HFutEngine2016 Simulation 2D Team Description Paper

Keting Lu, Jun Ma, Zhenfeng Cai, Hao Wang, BaoFu Fang
HeFei University of Technology
fangbf@hfut.edu.cn

1.Indroduction

HfutEngine2D was founded in 2002 and participated in China Robot Contest the same year. In the following years, HfutEngine develops rapidly and participates in various competitions. Since 2003, we used the UVA BASE2003 as our underlying code, along with the server version of the update in intercepting the ball on the learning, BP neural network algorithm is adopted, in terms of choice of action adaptive, learning algorithm based on value is adopted, in passing action learning opportunities perspective-taking reinforcement learning algorithm and a series of machine learning algorithm are also adopted, achieving good effect. We once won the 2007 Robocup China open runner-up, 2008 World Cup in Robocup seventh place, Iran open tournament the bronze in 2008, 2009 Robocup China open simulation 2d group of fourth place, 2009 Iranian open tournament the champion, 2011 Robocup China open simulation 2d group, we use Agent2D as our underlying code for HfutEngine2D construction in 2014. In the 2015 World Cup we won 8 place of the simulation of 2d group. If this year for the World Cup for the opportunity, will be our ninth consecutive World Cup simulation of 2d. Since the end of the 2015 World Cup, we did some effective work our perception of information and decision-making in all.

2.Improvement of Perception of Information

(1)Reinforcement Learning Based on Statistical History Visual Information

In Agent2D there exists over-believe problems concerning the process of visual update and usage. Visual update process in underlying code performs as follows: First world model get all kinds of information from the server after parsing visual information from the server, and then transform distance between the agent and other players as well as relative angle into global coordinate and global speed. In the history visual information, if we found this player then update its information. If not, set the variable velocity error and then according to the speed error of the current players seen by the agent and the very players in history, we update the player which has the minimum error and it is within the error range. If the first two steps do not updated player's history visual information, create new player's visual information. Finally, we sort the visual information of the player object based on the credibility and delete the players information which is lower than a threshold value of credibility(The fact is in the underlying code, the lower the threshold value is, the more reliable it will be. For perceptual intuition, we use the opposite rule in the following explanation).Here are the threshold formula, where X is the horizontal coordinate of the football field:

$$f(x)=\begin{cases}-0.003X^3+0.206X^2-3.110X+25, & \text{where } 0 < X <= 37\\ 0.1X^2-10.5X+277.6, & \text{where } 37 < X <= 52.5\end{cases}$$

Although the above-described process has filtered the low credibility of the historical information but not detailed enough, we found that you can use a credibility threshold which is lower the longer the distance the credibility of the current cycle, based on the credibility of the current visual information. After that, we build a mathematical model of success rate of passing information to reflect accuracy of the extent of different effects.Our statistical findings choose to use the highest success rate of passing the threshold. The data in Table 1 is a comparison of pass rates under different thresholds with the Yushan team.

Next requirement of visual history information varies at different time, for example, when Central backcourt ball is passed to upfield, we need three most recent period for visual information to update while scraping near the first field requires visual information of the current period and cannot believe the history information. As a result, we use the mechanism of dynamic threshold in passing, kicking and other movements generating function of the target point and add to the player x-coordinate of the visual information credibility threshold for different time filter confidence range of different visual history information.

Table1: Comparison of pass rate under different thresholds

threshold	Number of games	Total number of passes	Successful passes	Failed passes	Successful pass rate
30	5	614	571	43	92.99%
25	5	613	575	38	93.80%
20	5	624	573	51	91.83%
15	5	603	556	47	92.20%
10	5	628	576	52	91.72%
5	5	618	574	44	92.88%

(2) Dynamic Threshold

Above, thresholds are the same in whole stadium area which do not vary from region to region. We found that the threshold of a certain player can be different in different regions, so we can obtain information as much as possible in the event that we also determine the information is relatively accurate. Therefore, players in the rear has a lower threshold for more player information while players in the front higher because they need to pay more attention to the accuracy of information.

