

Opponent's Style Modeling Based on Situations for Bayesian Poker

Ruobing Li¹, Wenkai Li¹, Lin Shang¹, Yang Gao¹, and Mengjie Zhang²

¹ State Key Laboratory for Novel Software Technology,
Nanjing University, China

{l.ruobing,easerene}@gmail.com, {shanglin,gaoy}@nju.edu.cn

² School of Engineering and Computer Science,
Victoria University of Wellington, New Zealand
Mengjie.Zhang@ecs.vuw.ac.nz

Abstract. In a real poker game, one player can take actions of different styles in different situations. In this paper, a novel method is proposed to quantify and model the opponent's style in corresponding situation of a hand. Based on the proposed representation of Action Pair, the value of the style can be calculated and stored as "experience". When making a decision, the specific style will be obtained from the "experience". The style and the observable information will be used to estimate the value of the opponent's hand. In experiments, the obtained "experience" validates the correctness of our assumption that a player does not show an invariable style in all situations. The experimental results show that the agent player using our method can predict the value of the opponent's hand and earn more money in fixed hands comparing with the original agent.

Keywords: Style Modeling, Bayesian Poker, Hold'em Poker.

1 Introduction

Poker is essentially a process of reasoning under uncertainty. It provides a good platform for testing new ideas and approaches to dealing with uncertainty. In a poker game, the opponent's cards are unobservable. Thus, the value of the opponent's hand is unknown. The information which can be obtained from the opponent is only his action. If the opponent's style or other useful features can be effectively modeled, we will have a better understanding of the opponent's action. If we want to win more money in a poker game, we have to model the opponent's style, habits, characters or other features. These hidden information can help us better grasp the opponent's motivation. The opponent may show an obvious tendency in a number of hands. However, specific to each decision, a good player is likely to be inconsistent with his style. For example, a player may bluff in order to disrupt his opponent's thinking and make himself unpredictable. Thus, there is a strong need to adapt to the opponent well. Opponent modeling is then considered important in poker [1,2,3]. Researchers have been committed to find a reasonable and effective model to adapt poker agents to its opponents.

Opponent modeling in poker is mainly led by two ideas in current research [4]. In essential, both ideas are based on analyzing the opponent's action.

The first idea aims to classify the opponent into one of styles. As we know from [5], there are four recognized styles which are conservative-loose, conservative-tight, aggressive-loose and aggressive-tight in poker.

- **Conservative-Loose.** Players who over value their hands, but rarely raise, being fearful of large pots.
- **Conservative-Tight.** Players who play few hands, usually with a high probability of winning, but rarely raise, being fearful of large pots.
- **Aggressive-Loose.** Players who over value their hands, and who often raise to increase potential winnings.
- **Aggressive-Tight.** Players who play few hands, usually with a high probability of winning, and often raise to increase potential winnings.

Baker defines the opponents as above four distinctive styles and develops customized tactics to defeat each of the styles [6]. By analyzing the opponent's past actions, he classifies the opponent's style into one of four styles. After that, he chooses the most threatening opponent to play against and adopt that corresponding tactic when facing all the opponents of other styles. However, a player with a determinate style may act with an inconsistent style in a certain situation. Then the single style will be ineffective.

The second idea aims to model the distribution over the opponent's future actions or current state. The predictive model can be used to estimate hidden state in situations where the opponent has access to information only available to the observer through his actions.

Some researchers count the times of each player folded, called, or raised [7,8]. Then the distribution of the opponent's actions is constructed. Given the observed action and the board cards, the obtained distribution can be used to estimate the value of the opponent's hand.

Researchers have also studied on modeling the opponent's next actions. Billings [9,10] creates a statistics-based opponent modeling system. In this system, a probability distribution of actions is generated with each of the opponent's action. The opponent's next action is then predicted. The question is that the predictions of the next action are not so accurate in particular situations.

