The optimization and improvement of bridge game system

Yifan Song[1], Hongkun Qiu[2], Yajie Wang[3], Xiaodong Zheng[4]

1. School of Computer Science, Shenyang Aerospace University, Shenyang 110136,China
E-mail:15840836753@163.com

2. School of Computer Science, Shenyang Aerospace University, Shenyang 110136, China
E-mail:qiuhongkun_cn@sina.com

3. Engineering Training Center, Shenyang Aerospace University, Shenyang 110136, China
E-mail:qiuhongkun_cn@sina.com

4. School of Computer Science, Shenyang Aerospace University, Shenyang 110136,China
E-mail:1473146580@qq.com

Abstract: This article from the background of the bridge, improving the system analysis, experimental test three perspective, based on the idea of reinforcement learning, from two aspects of contract and scoring ability, through a large number of calculations and the original program and every time a new system of winning IMP value level to achieve the agreed order, accumulate experience, design a set of good rewards and punishment mechanism, greatly improve the efficiency of the bridge to play CARDS and intelligence, thus the bridge game system was optimized and improved. In addition, the traditional methods need to manually extract the features of poor expansibility, this paper combined with reinforcement learning algorithm, the ideas of game devised a new system, under the condition of different effective play the computer, the program has also reached a higher level of the game structure design, for the incomplete information game theory provides a reasonable method, application creates opportunities to people living in the future.

Key Words: Computer Games, Reinforcement learning, Incomplete information games, Bridge

1 Introduction

In recent years, with the public's praise of artificial intelligence and the promotion of computer technology, artificial intelligence has been applied in more and more fields, especially in the field of computer game. An endless variety of game software has greatly promoted the development of computer game and even artificial intelligence, and machine game, as one of the few items that have not been conquered by artificial intelligence, has been gradually perfected.[1-3]But with different types of chess, such as most chess programs with **incomplete information games**, there is room for improvement.

As one of the computer game cards, the traditional bidding stage of **bridge** is mostly Monte Carlo algorithm or database bidding, which has a large number of codes, which will lead to a long judgment time and some limitations.[3] Traditional call only according to the proportion of probability to estimate how much the right name is tasted, then after a lot of calculation or cumulative probability to get the accurate data, although this has advantages, but also a lot of limitations, the ability to design call this aspect, the machine machine almost to zero, this is a higher level of strategy, and everyone can fully embody the advantages between the game. In this context, we propose a bidding system based on **reinforcement learning**, which aims to make up for the deficiency of the traditional program through the intelligence and adjustability of reinforcement learning, multiple calculation and accumulation of experience value, and a set of well-adjusted "reward" and "punishment mechanism". The research results of this paper won the third prize in the 2020 National Computer Game Championship and the first prize in the 2020 Liaoning Computer Game Championship.

This work is supported by SAU College Student Innovation and Entrepreneurship Training Program X202010143070.

2 Background

MS Word Authors: please try to use the paragraph styles contained in this document.

2.1 Bridge

Bridge is a game played by four players using a standard deck of 52 cards, divided into four suits (spades, hearts, diamonds, and flowers), each of which contains thirteen cards. Players (usually referred to as north, south, east, and west) form opposing teams based on their positions. Bridge consists of two stages of play, called and played.[4]

2.2 Playing CARDS

In the game stage, a player puts a card in his hand face up in the middle of the table. The first card is played by the next house of the contracting party, also known as the first play. After the first play, the dummy side places all his cards face up on the table so that everyone can see the dummy's hand. During the game, the contracting party is responsible for playing for itself as well as for the dummy side. Each of the four directions plays a card, called a trick. Trick is a basic unit in bridge. The first player to play in each trick is called lead out. The next three players must play the same suit as lead out. The highest player of the cards wins the trick unless someone plays a trump. The player who wins this trick is the leader of the next trick. This cycle until the last card is played, and then the settlement is carried out. The winning or losing of the game is determined according to whether the contracting party has completed the number of tricks promised, and the score is calculated. In the process of playing cards, players can use a variety of strategies to win tricks, and the player's card power is closely related to whether he can flexibly use these skills.