We set 5 for scale interval of x coordinate so that the entire course can be divided into many regions. Respectively, the statistics within each region the highest success rate of passing the threshold will be calculated. After calculation, the scattergram will be shown as follows:

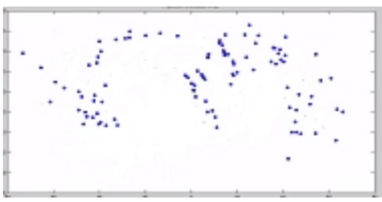

Figure 1: Distribution of Thresholds in different regions

We use K-means algorithm to cluster samples into k sets:

(1)Select randomly k cluster centroids $\mu_1, \mu_2, \ldots, \mu_k \in \mathbb{R}^n$

Repeat the following steps until the result converge:
(2)For each sample i, we calculate and cluster it into a certain set using the following formula:

$$c^{(i)} := \arg\min_j ||x^{(i)} - \mu_j||^2.$$

(3)For each set, we recalculate its cluster centroid using the following formula:

$$\mu_j := \frac{\sum_{i=1}^m 1\{c^{(i)} = j\} x^{(i)}}{\sum_{i=1}^m 1\{c^{(i)} = j\}}.$$

Finally, we dynamically adjust reliability threshold and select visual information according to clustering results in different x-coordinates.

(3) Improvement of Body Angle Relocation

Last year, the TDP [12], we have discussed the problem of improving the efficiency of passing by adjusting the angle of the body, but did not take into account the ball speed and prediction of passing point position as well as other considerations, so that the body adjusts to the ball fluency improvement is not obvious. Now we use the content of the work described in A above to obtain a more accurate historical information of teammate players, thereby reducing the error between passing predict points and the final passing point points, therefore, enhancing the importance of physical relocation. By adjusting the body angle of the agent who will get the passing ball, we improve effective use of speed of the ball which moves forward and reduce the power of the ball required, and also to some extent, reduce the waste of strength. To achieve this, we have adopted BP neural network, the location information of the players and the position of the ball, speed of the ball, location of the ball in next several cycles, etc. as input parameters X1, X2 ... Xn. w1, w2 ... wn are the corresponding weights of each parameter.

Input： $\mathrm{Net} = x1 * w1 + x2 * w2 + \ldots + xn * wn$

Output: $y = f(\mathrm{net}) = \frac{1}{1+e^{-net}}$

(4)Communication Protocol of hearing information

In traditional unused hearing team, the visual message of confidence occupy the vast majority of players. At the same time bring the visual world model wrong information due to the blind guesses. Obviously, this is fatal for a team high-level policy. Wrong information can lead players to lost a decision point which can be selected, or choose a wrong decision point, resulting in error of the entire decision-making .In all versions of Helios_Base code, there are design for hearing. Using 64bit to encode all the information, it can solve the whole problem of information coding, but process of decoding can be really fuzzy. For some teams just getting started, there exists difficulty in hearing mechanism learning and usage, the traditional encoding reuse also reduced because of excessive consolidation.We use a simplified propaganda mechanisms and encoding formats:

Sort	Fisrt Byte	Second Byte	Third Byte	Fourth Byte
BallMessage	b	pos_vel （5）		
PassMessage	p	unum_pos(4)	pos_vel(5)	
InterceptMessage	i	unum(1)	cycle(1)	
GoalieMessage	g	pos(3)	body(1)	
GoalieAndPlayerMessage	e	pos(3)	body(1)	num_pos(4)
OffsideLineMessage	o	x_rate(1)		
DefenseLineMessage	d	x_rate(1)		
PassRequestMessage	h	pos(3)		
StaminaMessage	s	rate(1)		
DribbleMessage	D	count(1)	pos(3)	
BallGoalieMessage	G	bpos_bvel(5)	gpos(3)	gbody(1)
OnePlayerMessage	P	unum_pos(4)		
TwoPlayerMessage	Q	unum_pos(4)	unum_pos(4)	
ThreePlayerMessage	Q	unum_pos(4)	unum_pos(4)	unum_pos(4)
SelfMessage	S	pos(3)	body(1)	stamina(1)
TeammateMessage	T	unum_pos(4)	body(1)	
OpponentMessage	O	unum_pos(4)	body(1)	
BallPlayerMessage	B	bpos_bvel(5)	unum_bpos(4)	pbody(1)

Figure2: Encoding Formats

When the player gets a bunch of hearing information, it looks for the corresponding decoding function from the head character, then call decoding program to obtain information represented by a string. Encoding methods for a single agent be divided into two as follows. First three characters and the other four to translate the same information into the following results:

Information	Encode str3	Encode str4	Decode str4
Position （20.9,20.9）	xEF	la78	Position(20.8611,20.9589)
Position （20.09,20.09）	xcW	Aa_7	Position(20.047,20.0274)
Position （20.009,20.009	xcW	za>7	Position(20.047,20.0274)

Figure3: Encoding Results of a Single Agent

For translation of speed location information, the maximum speed encoding is set to 3 because in the real environment the fastest on the pitch cannot exceed 3m / s, so the default setting is 3(if the speed exceed 3, we set it to 3), and encoded 5-character-long string.

In our case, experiment with different speeds shouted coding results under the same location are shown in figure 4 as follows:

Information	Encode	Decode
Position （20,20） Velocity （20,20）	Nv)nN	Position (19.9633,20.0274) Velocity (3,3)
Position （20,20） Velocity （10,10）	Nv)nN	Position (19.9633,20.0274) Velocity (3,3)
Position （20,20） Velocity （3， 3）	Nv)nN	Position (19.9633,20.0274) Velocity (3,3)
Position （20,20） Velocity （1.5， 1.5	Nv)9J	Position (19.9633,20.0274) Velocity (1.47619,1.47619)
Position （20,20） Velocity （1， 1）	Nv)5g	Position (19.9633,20.0274) Velocity (1,1)

Figure 4:Encoding and Decoding

For setting hearing information, in addition to considering whether it can be encoded and decoded, we also take into account impact of the error rate of the encoded information.

We have to determine the magnitude of the error and whether the error is an accidental one by many experiments, it can be seen from Figure 5, the error range is around 10^{-2} basic level :

Informatio	Encode	Decode
Player No.2， Position （20.0， 20.0）	elp*	Player No. 2 Position (19.9633,20.0274)
Player No.2， Position （20.1， 20.1）	em	Player No. 2 position (20.066,20.1605)
Player No.2， Position （20.2， 20.2）	eopE	Player No. 2 position (20.1686,20.1605)
Player No.2， Position （20.3， 20.3）	ep y	Player No. 2 position (20.2713,20.2935)
Player No.2， Position （20.4， 20.4）	erps	Player No. 2 position (20.3739,20.4266)
Player No.2， Position （20.5， 20.5）	es m	Player No. 2 position (20.4765,20.5597)
Player No.2, Position （20.6， 20.6）	eup0	Player No. 2 position (20.5792,20.5597)
Player No.2， Position （20.7， 20.7）	evZ*	Player No. 2 position (20.6818,20.6928)
Player No.2， Position （20.8， 20.8）	exo	Player No. 2 position (20.7845,20.8258)
Player No.2， Position （20.9， 20.9）	eyZU	Player No. 2 position (20.8871,20.9589)

Figure 5: Encoding and Decoding Information Table

3.Improvements of Decision-Making

(1) Application of Ngram Model in Team Decision Making

For each player, every decision made includes a number of factors under a certain probability. Let us take real soccer player for example, after getting the ball, he goes through careful consideration before deciding whether he will pass, move or shoot the ball, and to whom he will pass.

What's more, the player's decision will not always be the same even in the same environment. If there are multiple quite good choice, he can select one decision with certain probability. Also, there are many factors affecting decision-making process such as position, speed, angle and so on. And these factors can form successively relations which the former becomes the history information. The more history information is, the more restrictive to latter factors it will be. As an obvious example, the player standing on the forbidden zone has much more probability to shoot the ball than a player in the half-count. Also, a player will pass the ball to a closer teammate due to his lack of stamina, though a further one can get and shoot at once.

All kicking action have one thing in common, agents select a certain target point and perform the action. Suppose existing factor n, S1, S2, S3 ... Sn-1, Sn, then the probability of selecting one of the operation (including target point) can be expressed as follows:

$$P(\text{behavior}) = P(S1 | < BOS >) * P(S2 | S1) * P(S3 | S1S2) * P(S4 | S1S2S3) * ... * P(Sn | S1S2 \ldots Sn-1)$$

(<BOS> is the start tag)

So we want a player to do the action and the action with the highest probability will be executed. Thus how to obtain the probability of these actions will be the question. In fact, it can be considered the probability of selecting a target point. We have real data to get various statistical probability, thereby getting the probability of the player that best reflect the real situation and, ultimately, a model of the structure of FIG.This method has the advantage of not only being able to a certain extent, the reflection of the various team play and obtaining some new style of play through feedback learning, but also being possible to manually or automatic adjustment weights on some paths. After that, we can get a special playing method for fighting a certain team.

