

FCP_GPR_2014 Team: Joining Setplays from FCPortugal with Reinforcement Learning from GPR2D to Improve Decision-Making in Multi-Option Setplays

João Alberto Fabro^{1,3}, Bruno de Oliveira Oenning¹, Vinícius Brenner¹,
Luis Paulo Reis^{2,3} Nuno Lau⁴

fabro@utfpr.edu.br, {brunodfo, vini_brenner}@hotmail.com, lpreis@dsi.minho.pt,
nunolau@ua.pt

1. Robotic Research Group, Federal University of Technology-Paraná(UTFPR), Curitiba, Brazil
2. DSI/School of Engineering, University of Minho - Guimarães, Portugal
3. LIACC – Artificial Intelligence and Computer Science Lab., University of Porto, Portugal
4. IEETA, UA – Inst. Eng. Elect and Telematics of Aveiro, University of Aveiro, Portugal

Abstract. FCPortugal 2D is a team from Portugal (Universities of Aveiro, Minho and Porto) that participates in Robocup for a long time (since 1999, being the 2000 world champion in the 2D simulation category) with research focusing on multi-robot coordination methodologies. GPR, that stands for “Grupo de Pesquisa em Robótica” (Robotics Research Group as written in portuguese) is a 2D soccer simulation team from Brazil that uses Machine Learning techniques to improve the decision making of the robotic agents during the match. This document presents the approach developed by the joining of the two teams. From FCPortugal where used the base team, and the Setplay library. From GPR came the reinforcement learning techniques, that allow the decision making to be based on experience during simulated matches. Simulations results show that the proposed approach retains the good characteristics of the previous teams, and unite their strengths.

Keywords: Simulated Robot Soccer, Coordination, Machine learning, Setplays.

1 - Introduction

FCPortugal is a research team from 3 Portugal universities (Universities of Aveiro, Minho and Porto) that participates in Robocup since 1999, in various categories (Rescue, 2D and 3D soccer simulation, Small Size and Medium Size League). The main research goal of FC Portugal team is the development of techniques and tools that can be used in various dynamic competitive domains such as distinct RoboCup leagues. Recently, a complete framework for the specification, execution and graphic design of Setplays was introduced [1] [2] [3] [4] [5]. Setplays are sequences of actions that should be executed cooperatively by a set of players in order to achieve an objective during the match. The most common setplays are used for corner kicks and direct free kicks (for example after a fault in the game).

GPR-2D is the Robotic Research Group (“*Grupo de Pesquisa em Robótica*” in portuguese) robotic soccer team. This research group involves several universities

, p. 1, 2014.

from Brazil (the State University of South Paraná - UNIOESTE, Federal University of Technology - Paraná - UTFPR, and the Federal University of Santa Catarina - UFSC). The main research objectives are the development of adaptive intelligent approaches to control mobile robots, both simulated and real, to cope with complex, cooperative tasks [6] [7]. The focus of the research, specifically applied to the Robocup 2D simulation category, was the development of cooperation among distributed agents, using reinforcement learning techniques, such as Q-learning [8], to obtain cooperation and distributed problem solving, based on training during the matches.

During a period of one semester between the end of 2013 and the begin of 2014, one researcher from Brazil has spent its sabbatical leave at the University of Porto, Portugal, bringing together the efforts of the two research groups, and the result of this (ongoing) effort is briefly presented in this paper.

This paper briefly presents the Setplay Framework[1], developed by FCPortugal, in section 2. In section 3, the approach used by the GPR2D is presented. Section 4 highlights the main characteristics of the combined approach, highlighting some conclusions and discussions.

2. The Setplay Framework

Setplays (or set pieces) are commonly used in many team sports such as football, rugby and soccer¹. Their main purpose is to surprise the adversary team with a previously trained sequence of movements and passes, usually leading to a chance to shoot to the goal. Both human and robotic soccer players can benefit from such useful tool to obtain a high-level of coordination and cooperation.