Most existing contributions of above two ideas are giving a representation of style and revealing the relation between the action and style. However, it is insufficient to represent the opponent's styles in all situations. There has been no work conducted to consider the diverse styles showed by opponents in various situations of a hand. In this paper, we work on 2 Players Limit Texas Hold'Em. Bets are predefined at a fixed size in this form. Therefore, we can be focused on opponent's style modeling in this form of Texas Hold'Em Poker. We propose a novel method aiming to model the opponents style based on situations. We first propose Action Pair to describe the relation between actions and styles. Based on Action Pair, we will define some formulas to quantify and determine the opponent's style in various situations of a hand. The style model will then

be used to estimate the opponent's hidden information, which plays the decisive role in playing poker. Specifically, we intend to investigate:

- properly provide an objective and exact perception of the value of a hand,
- accurately model the style of the opponent in all situations,
- effectively use the modeled style of opponent.

The remainder of the paper is organized as follows. Section 2 introduces the new method. Section 3 describes the experiments and results with discussions. Section 4 presents the conclusions and future work.

2 The New Method

Understanding the opponent is the key to win in poker game. In order to make our agent better adapt to the opponent, we propose a method to quantify and determine the opponent's style based on situations. With the modeled style, the value of the opponent's full hand can be estimated.

The flow chart of the method is shown in Fig. 1, which has three modules.

The first module is *Inference*. Bayesian network [11,12] is used to model the probability distribution of win, lose and tie in this module. The distribution will provide us an objective and exact understanding of a hand. Korb has done much work in this field [9].

The second module is *Learning*. In this module, the opponent's style will be modeled for each situation of a hand.

The third module is *Tactic*. The modeled style will be used in the *Learning* module to help make decisions.

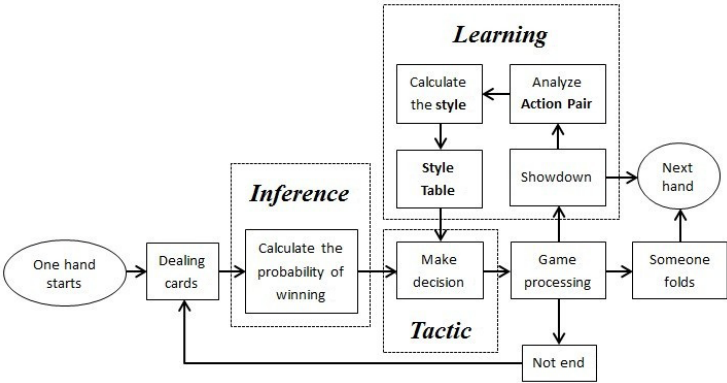


Fig. 1. The flow chart of the method

When a human player is dealing with a hand, he probably has two problems to consider: One is what the real value of a hand is, at least a relatively accurate estimate given observable information. The other problem is how to understand

the opponent's last action given the history of the opponent's actions. When a hand is over, he will start thinking again how the opponent plays. Reviewing last hand will provide experience for his future playing. The method we propose in this paper will follow such process.

In our method, Bayesian network is used to construct the probability distribution of win, lose and tie of a hand. After each showdown hand, Action Pair and the probability of winning are employed to calculate the style based on situations. The new style value is then merged with the historical data of the style. Style functions are designed to accomplish the style calculation. Then the decisions are made according to the opponent's style value. The details will be presented in the following subsections.

2.1 The Probability Distribution of Win, Lose and Tie

Firstly, we will model the probability distribution of win, lose and tie given the current information of a hand. With the observable information such as round, board cards and our hole cards, although it is possible to compute the probability distribution exactly under the assumption of a uniform probability over the cards on a fair deck, the high computational cost is still unbearable [13]. So the distribution can only be estimated. The distribution in this paper is inferred by a Bayesian network. It is the *Inference* module in Fig. 2.