2.3 Call

Bidding is the core of the whole bridge, and all of our playing process is based on the type of suit and trick number of the call, so a proper contract is the key to winning. At the same time, bidding is also very difficult, one is the problem of bridge itself, it belongs to the "incomplete information" game. Information gathering in bridge is much more difficult than in chess. Information about chess is not completely public.

Referring to the previous game algorithm, as shown in Fig.1, the traditional call can only approximate the appropriate call according to the proportion of probability, and obtain accurate data through multiple calculations or cumulative probability values. This has its advantages (we would expect the computer's odds to rise sharply if it had played cards instead of bidding), but its limitations are many. Therefore, the advantage of reinforcement learning lies in this, through a lot of calculation, accumulated experience value. Together with a well-tuned "reward" and "punishment" mechanism, the computer can make an accurate call in a wide variety of situations.[5]

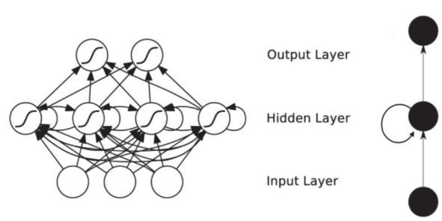

Fig. 1: Concept diagram of traditional game algorithms

2.4 The traditional Monte Carlo method

In old bridge games, when the first card is played by the player, also known as the first attacker, the hand is then revealed by the poker player, and the unsolved information is reduced to something in between, far less than the original three-way question. With the progress of the game, the possibility of the unsolved game is also slowly decreasing, which also represents the gradual visibility and clarity of the unknown information. For the card player, only to speculate the unknown card in which side hands, can take this as the root, the next step of thinking and judgment, to repeat. The application of Monte Carlo algorithm is formed through repeated speculation and judgment, including all situations and picking out the optimal line of playing cards.

When the user uses Monte Carlo algorithm to simulate playing cards, the design is carried out in accordance with the following process:

(1) Concretise the abstract bridge playing process into a concrete Monte Carlo probability problem, pick out the known hands and dummy hands, establish pseudo-random numbers through mathematical methods, and then create a bridge playing model of random dealing, so that the probability of unknown hand distribution is exactly equal to the probability of dealing in the game.

(2) In the process of repeated random licensing with computer code, a random sampling method is created for the random data, and then a large number of repeated experiments are conducted to judge and speculate the results of the experimental data.

(3) Review the analysis process and results, conduct backtracking pruning double paddle evaluation, get the optimal card route under the game situation, repeat this process for many times, synthesize the experimental results, and summarize the results with the most profits and the most times to the route selected by the program.

3 Bridge system under reinforcement learning

Reinforcement learning has been discussed in the fields of information theory, game theory and automatic control, etc., and has been used to explain the equilibrium state under the condition of bounded rationality, design the recommendation system and robot interaction system. Some complex reinforcement learning algorithms have a degree of general intelligence to solve complex problems, reaching human levels in Go and video games. In the future, deep reinforcement learning will be applied in many fields, which has unique research value and significance.

3.1 Review of reinforcement learning

The most important feature of deep learning is that deep neural networks can automatically find concise low-dimensional representations of high-dimensional data. References 5. Reinforcement learning model process belongs to Markov decision process, including state set, behavior set, state transition probability set and return set, as follows:

(1) State set S : the set of all states that can be reached by an agent.

(2) Behavior set A : the set of all behaviors available to learning agents.

(3) State transition probability set $P_{ss'}^{a}$: the probability of the action behavior a taken by the learning agent to move from state s to state s' .

(4) Return set $R_{ss'}^{a}$: the immediate return obtained after the probability $P_{ss'}^{a}$ changes from state s to state s' .

