Self-Optimizing Evaluation Function for Chinese-Chess

Xiangran Du[1,2], Min Zahang[1] and Xizhao Wang[2]

[1]*Tianjin Maritime College, Tianjin 300350, China*
[2]*Key Lab. of Machine Learning and Computational Intelligence, College of Mathematics and Computer Science, Hebei University, Baoding 071002, China*
duxiangran1226@126.com

Abstract

Computer game is a vibrant research area in artificial intelligence. Chinese chess game is an important part of computer game and it has become an important study area after chess game had reached its culmination when Deep Blue and its successors beat Kasparov. Some achievements acquired in Chinese chess game have applied into fields of medicine, economics and military. This paper presented a new method of optimizing evaluation function in Chinese-chess programming by particle swarm optimization. The process of training evaluation function is to automatically adjust these parameters in the evaluation function by self-optimizing method accomplished through competition, which is a Chinese-chess system plays against itself with different evaluation functions. The results show that the particle swarm optimization is successfully applied to optimize the evaluation function in Chinese chess and the performance of the presented program is effectively improved after many trains. We also examined the importance of the place control in the evaluation function by the comparison the optimizing results with and without the control of the place and showed the comparison result.

Keywords: Artificial intelligence; Chinese-chess; Particle swarm optimization; Self-learning; Evaluation function

1. Introduction

Artificial intelligence is a method of studying intelligent computer, which makes the machine have abilities of human thinking and judgment. Machine Game is an important field of research in artificial intelligence. Various searching algorithms, pattern recognition, and intelligent methods came from machine game research are widely applied to many fields and improved the development of these fields. Likewise the advance of these fields greatly promoted the development of artificial intelligence.

Computer chess is a two-player, zero-sum game with complete information, which has developed more than 70 years [1]. The famous achievements are chess champion Kasparov was defeated by supercomputer "Deep Blue" and computer Hydra easily defeated chess master Michael Adams [2]. Chinese chess is one of the most popular board games worldwide, being played by approximately 1.5 billion people in china and wherever Chinese have settled. Chinese chess program has became the next challenge for many artificial intelligence masters, because the complexity of Chinese-chess is between that of chess and Shogi (the complexity of chess is 123 (Allis(1994), the complexity of Shogi was estimated Bylida, Sakuta and Rollason (2002)) and there is a long distance between Chinese chess program and masters of Chinese-chess [3].

Currently, Chinese chess game research focuses on the evaluation function optimization and game system intelligences. Northeastern University's Wang Jiao used genetic algorithm to optimize evaluation function [4]. Changsha University's Fu Qiang successfully achieved the intelligent Chinese chess by enhance learning [5]. Yifei Wang graduated from Harbin Engineering University realized the method of combining neural network with TD into the Chinese chess program [6]. TD (λ) algorithm is introduced into Chinese chess Computer Game by Hebei University Professor Xizhao Wang and Yulin He [7-8].

The paper presented an intelligent Chinese chess system developed by a dynamic self-learning method. The particle swarm optimization is applied to dynamically optimize parameters of the evaluation function in this system according to the character of Chinese-chess. After a lot of trainings, Chinese-chess program has realized the intelligence. The result of the experiment shows that it is possible to optimize the parameters in the evaluation function and the power of Chinese-chess is improved effectively. We also make an examination about the effect of the evaluation function with and without the place control in order to simplify the evaluation function under the condition of ensuring efficiency and accuracy.

This paper is organized as follows. In Section 2, we give a brief review of the evaluation function [9]. Section 3 provides the process of training the evaluation function by the particle swarm optimization. In Section 4, the experimental results showing the performance of the optimized Chinese-chess program and the comparison results among different training stages are presented and discussed. Section 5 concludes this paper and points out the direction for future research.