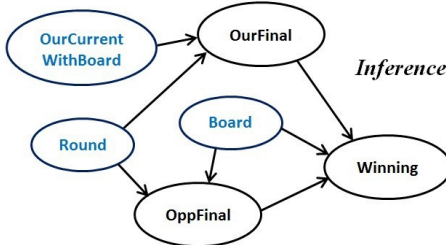


Fig. 2. The structure of the Bayesian network which is used to infer the probability distribution of win, lose and tie

Based on previous research [9], we refine the factors that remarkably influence the probability distribution of win, lose and tie given each hand. These factors are used to construct the Bayesian network. *OurCurrentWithBoard*, *OppFinal*, *OurFinal* and *Board* are the nodes which represent the types of corresponding hand. All the hand types are refined into 22 types as the nodes' states which are shown in Table 1.

The nodes *Round*, *OurCurrentWithBoard* and *Board* in Bayesian network are observed in Fig. 2. *Round* indicates the current round in a hand. *Board* indicates the current board cards on the table. *OurCurrentWithBoard* indicates the type of our hand including the board cards. *OurFinal* indicates our probability distribution of hand types inferred by *Round* and *OurCurrentWithBoard*. *OppFinal* indicates the opponent's probability distribution of hand types inferred by

Table 1. The 22 Types of Hand Cards. Low means the top rank 2 ~ 8 (in Busted Low means lower than Q). Medium means top rank 9 ~ J. High means top rank Q ~ A. L is short for Low. M for Medium and H for High.

Type	Example	Type	Example
Busted Low	♠J ♥10 ♠8 ♣4 ♥2	Pair H Second M	♠K ♥K ♠J ♣4 ♥2
Busted Q	♣Q ♠2 ♥10 ♠7 ♠4	Pair H Second H	♣A ♠A ♥K ♠7 ♠4
Busted K	♠K ♣4 ♥7 ♣Q ♠6	Two Pair Low	♠7 ♣7 ♥6 ♣6 ♠3
Busted A	♥A ♠K ♣10 ♥8 ♣2	Two Pair Medium	♥J ♠J ♣8 ♥8 ♣2
Pair L Second L	♠6 ♣6 ♥3 ♠7 ♣5	Two Pair High	♠A ♣A ♥7 ♠7 ♣K
Pair L Second M	♠5 ♥5 ♠2 ♣10 ♠9	Triple	♠5 ♥5 ♠5 ♣10 ♠9
Pair L Second H	♣3 ♥3 ♠A ♣10 ♠9	Straight	♣K ♥Q ♠J ♣10 ♠9
Pair M Second L	♥J ♣J ♥8 ♠3 ♠2	Flush	♥K ♥10 ♥8 ♥3 ♥2
Pair M Second M	♥J ♠J ♣10 ♣9 ♠3	Full House	♥K ♠K ♣K ♣9 ♠9
Pair M Second H	♥10 ♠10 ♣K ♠6 ♣5	Four of a Kind	♥6 ♠6 ♣6 ♠6 ♣K
Pair H Second L	♠A ♠A ♣3 ♠5 ♥6	Straight Flush	♠A ♠K ♠Q ♠J ♠10

Round and *Board*. *Winning* indicates the probability distribution of win, lose and tie. It is inferred by *OppFinal*, *OurFinal* and *Board* together.

To obtain the conditional probability table, 100,000,000 hands are randomly generated to be used to conduct statistics.

2.2 Model the Opponent's Style Based on Situations

Poker does not merely concern the probability of which hand can win. It is psychological warfare between players. The better we know the opponent, the more money we win. Different from existing work [7], it is limited to classify an opponent into one of styles. A player will not always take actions of an invariable style. Having studied match logs of some players of different styles, we find that a player will show different styles in different situations. While in similar situations the player will show a similar style. We assume that the player's style will be consistent in a similar situation. Then there comes two problems: we need a method to quantify the style; we need to define what "a similar situation" is.

As mentioned in Section I, there are four main styles of players which are recognized: *conservative-loose*, *conservative-tight*, *aggressive-loose* and *aggressive-tight*. In order to get rid of the constraints of limited style, we quantify the style and map them to a continuous interval $[0, 1]$. When the value of the style tends to be 0, the style tends to be aggressive. When the value of the style tends to be 1, the style tends to be conservative. Of course, when the value of style tends to be 0.5, the middle of $[0, 1]$, the style tends to be normal.