The FC Portugal team have developed a complete framework for high-level Setplay specification and execution, that has been successfully applied to at least 3 different Robocup soccer categories: 2D and 3D simulation, as well as Middle Size League [9][10]. At first, it was proposed a standardized, league independent, XML based, specification language, that defines set-plays in enough details so that it is possible to directly interpreted and executed in any soccer league (including human soccer).

The Setplay is defined by a set of basic parameters, such as its name, region of the field where it start, number of players involved, and the situation (that can be any situation when a ball is re-inserted in the play: Kick-off, Throw-in, Direct or Indirect Free Kicks, Corner Kicks and Goalie ball reposition). After the initial information, the remaining of the Setplay's definition consists of a sequence of "Steps". Each "Step" has a number, and a set of positions that the players involved should take. A "Step" also depends on a "Condition", that must be satisfied before executing it. Each Step have one or more "Transitions" that indicate a set of actions that all the players involved must execute in order to reach another "Step". As an example of setplay, suppose that

¹ See "<http://www.professionalsoccercoaching.com/free-kicks/soccerfreekicks2>" for a description and example of a complete Setplay.

during a free kick, the player in possession of the ball (kicker) have to kick it towards a companion (receiver 1), while two other players (receivers 2 and 3) run to predefined positions preparing themselves for a possible pass in the second step. Thus, when receiver 1 receives the ball, there are two options for the continuity of the set-play: to pass to receiver 2, or to pass to receiver 3.

Together with the specification language, a graphical tool was developed to allow the easy creation of even complex setplays – Playmaker [11] in its first version and SPlanner as a second more elaborated version [12]. This interface allows the definition of complete setplays. The situation described in the previous paragraph can be viewed as described in the SPlanner tool in Fig. 1. In this case, player 3 is the kicker, player 4 is receiver 1, and players 2 and 5 take the role of receivers 2 and 3. The selection of which player takes which role is also solved by the library, that selects the players based on their distance to the positions of the roles at the beginning of the set-play.



Fig. 1 - Setplays definition tool - SPlanner - and example of setplay.

3 - The GPR-2D Team

The GPR-2D team was developed using as base the Agent2D source code, available in [13]. Based in this complete team, a new decision making procedure were implemented allowing the agents that where inside the attack area to take action based on previous experience – a reinforcement learning technique. Only the player in possession of the ball would take action based on previous experience using the reinforcement learning approach – the Q-Learning algorithm – that allows the simulated agents to select the best action when in possession of the ball. The reinforcement approach “reinforces” actions that lead to goals scored. Details of the learning procedure used can be found in [6]. In section 4 the application of this algorithm to the selection of Transitions when in multiple-choice Steps of a Setplay is presented.

4 – The Proposed Approach – Machine Learning applied to Action Selection on Multiple-choices Setplays

The approach proposed for the team FCP_GPR_2014 is to enable the selection of the next Transition when on a given Step of a Setplay using Machine Learning. The example of Setplay presented in Fig. 1 has several such States in which there multiple Transitions (Steps 2, 3, 7 and 9) . The graphical representation of the setplay as a directed graph in the lower left corner of the figure is detailed in Fig. 2.

In the original Setplay framework, the selection of which transition to execute was defined by conditions, i.e., if the receiving player is positioned, or if the pass have low probability of being intercepted, the Transition was chosen. If more than one Transition was “enabled”, usually the first one was executed.

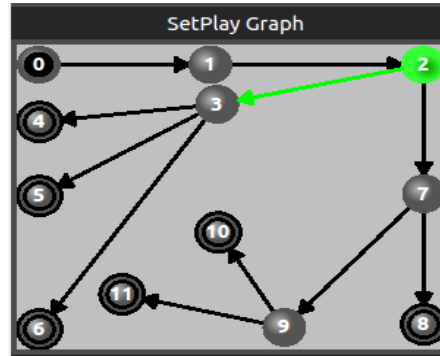


Fig. 2. Graph of the Setplay with multiple option of Transitions - example.

