement Learning is Q-Learning [4].

Because of the time restriction on soccer robots, it is necessary to define abstraction models to reduce the dimensionality of the state space. It is also important to develop models of learning and rewards that meet the specificities of what you want to learn, rewarding only those actions that result in 'good' states.

During the learning, the passes are validated and scored with efficiency values for each state. To provide interoperability between the policy decision of Reinforcement Learning and the policy already followed by the team, the act of following the original policy was abstracted as an action to be evaluated. Passing into space were not considered on the set of actions, only direct passes. Thus, the set of actions used in Reinforcement Learning is formed by six actions:

- direct pass to the nearest teammate;
- direct pass to the second nearest teammate;

* A. T. R. Silva is with the Federal University of Tocantins, Palmas, Brazil (arossini@uft.edu.br).
C. A. S. P. Rodrigues is with the Federal University of Tocantins, Palmas, Brazil (krezeberto@gmail.com).
V. P. Jardim is with the Federal University of Tocantins, Palmas, Brazil (vinipj@hotmail.com).
R. A. Nascimento is with the Federal University of Tocantins, Palmas, Brazil (rubenanapu@hotmail.com).
V. S. Silva is with the Fluminense Federal University, Niteroi, Brazil (vitorsousasilva@gmail.com).
H. B. Rodrigues is with the Federal University of Tocantins, Palmas, Brazil (herynson@gmail.com).
N. S. Marques Jr is with the Federal University of Tocantins, Palmas, Brazil (nilobohn@gmail.com).

- direct pass to the third nearest teammate;
- direct pass to the fourth nearest teammate;
- direct pass to the fifth nearest teammate;
- and policy decision of the agent.

Reinforcement Learning does not interfere with the political decision-native in situations where it is efficient. For this, the action modeled is initialized with an absolute bias, which always selects the native politics in states that have not yet received reinforcements.

The values of reinforcement are presented below.

- Success - pass successfully performed, a value of +1 point;
- Interception - when a pass is intercepted by the opponent, a value of -3 points;
- Offside position - receiver is offside, a value of -2 points;
- Poor pass - passes that should not be executed (e.g. receiver too far or out of bounds), a value of -7 points;
- Danger pass - receiver in the risk area (near the large area of defense), a value of -5 points;
- Own goal - a value of -60 points.

In the Q-Learning implemented, for each action performed by the agent we compute a reward and the expected value in following the best policy at a discount. The policy information is stored in an array (s, a), which stores the estimated values for each pair of state s and action a [5].

It is necessary to use a function of discretization to analyze the states. A mapping of continuous states would result in a very large number of states, making that Reinforcement Learning would not converge satisfactorily in computation time. The discretization of the state aims to reduce the number of states and make rapid analysis.

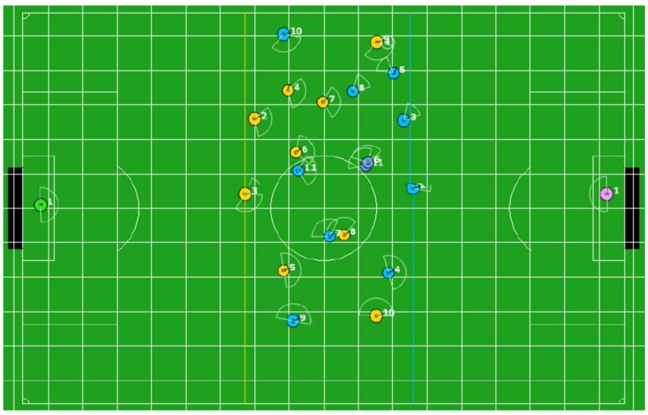

Figure 1. Division of the playing field in cells

The function of discretization implemented divides the playing field in cells of 6m x 6m. Figure 1 and 2 illustrate this process. The value of 6m of the cell size was defined to be the distance adopted as 'too close to the pass' in the iBots source code.

Figure 2. Example of discretizing the field in cells

To quantify the quality of the proposed solution implemented, there were 600 matches so the players could learn better passing situations. The matches were between the iBots team which has received the proposed solution with Reinforcement Learning and the iBots team without Reinforcement Learning [6] [7] [8].