When analyzing the opponent's style, we are finding the opponent's motivation behind his actions. We generalize the factors related with opponents'

actions: our action before the opponent, the opponent’s hand(hole cards and board cards) and round. These factors together form the opponent’s style. We call the process *Learning*.

In this module, we propose the notion Action Pair. It consists of our agent’s action and the opponent’s action, which are contextual. For example, “rc” means the opponent checks after our agent’s raise action. Leaving aside the hand, we can intuitively find interesting information hidden in the Action Pair.

If the opponent raises after our agent’s raise action, we assume that the opponent has a strong tendency of aggressiveness. If the opponent checks after our agent’s check action, we assume that the opponent has no particular tendency. Following this principle, we sort all types of Action Pair according to the order which is from “Aggressive” to “Conservative”. We conclude the Action Pair order in Fig. 3.



Fig. 3. The Action Pairs in the order of the style

We assume that if the opponent’s hole cards are unobservable, any analysis of the opponent’s style is unconvincing because we do not have enough exact information to infer the opponent’s motivation behind. Therefore, we only analyze the opponent’s style after a showdown hand when the opponent’s hole cards are observable.

In addition, *Round* is an important factor, which affects the thinking of an opponent. A player will have different considerations in different rounds. It can be easily understood that it is totally different when we have one pair in the first round from the last round.

With the three factors *Action Pair*, *Opponent’s Hand Cards* and *Round* above, the opponent’s style can be modeled.

We propose a method to compute the value of the opponent’s style. The Action Pair and the probability distribution of win, lose and tie are used to map the value of the style to the a continuous interval $[0, 1]$ as shown in Fig. 4.

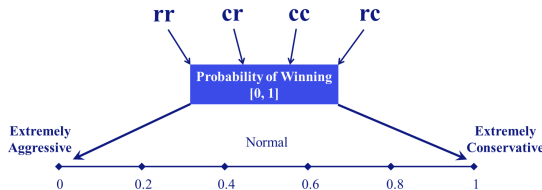


Fig. 4. The graphical representation of the mapping process

Corresponding to each Action Pair, a function $Style(WinP)$ is defined as follows to compute the value of the style:

- For Action Pair “rr” : $Style(winP) = 0.5 \times winP^{\frac{1}{3}}$, $winP \in [0, 1]$.
- For Action Pair “cr” : $Style(winP) = 0.5 \times winP^{\frac{1}{2}}$, $winP \in [0, 1]$.
- For Action Pair “cc” : $Style(winP) = winP$, $winP \in [0, 1]$.
- For Action Pair “rc” : $Style(winP) = \frac{1}{1+e^{-8 \times (-winP+0.5)}}$, $winP \in [0, 1]$.

It can be deduced that if the opponent has a high probability of winning, but he chooses to check, there is a good reason to assign a value representing a conservative tendency as his style. There are still some rules in defining style functions:

- If the opponent has a high probability to win and he raises, the value of style should be tended to be Normal.
- A small probability of winning and a nonconservative Action Pair produce aggressive style. This aggressiveness will be intensified as the Action Pair tends to be more aggressive.
- A big probability of winning and a nonaggressive Action Pair produce conservative style. This conservativeness will be intensified as the Action Pair tends to be more conservative.

When the opponent folds, his hole cards will be unobservable. Then such case of fold will not be considered.

Given the style functions, the calculations will only be operated after showdown. Such information for each style calculation should be ready: Action Pair, Current Round, Opponent's Hole Cards, Board Cards. $Style(winP)$ can be calculated by above information. E.g., $winP$ (the probability of winning) can be obtained in previous **Inference** module using *Round*, *Opponent's Hole cards* and *Board cards*. Then we can calculate the style of the opponent according to $Style(winP)$. This style is what the opponent shows in the current situation of a hand. It satisfies the distribution of the current information of a hand. The current situation consists of information which are observable by both the opponent and ourselves. *Round* and *Board cards* are obviously the most appropriate information within the situation. The table *StyleTable*, which is indexed by *Round* and *BoardCardsType*, is created to store the values of styles as “experience”. Each obtained value will be put in the position where is indexed by the combination of corresponding *Round* and *Board cards type* and calculated with the historical data together in a statistical way. The greatest advantage is that it's an online learning which can learn the opponent's style while in the play.

