

Alice2021: Team Description Paper

Fan Wang, YuTang Guo,XingQi Zheng and XinYu Liu

Robotics and Innovation Technology Lab
School of Computer and Information
HeiFei Normal University,P.R.China

Miracle.hftc@gmail.com

Abstract. This team description paper mainly explains the work of Alice2D at this stage.According to the characteristics of our team, based on the past research, our team is committed to studying the differences of players' roles and reinforcement learning algorithm,changes to the evaluator,formation optimization. After many tests, we continuously optimize the team code, which has greatly improved the performance of the court.

Keywords: Player role differences, Monte Carlo tree algorithm, Decision tree

1 Introduction

Alice2D was founded in 2016, by the Hefei Normal University Student Innovation Laboratory of a group of robot-loving soccer students. This is the fifth time we apply for the chance to participate in RoboCup 2D competition. Our team code is based on Agent2D3.1.0[1] which was released by Hidehisa Akiyama. It also includes fragments of released code of Helios2008[2]. We sincerely appreciate the assistance from Hidehisa Akiyama and all of the related open source developers. In the 2017 World Tournament, the team won the seventh place in the regular season and won the championship in the challenge, 7th place in the 2018 World Cup regular season.In China, our team is also named Miracle, winning the third place in China division of the 2019 RoboCup World Cup and the third place in China division of the 2021 RoboCup World Cup.

2 Focus of team science

On the basis of manual strategy research, our team made improvement, planned for the role of agents in the course, and adopted artificial intelligence methods, namely reinforcement learning and decision tree learning. In the process of reinforcement learning, by constantly trying actions and observing the rewards of actions, the agent gradually chooses the beneficial actions in various situations, so as to obtain a high cumulative reward value in the process of interacting with the environment.

3 Player role differences

Robot Soccer players have different positions, including striker, center forward, midfielder, defender and goalkeeper. Different players have different strategies. For example, strikers tend to attack, midfielders tend to organize, and defenders tend to defend. Due to time constraints, our team only made improvements on forwards and defenders, and built different models according to different strategies and movements. The following are introductions based on player roles.

Forward Player When the striker in the chart on the left side of the position, the goalkeeper is located in near the goal area, then shot range should be in the range of the two red and blue line, the shooting range larger, shooting success rate is high, but if the striker missed shot time, direction along the dotted line to the right, the other goalkeeper will choose attack, narrowing the scope of a shot. At this time, the shooting range is greatly reduced compared to the Angle of the shot just now, and the success rate is greatly reduced. Therefore, the success rate of the shot will change with the timing of the shot[3].

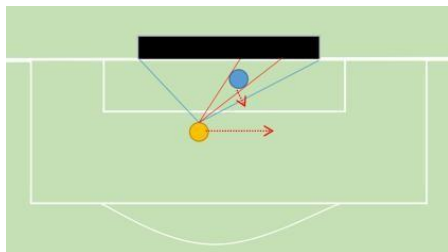


Figure 1:Before moving

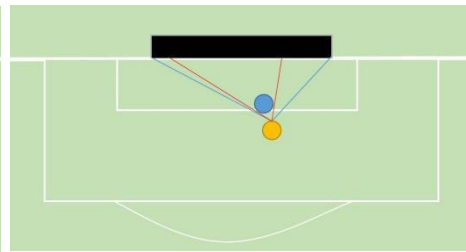


Figure 2:After moving

In RoboCup2D , shoot_success(dist, α , p, ε)is[4]:

$$\text{shoot_success}(\text{dist}, \alpha, p, \varepsilon) = \frac{1}{\varepsilon + 1} * \left\{ \frac{1}{(1 + \text{dist})^2} * \frac{\alpha}{\pi} + \frac{p * \alpha}{p_{\max}(1 + \text{dist}) * \pi} * \left[1 - \frac{1}{(1 + \text{dist})^2} * \frac{\alpha}{\pi} \right] \right\} \quad (2-1)$$

α indicates the perspective from which a forward player can shoot, dist indicates the shortest distance to shoot, p indicates the ability value of a forward player, ε indicates the interference factor of the goalkeeper[5].

$$\alpha = \frac{\arccos[(x_1 - x)(x_2 - x) + (y_1 - y)(y_2 - y)]}{\sqrt{(x_1 - x)^2 + (y_1 - y)^2} * \sqrt{(x_2 - x)^2 + (y_2 - y)^2}} \quad (2-2)$$

According to (2-1) and (2-2), players should try to choose left or right side to attack the goal, and try to avoid attacking the goal (a) in the area directly facing the goal. This method can be used for reference. On this basis, we optimized the shooting strategy based on this point. While choosing the left or right side to attack the goal, we circled with the opposing player to find the best shooting point (b).Practice with 100 shots:

Table 1. Shooting training results

Before and after improvement	goals	Percentage of goals scored
before improvement	49	49%
after improvement(a)	65	65%
after improvement(b)	71	71%

Defender player The quality of defense is the key to victory. The intercept plays an important role in the defense process. Whether the intercept is successful or not, players have to stop on the floor for several cycles to achieve the actual needs, so the strategy of intercept is bound to affect the cooperation of the whole team. At this point you need to change the probability of the tackle in the hope that the tackle will be less costly.The traditional player execution strategy is based on the decision tree to select the action[7,8]. The execution process of the decision tree is shown in the following figure:

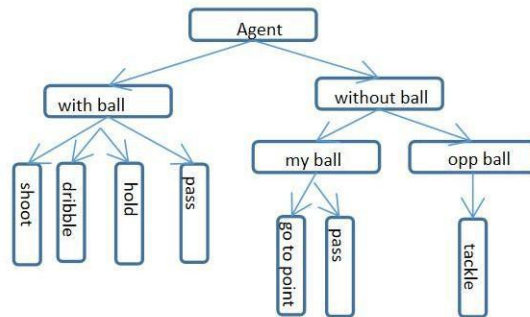


Figure 3:single agent decision-tree

Our team's defense is based on possession and off-possession strategies. Focuses on producing many different actions for the chain action model, the holder uses the evaluation function to assign different chains to operate on a point, and then selects the highest point of the chain of operations to be executed.

4 Reinforcement learning algorithm-Monte Carlo algorithm

Our Alice team has studied the combination of Monte Carlo algorithm and action chain in the past. Monte Carlo algorithm is similar to strategy iteration in dynamic programming, which also includes strategy iteration and strategy optimization[6]. Monte Carlo algorithm is suitable for learning problems in unknown environment.

Agent2d basic code generates two depth chain action, the basic model of evaluating the code only consider the first step in the operation, lead to can't choose the optimal solution, so we change the basic code evaluation model, based on the original to form the best action chain reaction, we added a chain action depth search depth and evaluation model[9]. The creativity of the algorithm is to follow up on the original basis, improve the decision tree for the position of the distance from the goal, such as outside the penalty area, inside the penalty area, and optimize the movement chain.



Figure 4:Best chain action is our player no.11 pass ball to no.10 and no.10 then pass ball to no.9.Then no.9 can shoot. Because of pass can be generated as the second action in chainaction and use Monte carlo tree search to find best

5 Changes to the evaluator

Our team modified the evaluator, including the modification of the goal area, the modification of the actions executed in the previous period and the evaluation of the player types. The evaluator was appropriately modified in combination with the formation of our team in order to expect the optimal actions to be executed.

6 Formation optimization

Analyze the RCG file[10] to obtain the data file containing the coordinates of each player in each period. The two-dimensional coordinate source data needed for the initial experiment can be obtained, and then the source data are mined and analyzed.

The server has been updated, which causes a deviation between the previous server version and the existing server version. Fix formation parameters, debug and change.

7 Summary and Outlook

Over the course of the year, our team made improvements in combination with formations and evaluators, continued to refine the decision tree for Monte Carlo algorithms, and combined strategy modifications for on-court player types. Today, our team still exists many problems need to solve, after last time communication with yushan team learning, inspire the study of digital twin framework[11], but our team merely in theory. Our team also attempts and manages to implement a more stable multi-layer action chain based on the relevant principles in TDP of HELIOS2012, and to optimize the state-space search by means of the addition of direct pass, short dribble and other actions between the two layers[12].

In the future, our team will continue to scripting language, data mining, data, deep learning log analysis, etc. On the basis of manual strategy, high level action decision making is studied.

The RoboCup is a good opportunity to learn from other teams, we will learn the strengths of other teams, improving our weaknesses. Wish the RoboCup simulation 2D soccer is getting better and better.

References

1. Akiyama, H.: Agent2D Base Code. <http://www.rctools.sourceforge.jp>. 2015
2. Hidehisa Akiyama, Tomoharu Nakashima, HELIOS Base: An Open Source Package for the RoboCup Soccer 2D Simulation. <http://pumath.dl.osdn.jp/rctools/32271>
3. Hui Zhou. Research on robocup2d simulation football team agent collaboration[D]. Nanjing University, 2013.
4. Potter M.A, De J.K. Cooperative co-evolution: an architecture for evolving coadapted subcomponents[J]. Evolution Computation, 2000, 8(1): 10-29.
5. Parker, Gary B, Blumenthal J. Punctuated Anytime Learning For Evolving A Team[C]. World Automation 2002 Congress Proceedings, 2002.

6. Shiyang chen.Research on Robocup_2d High-level Tackling Strategy Based on Reinforcement Learning[D]. Anhui University,2014
7. ZHAO Fajun, LI Longshu. Robo Cup regional cooperative strategy based on multi-Agent Q-learning. Computer Engineering and Applications, 2014, 50 (23) : 127-130.
8. Xiaofeng Gu,Daiyuan Zhang.Design of catching strategy for robot soccer game[J]. Computer Applications, 2005, 25(8): 1858-1859.
9. Deng Yong,Xiao Zhou,Siyang Chen,Zhengxi Gan,Alice Robocup Simulation 2D Team Description,ROBOCUP2017 SOCCER SIMULATION 2D LEAGUE,2017
- 10.YAMAMOTO A,OKA N,Noda I.Imitative learning from the logs of a soccer simulator[C].The 19th Annual Conference of the Japanese Society for Artificial Intelligence
- 11.Zekai Cheng,Ningyu Xie,Sichun Yang,Xingxing She.Application of digital twin framework based on Robocup simulation 2d[J].Journal of Suzhou University of Science and Technology,2019
- 12.H.Akiyama, H.Shimora, T.Nakashima, Y.Narimoto, K.Yamashita, HELIOS2012 Team Description Paper, The 16th annual RoboCup International Symposium, Mexico, 2012.