Cyrus 2D Simulation 2019

Team Description Paper

Nader Zare[1], Mahtab Sarvmaili[1], Omid Mehrabian[2], Amin Nikanjam[1], Seyed Hossein Khasteh[1], Aref Sayareh, Omid Amini, Borna Barahimi, Arshia Majidi, Aria Mostajeran

1 K.N.Toosi University of Technology

2 Ottawa University

Nader.zare88@gmail.com

Abstract. The following team description paper demonstrates the activities carried out in the most recent year by Cyrus 2D simulation team, in brief. During this year, some tasks have been conducted in order to improve the team's offensive and defensive behavior, using machine learning. Some of the mentioned tasks will be looked into in this paper, with our main focus being on the implementation of deep reinforcement learning in Cyrus's defensive agents' decision making.

Keywords: Robocup, Soccer 2d simulation, Agent, Neural Network

1 Introduction

Cyrus was founded in 2012 and participated the RoboCup 2013 for the first time. This team has managed to achieve 2nd place in RoboCup 2018 and fourth place in RoboCup 2017. Also, it has won the 1st place title for IranOpen twice, once in 2014 and the other time in 2018, followed by its last achievement, being the 1st place in RoboCup Asia-Pacific 2018. Cyrus team members have created the starter base by simplifying the agent2D base [1] in cooperation with IranOpen's technical committee members in order to further improve the soccer simulation 2D league and motivate new teams for participation. Cyrus's base is agent2d.

2 Activities Done in Other Teams

In the recent years, the following tasks have been done in other teams: HELIOS has employed an action sequence planning framework as well as a knowledge sharing system to further optimize its use [2]. FRA-UNIted has integrated TensorFlow into their agent [3]. Oxsy has extended its coach attribution with adding abilities such as adaptive offside trap [4]. Mt has developed a new game log mining method [5]. A formation detection system as well as a software called Tournament Planning and Analyzing Software(TPAS) have been developed by Namira [6] [7] [8]. Razi has organized its own system for rating states in chain action (along with improving its basic behaviors) [9]. Nexus has implemented Reinforcement Learning for decision making in penalty area [10].

3 Previous Works in Cyrus

- 2013: Dynamic formation changing, field evaluation system for offense [11]
- 2014: field evaluation system optimization [12]
- 2015: Optimizing defensive decision making by using message passing between agents
- 2016: Shooting behavior optimization using Deep Neural Network
- 2017: Predicting opponent's behavior using Rough Neural Network [13]
- 2018: Predicting the behavior of the agents with ball possession [14]

4 Released Software

4.1 Cyrus 2014 Source

Cyrus 2014 won the 1st place in the IranOpen RoboCup 2014 and also became 5th in the RoboCup 2014, later that year. This team's source code can be found in the following link:

https://github.com/naderzare/cyrus2014

4.2 Starter agent & Starter librcsc

Cyrus team members, in cooperation with IranOpen technical committee of 2D soccer simulation league, have simplified the agent base and the librcsc library for the 2D soccer simulation starter league, designed specifically for high school students. The mentioned base and librcsc have been used in 2D soccer simulation starter league during both IranOpen RoboCup 2018 and RoboCup Asia-Pacific 2018. More than ten teams participated in each of the competitions, with more than fifty participants in total. All of the participants have used the base developed by Cyrus and IranOpen committee of 2D soccer simulation league. High level behaviors like passing, dribbling, and shooting have been omitted from this base. The base can be found in the links below:

- Starter Agent 2D: https://github.com/naderzare/StarterAgent2D
- Starter LibRcsc: https://github.com/naderzare/StarterLibRCSC

5 Defensive Decision Making, using Deep Reinforcement Learning

In this section we are willing to suggest a method that involves an agent observing the environment's information and choosing a defensive behavior accordingly. Each agent can choose any of the behaviors mentioned below:

Block: Choosing a position for the purpose of preventing the opponent with ball possession from dribbling. In Cyrus's algorithm for block, an angle is chosen as the opponent's dribbling angle as well as an average pace of dribbling. Then, as illustrated below, after the point where the agent reaches fastest is found, it would be chosen as the block behavior target.

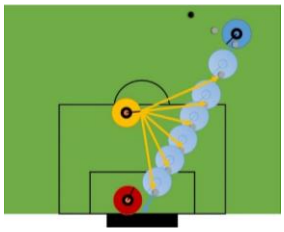

Fig. 1. Block Algorithm

Intercept: Moving to the position where the agent can reach the ball within the shortest amount of time.

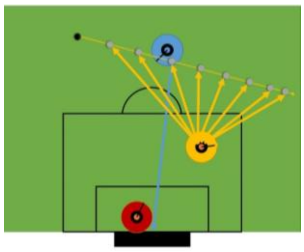

Fig. 2. Intercept Algorithm

Cooperation with Goalie: Moving to the position that aids the goalie in covering the goal. The agent, first finds the largest open angle between the ball and goal, then chooses a position on the bisector of the largest angle.