2. Evaluation Function

Computer Chinese-chess program includes mainly five parts: opening play, moving generation, searching algorithms, evaluation function, searching and playing the endgame (see Figure 1). Among of these parts, evaluation function assessed the current state of chess is most important and reflects intelligence. It is also the part where Chinese chess masters are superior to computer. If the evaluation function made an accurate assessment, the searching algorithm can quickly find a right next move that makes the state of the chess favorable to me, otherwise the searching algorithm only just finds wrong move.
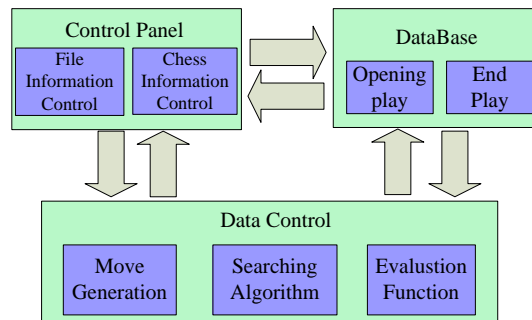

Figure 1. The Structure of Chinese Chess Program

The evaluation function in Chinese chess is the difference between the current assessed values of the two sides.

$$\text{Eval}(x) = \begin{cases} \text{Eval}(Red) - \text{Eval}(Black) & x = Red \\ \text{Eval}(Black) - \text{Eval}(Red) & x = Black \end{cases} \qquad (1)$$

If the x indicates Red and Eval(x) is greater than zero, it shows the current situation is advantageous for Red. Otherwise the situation is disadvantageous. Likewise, if the x represents Black.

The evaluation of a position in Chinese chess has approximately six elements [3]: (1) the strength of the pieces in play, (2) the positions' value, (3) the control of the places, (4) the pieces' flexibility, (5) the threat between pieces and the protection of pieces from threat, and (6) the piece features or dynamic adjustment according to the situation. Details of those elements are discussed blow.

1) The strength of the pieces

The strength of pieces represents the important level for a piece. According to Chinese chess rules, every piece having unique move means that its effect and important level are different in various situations. The common values for different pieces are shown in Table 1 given by Chinese chess masters. The strength of pieces can be computed relative to the value and dynamic weight for each piece on the board, it is form of

$$\text{Strength} = \sum_{i=1}^{16} w_i \cdot Value_i \qquad (2)$$

where wi the weight expresses the important level for the number i piece on a board. Each piece has wi that is changing based on the current situation for each piece.

Table 1. The Value of Each Piece

Piece:	King	Assistant	Elephant	Rook	Horse	Cannon	Pawn
Value:	10000	110	110	300	600	300	70

2) The value of the position

The position's value indicates the value of the pieces at a different place on the board. The place that can be occupied to threaten the enemy piece especially having a higher value is the best place. Chinese chess programs are usually equipped with a table stored the estimated value of possible position for each piece. Figure 2 shows the position value table of the Pawn on the board.

For example, the value of the Pawn is different with the changing position. The power of the Pawn is proportionate to the distance with the enemy King.

```
0    3    6    9    12   9    6    3    0
18   36   56   80   120  80   56   36   18
14   26   42   60   80   60   42   26   14
10   20   30   34   40   34   30   20   10
6    12   18   18   20   18   18   12   6
2    0    8    0    8    0    8    0    2
0    0    -2   0    4    0    -2   0    0
0    0    0    0    0    0    0    0    0

0    0    0    0    0    0    0    0    0
0    0    0    0    0    0    0    0    0
```

Figure 2. Position Values of a Pawn Used by ELP

3) The flexibility of the pieces

If a piece could not move freely under the rule of Chinese chess, its attacking or defending power is restricted. In other word, the more flexible the piece is, the more attacking or defending power it has. For example, if a piece stands on the position preventing two Elephants from protecting each other, the defending power of the Elephants is largely decreased. However estimating the flexibility of the pieces cost too much, Chinese chess program usually considers the flexibility of some important pieces.

Table 2. The Flexibility of Each Piece

Piece:	Assistant	Elephant	Rook	Horse	Cannon	Pawn
Value:	7	3	1	7	13	15

4) The threat between pieces and the protection of pieces from threat

The threat between pieces and the protection of pieces from threat in evaluation function is fully considered that the cooperation among the pieces in Chinese chess is important and makes these pieces form an interrelated whole. When a piece on the board is threatened, we could move the piece to another place or use a piece to protect it. The security of a piece is determined by the number and kind of protectors and threatening enemies.