The approach proposed for the joining of the two teams for this year's competition is to take that decision based on previous experiences, using the Q-Learning algorithm to devise the better option. In order to obtain that adaptive behavior, and based on the approach used by GPR in previous competitions, a matrix relating Steps with Transitions was proposed. For the example presented in Figs. 1 and 2, the resulting matrix is shown in Table 1.

Thus, for this specific Setplay, 10 States were defined (lines), and 11 possible actions or Transitions (columns). This matrix Q is where the reinforcements are assigned by the Q-learning algorithm (only in the positions with the initial values “0” assigned to them). If, during the execution of the Setplay, the player with the ball chooses a transition, and the action is correctly executed (the setplay continues its execution), a reinforcement is realized. In our approach, the reinforcement is constant and values 100. To allow for execution of actions not tried before, or with low Q value, an ϵ -greedy approach is used (20% of the times, a Transition is chosen randomly, allowing execution of Transitions with lower Q values).

Table 1 – Q-Learning Matrix for the complete Setplay of Fig. 2.

Transition to→ Step:↓	1	2	3	4	5	6	7	8	9	10	11
0	0	-	-	-	-	-	-	-	-	-	-
1	-	0	-	-	-	-	-	-	-	-	-
2	-	-	0	-	-	-	0	-	-	-	-
3	-	-	-	0	0	0	-	-	-	-	-
4	-	-	-	-	-	-	-	-	-	-	-
5	-	-	-	-	-	-	-	-	-	-	-
6	-	-	-	-	-	-	-	-	-	-	-
7	-	-	-	-	-	-	-	0	0	-	-
8	-	-	-	-	-	-	-	-	-	-	-
9	-	-	-	-	-	-	-	-	-	0	0
10	-	-	-	-	-	-	-	-	-	-	-

4 – Conclusions and Discussion

The main advance proposed for this year's competition is the joining of the approaches of Setplays (from FCPortugal) with Reinforcement Learning (from GPR2D). In order to evaluate the proposed approach, several setplays with multiple options, similar to the one presented here, are being created. For each setplay, a file containing an empty matrix is created. After that, several matches must be executed, so that the Q-Learning algorithm is able to evaluate every option of every setplay during matches against different adversary teams. The obtained matrices correspond to the “optimum” policy, if enough game executions are realized. After that, the matrices can be “frozen”, and only the best option for each step is executed, according to the highest value present on the matrix. These simulations are being realized, and will continue until the time of the competition, but a few conclusions can already be drawn:

- The use of just one matrix (or one set of matrices, one for each setplay) can achieve the best selection of actions when all the games executed to obtain that matrices are always against the same team. When executing matches with diverse adversary teams, sometimes the “best” action, that works against one team, is not optimal against other team. This indicates the need to develop matrices that are “adversary specific”, i.e., that are adjusted by several matches against the same team (for example, using the binary team from the previous year's competition);

- As can be seen in Table 1, if the setplays build so that there are no cycles in the graph, the graph can be seen as a tree structure, and thus only one line of the matrix is needed to store all of the parameters, providing an economy of memory and time to load and save the matrices;

Future improvements of the approach include the use of more flexible reinforcements, for example using the position of the player to indicate higher or

lower values (players that receive the ball closer to the adversary goal area should receive bigger reinforcements). Other approaches are also under consideration for future development, such as using heuristic search to automatically adjust the positions of the players during the execution of each setplay.

Acknowledgements

The first author would like to thank CAPES for his scholarship, process No. BEX 9292/13-6. We would also like to acknowledge the Robocup community for its support, specially teams Helios [13] and CMUnited1999[14] for their excellent contributions to the community and source releases that are the base for our teams.