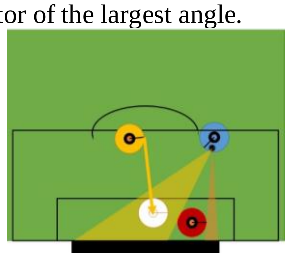

Fig. 3. Goal Defender Algorithm

Move: Getting to the point that is chosen according to the formation.

In order to reach the wanted target we implemented deep reinforcement learning, specifically DDPG. In this method two Actor-Critic neural networks are used for choosing a behavior and evaluation of the chosen behavior. The Actor neural network determines the best behavior according to the game state, while the Critic neural network evaluates the behavior in a game situation.

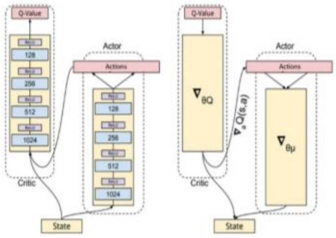

Fig. 4. Actor-Critic architecture [15]

Experience queue, target network, bounded parameter space learning, and Monte Carlo algorithm have been utilized in order to stabilize the learning process. Each episode consists of a number of steps and it would end if any of the following conditions occurs and a particular amount of points would be rewarded to the agent by the environment.

Passing of 150 cycles	10
Ball leaving the field	10
Gaining possession by the goalie or defenders	10
A goal being conceded	-10

As mentioned earlier, the actor-critic method uses two neural networks to first choose the best behavior, and then the critic neural network comes in for evaluating the chosen behavior. We have applied the following equations for these two neural networks' learning.

$$Q(s,a) = Q(s,a) + \alpha(r + \gamma \max_{a'} Q(s',a') - Q(s,a))$$

$$L_Q(s,a|\mathrm{q}^Q) = (Q(s,a|\mathrm{q}^Q) - (r + \gamma \max_{a'} Q(s',a'|\mathrm{q}^Q)))^2$$

$$L_Q(s,a|\mathrm{q}^Q) = (Q(s,a|\mathrm{q}^Q) - (r + \gamma Q(s',\mu(s'|\mathrm{q}^\mu) \mid \mathrm{q}^Q)))^2$$

$$\nabla_{\mathrm{q}^\mu}\mu(s) = \nabla_a Q(s,a \mid \mathrm{q}^Q)\, \nabla_{\mathrm{q}^\mu}\mu(s \mid \mathrm{q}^\mu)$$

It can be concluded that the critic neural network is taught in the following step according to the received evaluations of environment by the agent and the actor neural network's output, whereas the actor is taught by the critic's network error. In order to execute this method, HFO [16] and caffe [17] have been used. This method has also been used on offensive agent's decision making, goalie cooperating defender's decision making, and multi-agent decision making while defending with a little adjustment added to it. In this paper we will only explain its execution on the goalie cooperating defender, parametrically, using high level behaviors.

5.1 Goalie cooperating defender against offensive opponent

In this experiment, a defender must carry out any of the four behaviors being intercept, move, block, and goal defending. We also have used Cyrus's goalie and HELIOS's forward for the practice. This approach had been taken in [15], using the kick, dash and turn behaviors. High-state was used during the experiment as suggested in [15]. The experiment was carried out using three different approaches.

First approach: The defender chooses one of the four behaviors listed earlier. Cyrus 2018 was used for the execution of these behaviors.

Second approach: Defender chooses a behavior, using Cyrus 2018 for executing it, just like the first approach, but before doing so, some random parameters would be added to the behavior target.

Third approach: One of the behaviors is chosen and then carried out by the defender, using one of their parameters, which are shown here:

- **Block:** For dribbling; average pace of dribbling
- **Formation:** Switching the formation target parametrically in the x-y axis
- **Goal defender:** Altering the agent's distance from goal on the bisector of the largest open angle.

As illustrated in the graph below, Goal Percent has been shown in connection to the Actor Update Step.

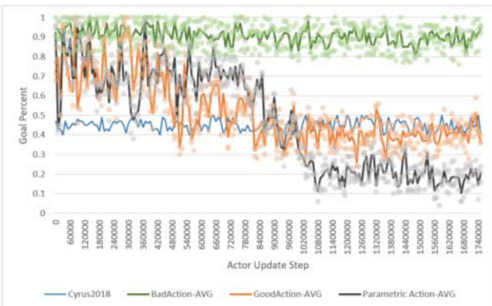

Fig. 5. First Approach (Good Action-AVG), Second Approach (Bad Action-AVG), Third Approach (Parametric Action-AVG)

It is concluded from the shown results that learning the parameters led to a better outcome for the agent's learning process. This method was executed using the following items in the multi-agent situation and will be added to Cyrus 2019:

- Addition of mark behavior
- Opponent players' behavior predictor
- Friendly agents' behavior predictor

6 Future Works

Cyrus team members are currently working on several projects with the hope of seeing these projects in practice for the upcoming RoboCup 2019. Some of these projects are as followed:

- Predicting opponents' behavior with ball possession
- Predicting opponents' movements as well as friendlies'
- Optimizing body smart kick
- Optimizing defensive decision making, implementing reinforcement learning
- Opponent formation detection

7 References

[1] H. Akiyama, "Agent2D Base Code," [Online]. Available: http://www.rctools.sourceforge.jp .