2.3 Make Decisions with Modeled Style

We call this process the **Tactic** module in Fig. 1. When estimating the value of the opponent's hand cards in a certain stage of a hand, information such as *Round*, *Board cards* should be first observed. Meanwhile, the historical style will be obtained from *StyleTable*. The style is what the opponent shows in the similar situation based on “experience”. Given the latest Action Pair in the same round, we can restore the opponent's approximate probability of winning. The approach is almost the inverse process of obtaining the style of opponent. Following functions are designed to map the style to the probability of winning:

- For Action Pair “rr” or “cr” : $EWinP(style) = style^{\frac{2}{3}}$, $style \in [0, 1]$
- For Action Pair “cc” : $EWinP(style) = style$, $style \in [0, 1]$
- For Action Pair “rc” : $EWinP(style) = 4 \times (x - 0.5)^3 + 0.5$, $style \in [0, 1]$

Note that all the calculations can only be operated when all the required information comes from the same round of a hand. This is consistent with what we have always emphasized that the opponent will show a similar style in a similar situation.

After having obtained the estimated opponent’s probability of winning, we can have many ways to adaptively change our action. The simplest way is that if the estimated value is very high, the opponent may have a very good hand. Because the value of the style is an average calculated in a number of hands, we do not need to consider whether the opponent is bluffing this time.

3 Experiments and Analysis

There are two issues to be verified in the experiments. The first is whether the style shows a significant difference in different situations. The second is whether style modeling can improve the performance of an agent. These two issues guide the experiments. To compare the results, we use three agents as opponents with different styles, which will be described in the following section.

3.1 Contrast Agents and Opponents

Contrast agents. The contrast agent bases on previous *Inference* module. 100,000,000 hands are randomly generated to train its Bayesian network. It raises when its probability of winning is over 0.6 and folds when its probability of losing is over 0.6. It is designed to carry normal style.

Rule-Based Opponents. By adjusting the threshold of the probability of winning to take actions, we simulate the aggressive and conservative opponents based on previous *Inference* module. The aggressive agent raises when its probability of winning is over 0.4 and folds when its probability of losing is over 0.8. The conservative agent raises when its probability of winning is over 0.8 and folds when its probability of losing is over 0.4.

Monash BPP(Bayesian Poker Player) Opponent. We introduce the Monash BPP of *ACPC*(Annual Computer Poker Competition) 2006 version as another opponent [14]. Its complete version got the third prize in the competition. For simplification, we do not include its opponent modeling part.

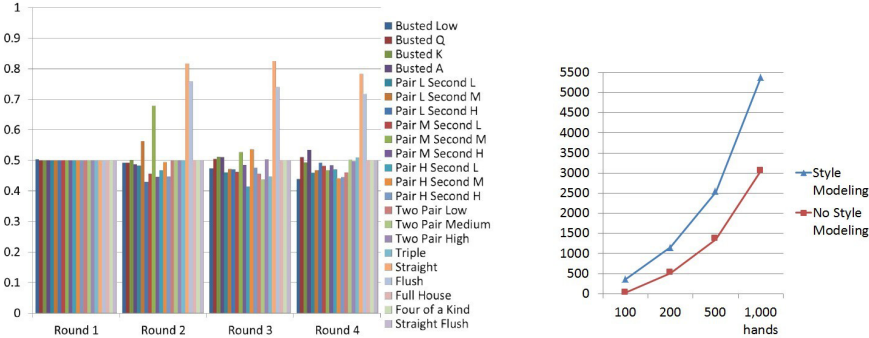
3.2 Results and Analysis

Each experiment is performed in 100, 200, 500 and 1,000 hands. The performance of the agents with and without style modeling are compared by the

winning money after each experiment. After each opponent’s experiments, the corresponding *StyleTable* will be analyzed. The experiments are carried on the *ACPC* Server.