References

1. Mota, L.; Lau, N.; Reis, L.P.; Co-ordination in RoboCup's 2D simulation league: Setplays as flexible, multi-robot plans, 2010 IEEE Conf. on Robotics, Automation and Mechatronics, RAM 2010, pp. 362-367.
2. Mota, L.; Reis, L.P.; An Elementary Communication Framework for Open Co-operative RoboCup Soccer Teams, in Sapaty P; Filipe J (Eds.) 4th Int. Conf. on Informatics in Control, Automation and Robotics - ICINCO 2007, pp. 97-101, Angers, France, May 9-12, 2007
3. Mota, L.; Reis, L.P.; Setplays: Achieving Coordination by the appropriate Use of arbitrary Pre-defined Flexible Plans and inter-robot Communication, RoboComm 2007 - First Int. Conf. on Robot Communication and Coordination, Athens, Greece, October 15-17, 2007.
4. Mota, L.; Reis, L.P.; A Common Framework for Cooperative Robotics: an Open, Fault Tolerant Architecture for Multi-league RoboCup Teams, Int. Conf. Simulation Modeling and Progr. for Aut. Robots (SIMPAP), Springer, LNCS/LNAI series, pp. 171-182, Venice, Italy, Nov, 2008.
5. Lau, N.; Reis, L. P.; Mota, L.; Almeida, F. FC Portugal 2D Simulation: Team Description Paper, online, available at: <http://staff.science.uva.nl/~arnoud/activities/robocup/RoboCup2013/Symposium/TeamDescriptionPapers/SoccerSimulation/Soccer2D/>, consulted on Jan/2014.
6. Neri, J.R.F.; Zatelli, M.R.; Farias dos Santos, C.H.; Fabro, J.A.; , A Proposal of QLearning to Control the Attack of a 2D Robot Soccer Simulation Team, 2012 Brazilian Robotics Symposium and Latin American Robotics Symposium (SBR-LARS), pp.174-178, 16-19 Oct. 2012.
7. Fabro, J. A.; Botta, A. L. C.; Parra, G. A. P.; Neri, J. R. F.; The GPR-2D 2013 Team Description Paper, online, available at: <http://staff.science.uva.nl/~arnoud/activities/robocup/RoboCup2013/Symposium/TeamDescriptionPapers/SoccerSimulation/Soccer2D/>, consulted on Jan/2014.
8. Dayan, P. Technical Note Q-learning, Centre for Cognitive Science, University of Edinburgh, Scotland: University of Edinburgh, 1992.
9. Lau, N.; Lopes, L. S.; Corrente, G. and Filipe, N.; Multi-Robot Team Coordination Through Roles, Positioning and Coordinated Procedures, Proc. IEEE/RSJ Int. Conf. on Intelligent Robots and Systems – IROS 2009, St. Louis, USA, Oct. 2009.
10. Lau, N.; Lopes, L. S.; Corrente, G. and Filipe, N.; Roles, Positionings and Set Plays to Coordinate a MSL Robot Team, Proc. 14th Port. Conf. on Artificial Intelligence, EPIA'2009, Aveiro, LNAI 5816, Springer, pp 323-337, October 12-15, 2009.
11. Reis, L.P.; Lopes, R.; Mota, L.; Lau, N.; Playmaker: Graphical definition of formations and setplays, Proc. 5th Iberian Conf. Information Systems and *Technologies (CISTI)*, 2010.
12. Cravo, J. G. B. (2012). SPlanner: a graphical application for the Flexible definition of Setplays in Robocup (in portuguese). MsC. Dissertation, Integrated Master in Computer and Informatics Engineering, Faculty of Engineering, University of Porto. Online, available at: <http://repositorio-aberto.up.pt/bitstream/10216/62120/1/000149781.pdf>, consulted on Jan/2014.
13. Hidehisa Akiyama, Helios RoboCup Simulation League Team, online, available at: <http://rctools.sourceforge.jp/pukiwiki/>, consulted on: Jan/2014.
14. Peter Stone, Patrick Riley and Manuela Veloso, CMUnited-99 source code, 1999, online, available at: <http://www.cs.cmu.edu/~pstone/RoboCup/CMUnited99-sim.html>, consulted on Feb/2000.