5) The control of the place

The control of the place on the board usually indicates that a place on the next move of a piece is controlled by it. If a place on the next move of a piece, it is usually considered the place is controlled by the piece. If a place is control of the both sides, a simple way to estimate which side controlling of the place is using the number and kind of pieces of both sides and sequence. The control of the place is valuable when Chinese chess is on opening and middle game, but the significant value is slightly decreased with the development of the game. Because the number of the place controlled by a piece is largely increased when the game is closing to end, some Chinese chess masters suggest that some important place, especially closing to the King, is seriously considered on the evaluation function.

6) The piece features or dynamic adjustment according to the situation

The piece features indicate the coordination among these pieces, especially for Cannon, Horse and Rook. The coordination is not simple sum of these pieces and makes these pieces cooperate as a whole.

Figure 3. Horse-cannon Shows the Power of Cooperative Relationship

The value of the sum of one horse and cannon is lower than the value of two cannons from Figure 1. Two cannons have the value of 600 and one horse and one cannon have value of 580. Two cannons have lower value than one cannon and one horse in fact, because one cannon and one horse have more tactical skills than two cannons. One of the tactical skills is called cannon behind a horse. The main tactics applied in Chinese chess include catching two pieces, pin down, sacrificing or exchanging pieces. These tactics should receive more value and make Chinese chess program stronger.

3. The Optimization of the Evaluation Function

Particle Swarm Optimization (PSO) is proposed by doctor Kenndy and Eberhart, it realized the analogues of bird flocks searching for corns to produce computational intelligence [9]. The advantages of the PSO are a simple realization, a small number of the parameters and higher speed. The results of the PSO have applied into more and more fields in science. The PSO has been received more attention and became a new hotspot in optimization algorithms after GA, ACS, *etc.*, [10].

3.1 Particle Swarm Optimization

In PSO each particle includes the location information and the velocity information are respectively presented by N dimension $L = (l_1, l_2, \cdots, l_N)$ and $S = (s_1, s_2, \cdots, s_N)$. The location information indicates these optimized parameters that are feasible solutions in solution space, and the velocity information for each particle is the optimized speed dynamically regulated by studying the surrounding environments. Each particle is moved through the search space by combing some information of the history of its own current and best location with those of one or more members of the swarm, with some velocity information and random perturbation. The next iteration occurs after all members in swarm have been moved. Each particle was programmed to update its velocity and velocity information in terms of the equation (1) and (2):

$$S(t+1) = S(t) + C_1 * rand * (LBest - S(t)) + C_2 * rand * (GLBest - S(t)) \tag{3}$$

$$L(t+1) = L(t) + S(t+1) \tag{4}$$

where w is termed the "inertia weight" and is set a dynamic number from a relatively high value, to a much lower value. The parameter w with a high value (*e.g.*, 0.9) means that particles take place global search, but when w is given a low value (*e.g.*, 0.4) the particles in swarm assemble toward local optima. The updating formula of the parameter w is the following:

$$w_t = \frac{(w_{fin} - w_{ini}) \times (T_{max} - t)}{T_{max} + w_{ini}} \tag{5}$$

where T_{max} is the maximum iteration, w_{ini} and w_{max} is respectively the initial value and the maximum value of the parameter w.

The parameters C1 and C2 in (1) are often called acceleration coefficients and determine the magnitude of the random forces in the direction of local optimum (called LBest) and global optimum (called GBest). The parameter rand is a random number from 0 to 1. The value C1=C2=2, almost ubiquitously adopted in PSO research.

3.2 Optimizing Evaluation Function

Each particle in particle swarm represents a set of parameters needing to be optimized when the PSO optimizes the evaluation function. The fitness function of the particle is the probability of winning in matches among particles, which is the ratio of the winning number to the total number of game. The matches among different particles are the Chinese chess games applied different evaluation functions, which are measured by the chance of victory. The one of the purpose for optimization is to make the evaluation function estimate the advantage and disadvantage state of the current position more exact. Another one is to greatly coordinate with the searching algorithm applied in Chinese chess and finally promote the power of Chinese chess program. The highest winning percentage in particle swarm is considered as the global optimum and the local optimum of a particle is regarded as the best value in the process of updating itself.

Every particle competes with the others in the particle swarm twice by exchanging the sequence of chess. The evaluation function is composed of the location information for each particle. Chinese chess programs applied different evaluation functions circularly game with each other, every two opponents compete twice with different sequence of chess. The result of the game is recorded in database.