Based on the above, reinforcement learning model is formed. A strategy is adopted by the agent in the state to make the environment change, so as to obtain the corresponding reinforcement learning signal. As shown in Figure 2, when the agent adopts action A, the environmental state changes from to, and at the same time, the agent gets a feedback signal.[5]

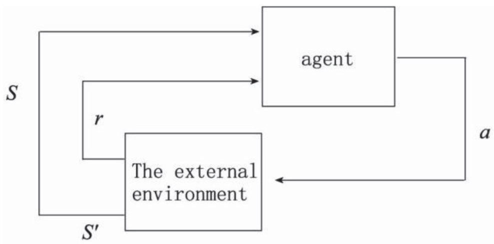

Fig. 2: Schematic diagram of reinforcement learning model

If it is applied to reinforcement learning, then reinforcement learning can be extended to the previously difficult problem, that is, to solve the problem of the environment with high dimensional state and action space, thus the birth of deep reinforcement learning.

3.2 Design of bridge game system based on reinforcement learning

Using the ideas of reinforcement learning, the team designed a new system for acquiring bridge cards, in which the initial state of $6.35*10^{11}$ plays were possible during the bidding process. In terms of the overall structure design of the software, the developer designed four functional blocks, which are log chess score, public module, application library, and platform interface packaging. The four functional blocks perform corresponding functions respectively, realize docking through interfaces, complete the conversion between functions, and interact with the code base, as shown in Figure 3.

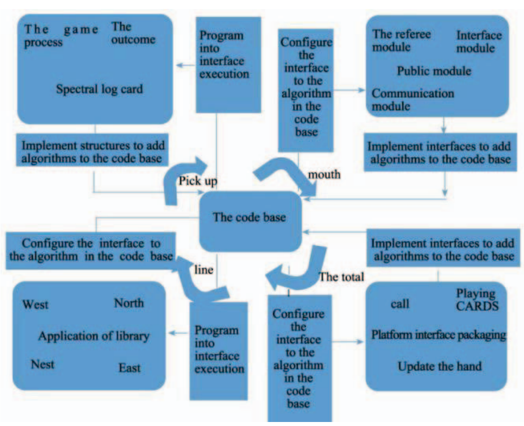

Fig. 3 :Schematic diagram of the overall design of the system

In this paper, the bidding procedure of the first opponent card points of the points of the conditions of classification, and each category of the situation of the card type classification, detailed into 14 categories, to determine the opening of the call. At this time, each card situation corresponds to a unique call, the partner can accurately locate the call by the call of the corresponding card type and card point range.

In the improved bridge system, instructions such as BRIDGEVER, INFO, DEAL, BID, CONTOVER, DUMMY, PLAY, GAMEOVER and ERROR can be issued. The player engine program can issue instructions such as BID, PLAY and OK. The communication flow chart between the system platform and the engine is shown in Figure 4.