This method is not limited to the above applications, but also can be applied to many places, such as judgment of passing conditions, selection of defensive dynamic point moves, dribbling intensity adjustment, etc., these are the follow-up work for us to finish.

(2) Improvement of Pre-catch Moves

In previous research, we let the players to ensure the pre-catch point of getting ball and make them move there in order to get the passing ball more easily. But the effect is not very obvious due to the method may break the original attack formation. Target point of ball holder and catcher can be different because they have different visual information. Another reason is that their decision-making process varies much, leading to the result of different target point, even if under the same world model.

Now, we integrate these two modules so that the decision-making results of the ball-holder will directly pass to the catcher, therefore the collusion of point from the catcher is a subset of the ball-holder's, reducing the possibility of breaking the original attack formation and improve the accuracy of target-point prediction.

4.Conlusion

Our team currently mainly focus on the improvement of decision-making on visual, hearing. But there still remains some problems to be solved, for example, how to determine the range of activities in next few cycles and how to adjust assessment value to complete the conversion of personal decisions and multi-agent decisions. It is also known as making individual decisions when cannot achieve teammates' cooperation while making decisions with teammates if they can.

Since we are still new comers to Agent 2D, we have not involved ourselves into modifying the action code in underlying code yet. These are the follow-up works in our further research.

Reference

[1]方宝富. 机器人足球仿真[M]. 合肥工业大学出版社, 2011.

[2] Zhou H B, Zhou Y H, Zhu L J. Implementation and Comparison of Improving BP Neural Network Based on MATLAB[J]. Computing Technology & Automation, 2008.

[3] Gao Y, Liu J. A Modified Differential Evolution Algorithm and Its Application in the Training of BP Neural Network[C]// Advanced Intelligent Mechatronics, 2008. AIM 2008. IEEE/ASME International Conference on. IEEE, 2008:1373-1377.

[4] Cui G X, Liang L I, Wang K K, et al. Research and improvement on Apriori algorithm of association rule mining[J]. European Aquaculture Society Special Publication, 2010, 30(11):2952-2955.

[5] Liu G, Chen X, Wang X, et al. On Multi-Agent Collaborative Planning And Its Application In Robocup[J]. Intelligent Control & Automation Proceedings of World Congress on, 2000, 1:184 - 188.

[6] Mota L, Lau N, Reis L P. Co-ordination in RoboCup's 2D simulation league: Setplays as flexible, multi-robot plans[C]// 2010 IEEE Conf. on Robotics, Automation and Mechatronics, RAM 2010. 2010:362--367.

[7] 李龙. 基于价值的机器学习方法及其在 RoboCup 仿真 2D 中的应用[D]. 合肥工业大学, 2009.

[8] Li X, Chen W, Zhai Z, et al. The Application of Hybrid Distributed Reinforcement Learning Algorithm in RoboCup 2D Soccer Simulation System.[J]. Lecture Notes in Computer Science, 2008, 5314:651-659.

[9] Mota L, Lau N, Reis L P. Co-ordination in RoboCup's 2D simulation league: Setplays as flexible, multi-robot plans[C]// 2010 IEEE Conf. on Robotics, Automation and Mechatronics, RAM 2010. 2010:362--367.

[10] Pourmehr S, Dadkhah C. An Overview on Opponent Modeling in RoboCup Soccer Simulation 2D[C]// Robot Soccer World Cup XV. Springer-Verlag, 2012:402-414.

[11] Taheri S, Mammadov M. Learning the naive Bayes classifier with optimization models[J]. International Journal of Applied Mathematics & Computer Science, 2013, 23(4):787-795.

[12] Akiyama H, Nakashima T. HELIOS Base: An Open Source Package for the RoboCup Soccer 2D Simulation[M]// RoboCup 2013: Robot World Cup XVII. Springer Berlin Heidelberg, 2013:528-535.