The location information of the particle with the best winning number in database is the global optimum at the present particle swarm. The distance of these particles is computed according to the location information and the nearest particle from the updating particle is selected on the basis of distance. The local optimum for every particle is obtained by averaging location information of particles that the winning number is no less than the updating particle's from the nearest particle.

The velocity and location information are respectively updated by the formula (1) and (2). The global optimum is eventually regarded as the best parameters for evaluation function when the global optimum is a sufficiently good fitness or a maximum number of iteration is reached.

4. Experimental Results and Discussion

The presented particle swarm optimization, training method and local or global optimum have been implemented. Our goals are first, comparing results of optimizing parameters of the evaluation function at the different generations and pick out the one which performs best when playing against the other particles, and second, utilizing the best program as an expert to play against the unoptimized Chinese chess program, and the last, examining the feasibility of the evaluation function without the control of the place.

4.1 Particle Swarm Optimization

In first training stage, the particle swarm includes 20 particles and every particle is composed of 90 parameters, the parameter C1 and C2 are set 2 respectively and rand is initially 0.5. Every particle competes against the others in the swarm 38 games in every generation with different sequences. There are 380 games in a generation for all particles. After 400 generations and 152000 games, a new evaluation function is got coming from the global optimum with the best fitness in particle swarm. We use Chinese chess programs with 40 different evaluation functions that come from 40 particles having the best fitness every ten (decade) generations in the optimizing process to compete each other at different sequence for testing the optimizing performance. The result of competition is shown in Figure 4.
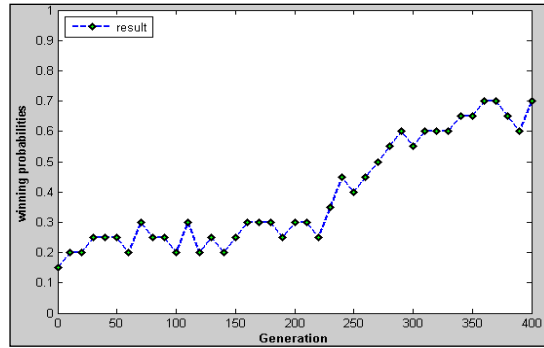
Figure 4. The Result of Games in Various Generations

In Figure 4, we notice that the power of the new evaluation function has advanced clearly although there is some zigzag phenomenon. We can also see that the learning efficiency of the particles is low at the beginning phase in the learning process and the results of the competition are between 20 percent and 30 percent. As the training proceeds, the performance of Chinese chess progressed clearly after 220 generations and the winning probability increased efficiently from 25 percent to 70 percent. The fastest escalating trend in the first learning stage is from 220 generation to 290 generation and after 300 generation, the winning probability is stable by and large at 65%. However, in the training process, we found that the optimized evaluation function can't judge clearly some complex situations and make the Chinese chess repeat some moves and sometimes infinite loops.

Based on weights obtained above besides parameter $C1 = C2 = 1.5$ and rand = 0.3 in the second training stage and a second 1000 generations are performed. Figure 5 shows that the process of training the evaluation functions and signs the global optimums for every ten generation. We noted that the performance advanced fast at the begin of the learning process and that the speed of learning knowledge gradually decreases as the training carries on from Figure 5. The global optimum has been largely evaluated from 10% to 70%. The power of the evaluation function increases in a zigzag until 1000 generation. After 380000 training games, the power of Chinese chess has been strengthen clearly. The moves evaluated by the new evaluation function are much intelligent than before, even though occasionally some moves might be badly computed.
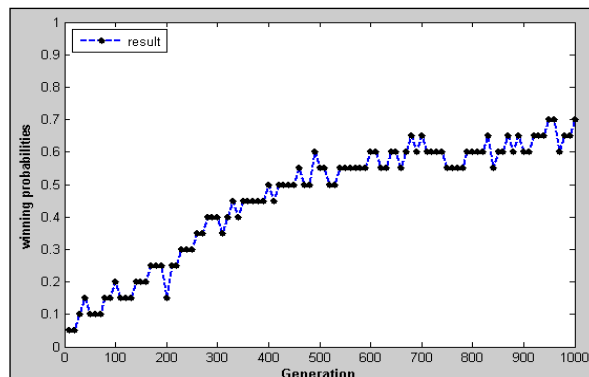

Figure 5. The Global Optimums for Every Ten Generation

In last training stage, the evaluation function is continuously trained by the PSO for another 1000 games with parameter $C1 = C2 = 1$ and rand = 0.1 and the results achieved are

little better than those before. According to the experiment results, we found that the presented method can optimize the evaluation function but there are some severe defensive problems in endgame and the evaluation function occasionally makes the game into the situation of the endless loop or insignificantly repeats attack on the opponent's king, especially in the endgame.