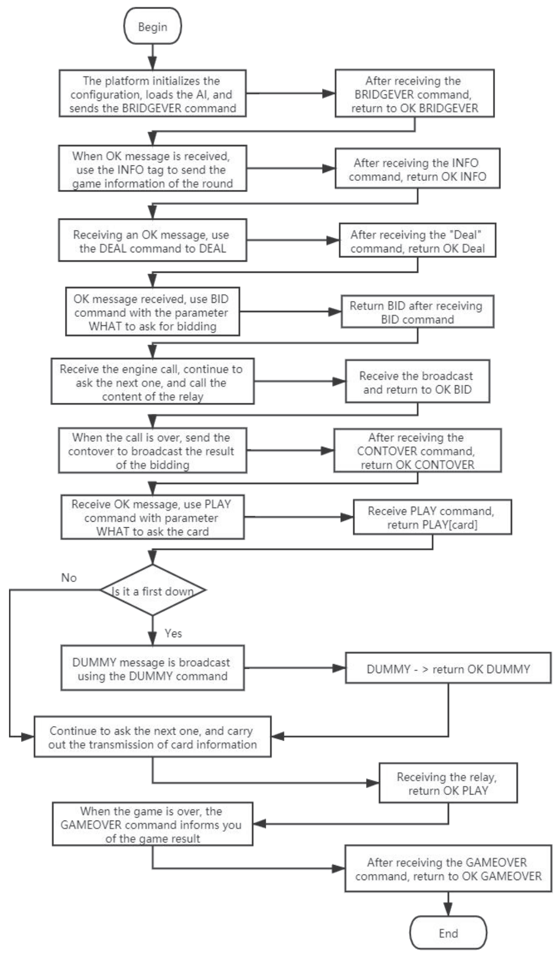

Fig. 4:Flow chart of interaction between bridge game system and engine

3.3 Example analysis of card type

In our improved program, the incomplete information game and Monte Carlo algorithm are combined with the bridge game program, so as to get a more accurate card play and improve the intelligence of the bridge program.

1) Run the system program of the improved East and West. After the game starts, the generated game is shown in Figure 5.

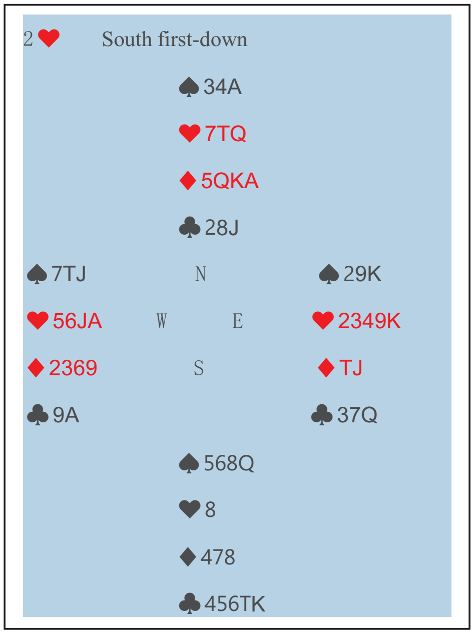

Fig.5:card type diagram

2) At the beginning of the game, the southern AI first attacked A ♣6, and it was the turn of the western AI to play (the improved program). At this time, the west had only ♣9 and ♣A in their stream. In the process of bidding, east, west two party programs call CARDS are to be carried out on the red peach, so western get Oriental AI AI on the plum blossom is not strong, our ♣9 will not win the trick, so with A respond, to seize the opportunity of the next round of the brand, to the east response, nature is the least of the plum blossom plum flower 3. (First round play C6, CA, C2, C3)

3) Halfway through the game, the Eastern and Western AI has locked the game. As shown in Figure 6, although the Oriental AI ♥2 and ♥3 cards are very small, but the north and south have no hearts, the two trick card is bound to get.

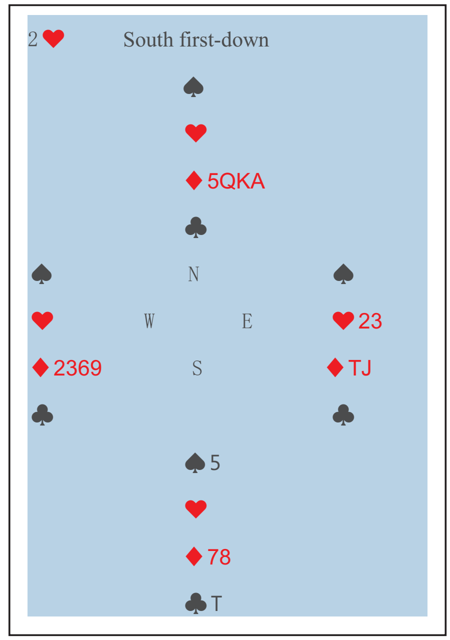

Fig.6:card pattern after half of the game

4) The game ends and East and West win. Through the above results, it is not difficult to see that the analysis ability of the optimized bridge game system for the game has been significantly improved. At the beginning of the game, the system will choose to win the trick as soon as possible, and then get the priority card opportunity, and then through the call again, further get new information, focus on the heart, as soon as possible to consume the opponent's heart, so that their side in the heart of the very small card, can also eat to win the trick, the victory of the game.