To improve the visualization of the data, matches were grouped according to the temporal realization and the averages for each group were calculated. The first group contains 1-50 matches; the second group 51-100; the third group 101-150; the fourth group 151-200; the fifth group 201-250; the sixth group 251-300; the seventh group 301 350; the eighth group 351-400; the ninth group 401-450; the tenth group 451-500; the eleventh group 501-550; and the twelfth group 551-600. The result, in figure 3, is expressed in percentage (wrong passes number divided by the passes number).

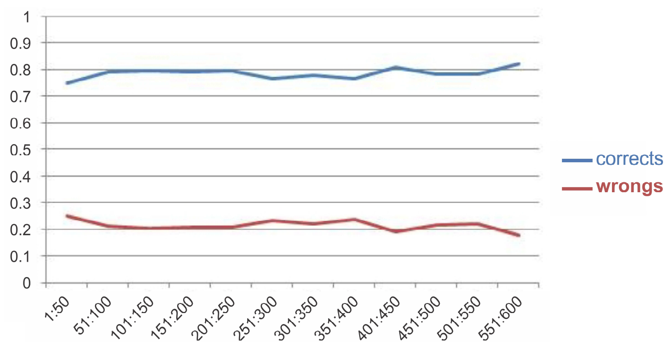

Figure 3. Percentage of average passes per group

We observed improvement in the quality of passes. However, we found that a greater number of matches is required for the algorithm to learn more states and improve what was already learned but did not reach the optimum condition. Other results can be found in [9].

B. Passing into space

It is used when a game is played at high speed. The player without the ball starts a movement, running to the empty space. The player with the ball must observe the pace and direction of the receiver. Timing and accuracy are essential.

Passing into space, exploring the length of the field, are passes designed to send the ball considering a projection of some teammate in relation to the goal. In other words, the ball is passed to some future position of a teammate in order to leave him in an empty space, preferably in a good position to shot on target. They can be divided into [10]:

- Leading Pass;
- Through Pass.

Communication between the agents involved in these types of passes is very important to optimize the timing of the movement. The new types of passes implemented by the iBots team are described bellow.

Leading Pass

Passes the ball to an empty space in order to find a teammate (leader) to receive the ball, which means, passes the ball to a space in front of the receiver. It usually occurs on the sides of the field.

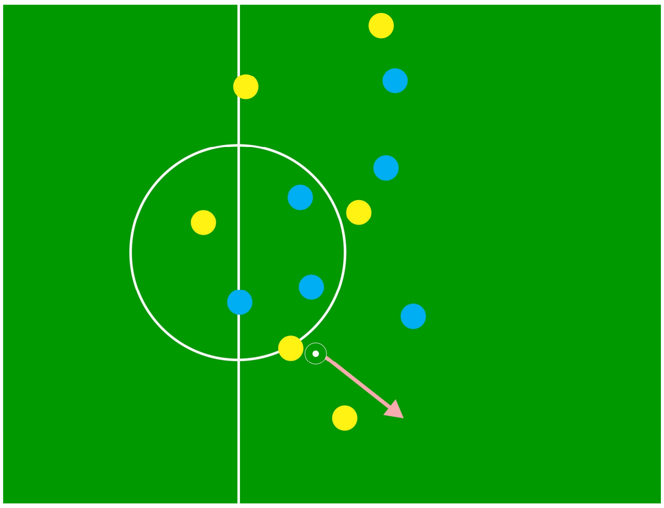

Figure 4. Example of a Leading Pass

Through Pass

Through Passes are also called Tunnel Passes or Piercing Passes. The ball moves between two adversaries toward an empty space, preferably between the line of defenders and the goalkeeper. It is used to let the striker alone against the opposing goalkeeper. Usually causes some type of disruption to the opposition defense.

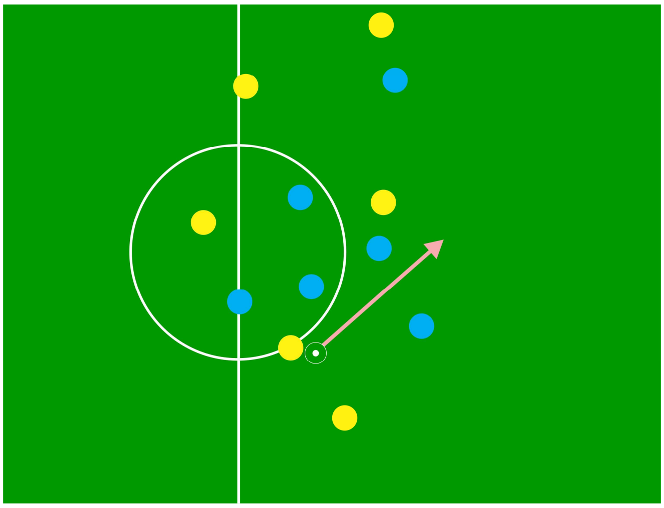

Figure 5. Example of a Through Pass

Implementation

To determine which point will be sent the ball we considered the following parameters:

- position of the player with the ball;
- central position (y-axis) of the goal;
- position of opponents in the center of the game (field region between the ball and the goal of the team that is defending);
- position of the teammates who are in the center of the game.