4.2 Comparison of the Optimizing Results

For showing the effect of optimization, we let Chinese chess programming optimized by the last generation's evaluation function and the unoptimized one plays against each other. As a result, the optimizing programming won 22 games and got 8 draws without loses. The result shows that the particle swarm optimization is useful to optimizing the evaluation function. The last step is to examine the degree of accuracy of the evaluation function without the place control is whether to lose.

The control of the place on the board usually indicates that a place on the next move of a piece is controlled by it. Some Chinese chess masters consider that the place control is less important than the others in the evaluation function and the importance of the place control is limited. The speed of the evaluation function without thinking the control of the place can be promoted and the time of searing algorithm is advanced, the power of Chinese-chess programming is whether to increase or to decrease? The result of the next experimentation is to answer this problem to some extent.

We repeat the above train process under the condition of without the control of the place in the evaluation function. After 2400 generations and 912000 training games, an optimizing evaluation function is realized. For examining the accuracy of the evaluation function without the place control, we make a result comparison in the last training stage between twice examinations. The concrete method is to select respectively 100 particles with best fitness from last 1000 generations in two training experiments for every ten generation, and these 200 particles competes each other with different sequences. After 39800 games, the result of competition is shown in Figure 6. The red line expresses the competing results of the particles with the evaluation function having no the control of the place and the results with the place control is indicated by the blue line from Figure 6.
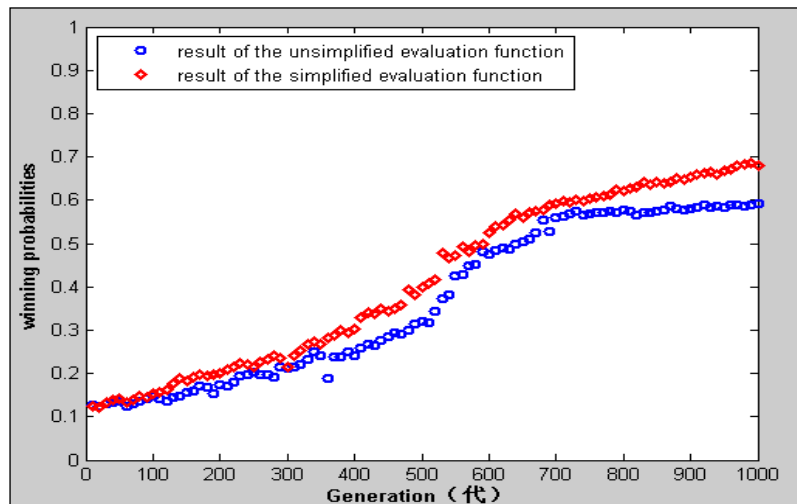

Figure 6. Comparison the Wvaluation Function Between the Unsimplified and the Without

From the results of the comparison between the evaluation function with and without the control of the place, we can notice that the performance of the simplified evaluation function is better than the un-simplified evaluation function especially for the generations from 300 to 400. In the first 100 generations, the winning probability is almost same for the different evaluation function. Then the winning probability increases rapidly for the particles from 200 to 500 generation and the winning growth of the particles with the simplified evaluation function is faster than the unsimplified ones. From 700 to 1000 generations, the performance of the optimizing particles is clearly obvious and the advantage of the particle without thinking the control of the place is enlarged. The reason might be that the simplified evaluation function can save time and cost more time on the searching algorithm that can search deeper from the game tree. Therefore the evaluation function without considering the place control is suitable for the Chinese-chess programming.

5. Conclusions and Future Direction

This paper puts forward an intellectualized method that is able to strengthen the evaluation function in Chinese chess without manual intervention. The method is carried out by optimizing the evaluation function with the particle swarm optimization. The experiment shows that the power of Chinese chess has been clearly advanced using the method presented in the paper. The optimized Chinese chess can easily overcome the former one. However, the optimizing method only applied to a class of Chinese chess. We have been carrying on our research on the influences of reducing the parameters in the evaluation function and transforming the construction method of the local optimum in particle swarm optimization.