[2] N. Tomoharu, H. Akiyama, Y. Suzuki, A. Ohori and T. Fukushima, "HELIOS2018: Team Description Paper," in *RoboCup 2018*, Nagoya, Japan, 2018.

[3] T. Gabel, P. Kloppner and E. Godehardt, "FRA-UNIted — Team Description 2018," in *RoboCup 2018*, Montreal, Canada, 2018.

[4] S. Marian, D. Luca, B. Sarac and O. Cotarlea, "OXSY 2018 Team Description," in *RoboCup 2018*, Montreal, Canada, 2018.

[5] Z. Yang, Z. Liu, X. Wang, N. Dong, X. Hu, J. Shengbing Chen and G. Lv, "MT2018: Team Description Paper," in *RoboCup 2018*, Montreal, Canada, 2018.

[6] E. Asali, F. Negahbani, S. Tafazzol, M. S. Maghareh, S. Bahmeie, S. Barazandeh, S. Mirian and M. Moshkelgosha, "Namira Soccer 2D Simulation Team Description Paper 2018," in *RoboCup 2018*, Monreal, Canada, 2018.

[7] E. Asali, A. Moravej, S. Akbarpoor, O. Asali, M. Katebzadeh, S. Tafazol, F. Negahbani, M. Valipour, S. Mirian, H. Mehran, A. Bitaraf Haghighi and Y. Khoshghadam Borazjani, "Persian Gulf Soccer 2D Simulation Team Description Paper 2017," in *RoboCup 2017*, Nagoya, Japan, 2017.

[8] E. Asali, M. Valipour, N. Zare, A. Afshar, M. Katebzadeh and G. Dastghaibyfard, "Using Machine Learning approaches to detect opponent formation," *Artificial Intelligence and Robotics (IRANOPEN)*, pp. 140-144, 2016.

[9] M. Noohpisheh, M. Shekarriz, S. Barzegar, D. Borzoo and A. kariminia, "Razi Soccer 2D Simulation Team Description Paper 2018," in *RobCup 2018*, Montreal, Canada, 2018.

[10] M. Abolfazli Esfahani, M. Ghafouri, M. Jamili, S. Askari, R. Etemadi, H. Mohebbi, M. Dara, M. Erafani and M. Malekabadi, "Nexus2D Team Description Paper," in *RoboCup 2017*, Nagoya, Japan, 2017.

[11] R. Khayami, N. Zare, H. Zolanvar, M. Karimi, P. Mahoor, F. Tekrar, E. Asali and M. Fatehi, "Cyrus Soccer 2D Sumilation Team Description Paper," in *17th annual RoboCup International Symposium*, Eindhoven, Netherlands, 2013.

[12] R. Khayami, N. Zare, M. Karimi, P. Mahor, A. Afshar, M. S. Najafi, M. Asadi, F. Tekrar, E. Asali and A. Keshavarzi, "CYRUS 2D simulation team description paper 2014," in *RoboCup 2014 Symposium and Competitions: Team description papers*, Joao Pessoa, 2014.

[13] N. Zare, A. Najimi, M. Sarvmaili, A. Akbarpour, M. NaghipourFar, B. Barahimi and A. Nikanjam, "CYRUS 2D Simulation Team Description Paper 2017," in *21th annual RoboCup International Symposium*, Nagoya, Japan, 2017.

[14] N. Zare, M. Sadeghipour, A. Keshavarzi, M. Sarvmaili, A. Nikanjam, R. Aghayari, A. Firouzkoohi, M. Abolnejad, S. Elahimanesh and A. Akhgari, "Cyrus 2D Simulation Team Description Paper 2018," in *RoboCup 2018*, Montreal, Canada, 2018.

[15] M. Hausknecht and P. Stone, "Deep Reinforcement Learning in Parameterized Action Space," in *International Conference on Learning Representations (ICLR)*, San Juan, Puerto Rico, 2016.

[16] K. Shivaram, L. Yaxin and S. Peter, "Half field offense in RoboCup soccer: A multiagent reinforcement learning case study," *RoboCup-2006: Robot Soccer World Cup X*, vol. 4434 of Lecture Notes in Artificial Intelligence, no. ISBN 978-3-540-74023-0, pp. 72-85, 2006.

[17] J. Yangqing, S. Evan, D. Jeff, K. Sergey, L. Jonathan, G. Ross, G. Sergio and D. Trevor, "Caffe: Convolutional Architecture for Fast Feature Embedding," *arXiv preprint arXiv:1408.5093*, 2014.