Some candidate points are collected and analyzed. The points that have lower quality are discarded and the most promising are selected. Then, the player with the ball calculates the force that would have to apply on the ball to make the pass. It checks the degree of security to pass the ball to the empty space either to Leading Pass or Through Pass.

In order to test the passes implemented, five matches were played against Agent2D. With a total of 25 passes, there were 13 correct passes (computed receiver received the ball) and 12 wrong passes (loss of ball possession, a teammate received the ball or the ball was out of play), see Table I. We obtained the mean of 52% correct passes and 48% wrong passes. In addition, considering only correct passes, the ball was received with mean error of 6.92 meters of distance to the processed position and mean angular errors of 8.21°.

TABLE I. RESULTS OF PASSING INTO SPACE

Leading and through pass	***Value***	***%***
Correct passes	13	52
Wrong passes	12	48
Total	25	100.00

III. CONCLUSION

The iBots team improved the passes because of their importance for soccer. This paper presented that wrong passes generate counter attacks, which are the most effective offensive.

We have implemented a Reinforcement Learning algorithm to increase the accuracy of the direct pass. There was improvement in the quality of passes, but it was necessary a lot of matches for the algorithm to converge all states to an optimum condition.

It was also implemented passing into space. These passes are used offensively and near the line of strikers in the tactical system. This is because passing into space are less accurate than the direct pass. Wrong pass in the defensive zone is very dangerous for the team.

Even less accurate, passing into space are used offensively because it destabilizes the opposition defense and tends to leave the attackers in privileged position to shot on target.

It is suggested as future works:

- extension of Reinforcement Learning for other skills besides passing, e.g. dribbling and strategic positioning;
- tests with different cell sizes and values of reward.

REFERENCES

[1] H. Akiyama (2012a), *Agent2D*, http://pt.sourceforge.jp/projects/rctools/downloads/55186/agent2d-3.1.1.tar.gz/

[2] H. Akiyama (2012b), *Librcsc*, http://pt.sourceforge.jp/projects/rctools/downloads/51941/librcsc-4.1.0.tar.gz/

[3] P. F. A. F. Barbosa, "Eficácia do Processo Ofensivo em Futebol: Estudo comparativo das equipas classificadas nos primeiros e segundo lugares das ligas nacionais de Espanha, Inglaterra, Itália e Portugal, na época de 2008/09". Bachelor's thesis, University of Porto, Porto, 2009.

[4] C. J. C. H. Watkins, "Learning from Delayed Rewards" Doctor's thesis, King's College, London,1989.

[5] W. Uther, and M. Veloso, *Adversarial reinforcement learning*, Technical report, Carnegie Mellon University, 2003.

[6] G. B. Ferreira, "Uso dinâmico de esquemas taticos com redes neurais perceptron de multiplas camadas na equipe ibots de futebol de robôs simulados" Bachelor's thesis, Federal University of Tocantins, Palmas, 2010.

[7] T. D. Santos, "Jogadas coletivas por meio de comunicação multi-agente para a equipe iBots da categoria de simulação 2D da robocup" Bachelor's thesis, Federal University of Tocantins, Palmas, 2011.

[8] A. T. R. Silva, G. B. Ferreira, T. D. Santos, V. S. Silva, E. G. Santos, H. G. Silva, C. A. S. P. Rodrigues, and T. S. Arruda. "iBots 2011: Descrição do time" presented at the Brazilian Robotics Competition, São João Del Rei, September 18–21, 2011.

[9] V. S. Silva, "Sistema de aprendizado por reforço para aprimoramento de passes na equipe iBots de futebol de robôs simulados 2D" Bachelor's thesis, Federal University of Tocantins, Palmas, 2011.

[10] A. Tavafi, N. Nozari, S. Rahmatinia, R. Vatani (2010) *LEAKIND'DROPS 2010 Soccer 2D Simulation Team Description Paper*.