The Opponent with Aggressive Style. As mentioned before, the threshold adjustment will make the opponent tend to take more risks to play more hands. Thus, the opponent will act aggressively even if it has no big hand in many situations of a hand. The performances of the agent with and without style modeling are then observed. A tactic is tried so that if the estimated opponent’s probability of winning is larger than the agent’s, we assume that the opponent has a larger hand than our agent and it will not raise in any way.

Note that the value of the style is 0.5 in some situations. The reason is that, in the first round, two board cards cannot form the combination of some hand types. In the following round, some hand types with a low probability cannot appear. Thus, a priori value 0.5 has to be assigned to this situation.



(a) The *StyleTable* for the aggressive opponent. The vertical axis indicates the scale of the style. (b) The result of comparison with style modeling against the aggressive opponent.

Fig. 5. Experiments for the aggressive opponent. In (b), the vertical axis indicates the scale of money won by the agents. The abscissa indicates the scale of the number of hands the agents play.

In Fig.5(a), the opponent shows an obvious aggressive style in most of situations. The values of styles are generally less than 0.5. In addition, the style is variable in all situations. It can be justified by the experiments that the opponent shows various styles in different situations.

According to Fig.5(b), a big advantage is shown to the aggressive opponent. Although the aggressive opponent loses money more easily, the agent with style modeling wins averagely more than 50% of the agent without style modeling in each match. By analyzing the log of matches, we find that the agent with style modeling does well in reducing its threshold of raising to adjust itself to the opponent’s aggressive style and win more.

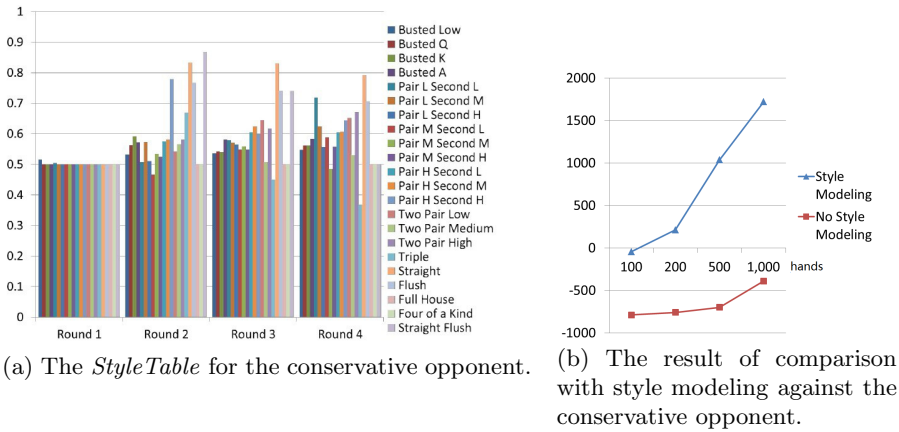


Fig. 6. Experiments for the conservative opponent

The Opponent with Conservative Style. The conservative opponent tends to take fewer risks so that it will play only sure hands and raise only when it has a greater assurance. We make slight adjustment that if the estimated opponent's probability of winning is larger than the agent's, the agent will directly fold.

In Fig.6(a), the conservative opponent clearly shows the opposite tendency from aggressive opponent. The values of styles are generally greater than 0.5. It indicates that the opponent shows an obvious conservative style in most of situations. The obtained style from the style model is able to reasonably demonstrate the tendency of the opponent. Furthermore, we find that again the style in different situations of a hand is variable as in last experiment.

In Fig.6(b), the compared agent loses all the matches. In contrast, the agent with style modeling wins money at a steady rate along with the increase of the hands. By analyzing the log, we conclude that the conservative opponent gives no chance to us since it only plays a good hand. So the compared agent with no style modeling has no advantages. When it raises, the opponent either folds or has a greater hand. However the agent with style modeling will fold once the estimated opponent's probability of winning is greater than its hand. Thus, it is not difficult to explain why the agent with style modeling performs better.