4 The experimental test

After a number of groups of experimental data tests, from the contract ability, scoring ability two aspects of comparison.

1) Comparison of contractual capacity

In order to judge whether the program based on reinforcement learning can improve the bidding ability to some extent, we conducted ten tests on the source program and the improved program under the same conditions. It's relatively easy for the source program to call up to levels 1 and 2, but it's a little more laboring for calls up to levels 3, 4 and above, as shown in Figure 7.

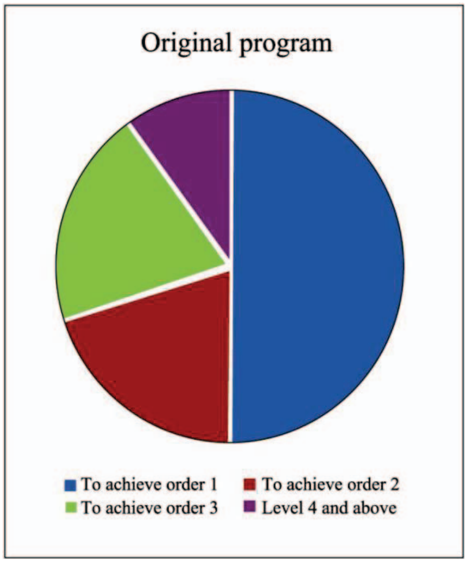

Fig. 7 :Schematic diagram of test results of source program

Figure 8 is the agreed pie chart that the source program can achieve in the actual test. It is not difficult to see that order 1 and order 2 occupy the vast majority of the area, while achieving order 4 is very little. Therefore, in this case, the scores we can get naturally tend to shrink sharply.

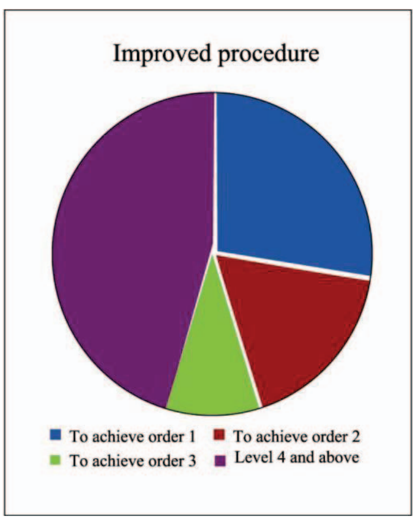

Fig. 8:Test diagram of the improved program

In Figure 4, we can see that the probability of a program reaching level 4 and above is significantly increased. Being more accurate in judging our own program to reach a given number of points, in practice, allows us to score more points in a round of matches, significantly reducing the likelihood that our team will win more games than our opponent will win a single game.

2) Comparison of scoring ability

The test data of the source program and the improved program were compared for 10 rounds, and the winning IMP value of each program was recorded to test the card playing ability of the improved program, as shown in Figure 9.

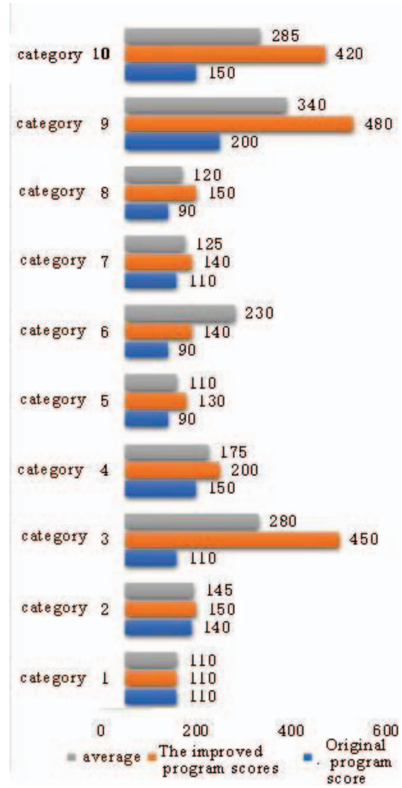

Fig.9: Bridge program ability comparison diagram

Contrast figure 9 zhongyuan program and improved, the optimized program in each round against the IMP values obtained is higher than the original program, and in some cases, the improved program can get more than three times the fraction of the original program, this can show in our improvement, process for match analysis more accurate, have stronger ability to play CARDS.