## Acknowledgement

References

[1] Samuel, "Some Studies in Machine Learning Using the Game of Checkers", IBM Journal on Research and Development, USA, **(1959)**, pp. 210-229.
[2] N. L. David, "Computer Games", New York: Springer New York Inc., **(1988)**, pp. 335-365.
[3] S. J. Yen, J. C. Chen and T. N. Yang, "Computer Chinese Chess", ICGA Journal, vol. 3, **(2004)**.
[4] W. Jiao, W. T. Shi, Y. H. Luo and X. H Xu, "Implement of Adaptive Genetic Algorithm of Evaluation Function in Chinese Chess Computer Game System", Journal of Northeastern University (Natural Science), Shen Yang, vol. 26, **(2005)**, pp. 949-952.
[5] F. Qiang and H. W. Chen, "A Design and Implement of Chinese Chess game, Journal of Changsha University of Science and Technology (Natural Science)", vol. 4, no. 4, **(2007)**, pp. 73-78.
[6] D. B. Zhao, Z. Zhang and Y. J. Dai, "Self-teaching Adaptive Dynamic Programming for Gomoku, Neurocomputing", Elsevier, **(2012)**, pp. 23-29.
[7] T. B. Trinh, Anwer and S. Bashi, "Temporal Difference Learning In Chinese Chess", Department of Electrical Engineering University of New Orleans, **(1996)**, pp. 7-11.
[8] J. Baxter, A. Tridgell and L. Weaver, "Learning To Play Chess Using Temporal Differences", Machine Learning, **(2001)**, pp. 243-263.
[9] J. Kennedy and R. C. Eberhart, "Particle Swarm Optimization. Proceedings of the IEEE International Conference on Neural Networks", Piscataway, NJ: IEEE Service Center, **(1995)** April.
[10] P. N. Suganthan, "Particle Swarm Optimizer with Neighborhood Operator", Proceedings of the Congress on Evolutionary Computation, Washington DC, USA, **(1999)** May.

[11] Y. Shi and R. C. Eberhart, "Parameter Selection in Particle Swarm Optimization", Evolutionary Programming VII. Lecture Notes in Computer Science, Springer, **(1998)**, pp. 591-600.
[12] A. Chatterjee and P. Siarry, "Nonlinear Inertia Weight Variation for Dynamic Adaptation in Particle Swarm Optimization, Computer & Operations Research", Elsevier, **(2006)**, pp. 859-871.
[13] M. A. Wiering, "Self-play and Using an Expert to Learn to Play Backgammon with Temporal Difference Learning", J. Intell. Learn. Syst. Appl., vol. 2, **(2010)**, pp. 57–68.
[14] D. B. Zhao, J. Q. Yi and D. R. Liu, "Particle Swarm Optimized Adaptive Dynamic Programming", Proceedings of the 2007 IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning, Honolulu, Hawaiian Islands, **(2007)** April 1-5.
[15] A. G. Barto, R. S. Sutton and C. W. Anderson, "Neuron Like Adaptive Elements that can Solve Difficult Learning Control Problems", IEEETrans. Syst. ManCybern, vol. 13, **(1983)**, pp. 834–847.

Authors

Xiangran Du received the B.S. degree in Heilongjiang University of Science and Technology, China in 2006, and M.S. degree in College of Mathematics and Computer Science, Hebei University. He works at Tianjin Maritime College and Key Lab. of Machine Learning and Computational Intelligence. His main research interests include the application of the particle swarm optimization and neural network to Chinese chess system and reinforcement learning to traffic control in urban areas.
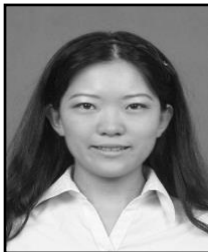
Min Zhang received the B.S. degree from Dalian Nationalities University, China in 2005, and M.S. degree in Transportation planning and management from Dalian Maritime University, China in 2007. She has published one book and 15 journal papers. Her current research interests include the assessment of port informatization and adaptive dynamic programming in port informatizational system.