5 Conclusion

Through a large number of experiments, it is proved that the optimized and improved bridge game system has a great leap forward in both feasibility and intelligent program compared with the traditional game system. The improved program is more comprehensive in consideration and has more prominent advantages.

Bridge is a game of incomplete information, and computers are good at analyzing through lots of calculations. In the process of judgment chess program, this article optimized and improved game system with a complete set of rewards and punishment mechanism, the precision of calculation by the program, to be able to continue to accumulate experience value, can bring a better analysis result to the program, make the program more intelligent, play more efficient and make the computer in the face of a variety of circumstances, make an accurate bid.

References

[1] J. C. Halsoni, Yi Xianrong, *Incomplete Information Game*, Economic Translation,1995.

[2] Liu Quan, Zhai Jianwei, Zhang Zongchang, Zhong Shan, Zhou Qian, Zhang Peng, Xu Jin.*A review of deep reinforcement learning*[J]. Chinese Journal of Computers, 2018, 41(01): 1-27.

[3] Wang Yulou. *Current situation and future development trend of artificial intelligence*[J]. Science and Technology Outlook, 2016,26(22):299.

[4] Deepmind.*Human-level control through deep reinforcement learning*[P].Nature.

[5] Wang Pengcheng. *Research on Machine Game with Incomplete Information Based on Deep Reinforcement Learning*[J], TP391.4/621.3.

[6] Yu Jianpeng, Gui Jianping. *Review of Reinforcement Learning*[J]. Computer Knowledge and Technology, 2008, 002(015): 1094-1095.

[7] Liu Zhong, Li Haihong. *Research on Reinforcement Learning Algorithm*[J]. Computer Engineering and Design, 2008, 029(022): 5805-5809.

[8] Mo Jianwen, Lin Shimin. *Design and Implementation of Intelligent Game Program Based on TD-Reinforcement Learning*[J]. Journal of Computer Applications, 2004, 024(0Z1): 287-288.

[9] Wang Yun, Han Wei. *Multi-agent Reinforcement Learning for Symmetrical Coordinated Game Problem*[J]. Computer Engineering and Applications, 2008, 044(036): 230-248.

[10] Liu Weibing, WANG Xianjia. *A multi-agent reinforcement learning model for evolutionary games*[J]. Systems Engineering-Theory & Practice, 2009, 029(003): 28-33.

[11] Liu Bingyan, Ye Xiongbing, Gao Yong, Wang Xinbo, Ni Lei. *Strategy solution of non-cooperative target pursuing game based on branch deep reinforcement learning*[J], Acta Aeronautica et Astronautica Sinica, 2020,41 (10).

[12] Cai Wei, Bai Guangwei, Shen Hang, Cheng Zhaowei, Zhang Huili. *A Win-Win Game Based on Reinforcement Learning in Mobile Group Intelligence Perception* [J]. Computer Science, 2019, 47(10).

[13] Xiang Yutao, Zhu Daoyi, Wang Zhongtao, Dong Yu. *Research on the algorithm of Einstein chess based on reinforcement learning*[J], Computer Knowledge and Technology, 20, 16(22).

[14] Cui Wenhua, Li Dong, Tang Yubo, Liu Shaojun. *Deep Reinforcement Learning Based Decisionmaking Method Framework of Military Game Dedication*[J], Defense Science and Technology, 20, 41(02).

[15] Wang Renquan, Ding Meng, Li Shuqin, Shi Luying, Qi Yizhong, et al. *Surakarta Game Algorithm Based on Reinforcement Learning*[J], Intelligent Computers and Applications, 20, 10(04).