The Opponent Monash BPP. Based on Bayesian network, it first infers the probability of winning through belief propagation. Combining the probability and its next action, the final decision is made. In addition, it adds randomization to make itself unpredictable.

In Fig.7(a), the opponent shows no clear general style. In this case, a single style modeling will not work. Considering that the styles in different situations are variable, style modeling will take effect.

In Fig.7(b), the agent with style modeling begins to show significant advantage from 500 hands match. Based on the analysis of the log of matches, the "experience" *StyleTable* does not help much in the first 200 hands. Since the statistical effect is not obvious at the very beginning. After that, the agent with

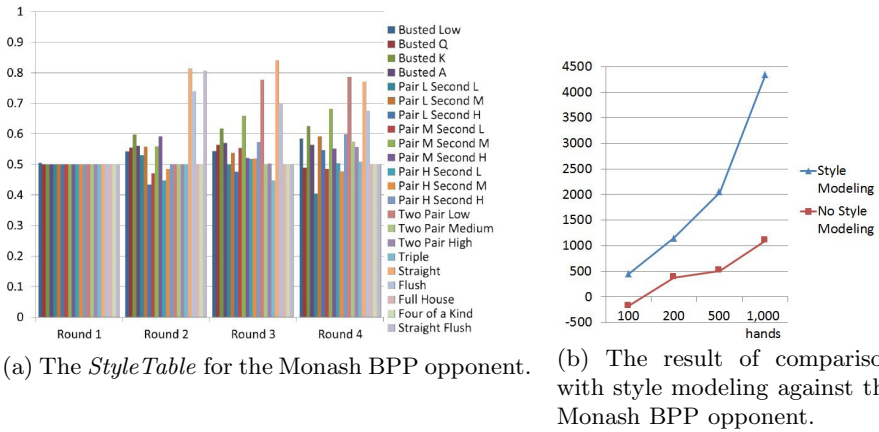


Fig. 7. Experiments for the Monash BPP opponent

style modeling successfully avoids to raise when facing the Monash BPP's larger hand. Meanwhile, it succeeds in finding the opponent's poor hand and stays close to it even if the hand of the agent with style modeling is also poor.

3.3 Discussion

Even if the opponent has a distinct style, it is hard for him to avoid showing inconsistent styles in different situations of a hand. As a result, we cannot use his single style to guide all of our actions. This is why we propose the method to model the style in all situations. The obtained *StyleTable* and the estimated opponent's probability of winning effectively help the agent improve the performance. The improved agent adapts to the opponent and performs better than its original version, which is very encouraging.

Additionally, the agent with style modeling will adjust with the increase of the hands. We find that the *StyleTable* tends to be a stable distribution after 1000 hands. It indicates that the opponent's style modeling based on situations will not change much after that.

4 Conclusions and Future Work

This paper proposed a novel method to model the opponent's style based on situations and a detailed procedure centered at the modeled style. In order to achieve this goal, a notion of Action Pair and a calculation method are first given to obtain the value of the style and known as "experience". The specific style will be obtained from the "experience" according to the current situation of the hand. Then the style was used to help estimate the value of the opponent's hand. In this way, the final decision can be made by some tactics.

The experimental results show that the modeled style validates the correctness of our assumption that a player does not always show an invariable style

in situations. This style modeling subdivides the style into different situations. Having obtained the style values in all situations, we can have an insight into the opponent's style in detail, which cannot be easily observed by existing methods. It's encouraging to find that the obtained style and the estimated opponent's probability of winning can effectively improve the performance of the original agent. They jointly help the agent predict the hidden information of the opponent's hand and earn more money in fixed hands.

Future work will pay attention to building the style quantifying process more accurately. The method will also be extended to No-Limit Hold'Em Poker where the pot is an important factor which has a great impact on player's decision.

Acknowledgments. We would like to acknowledge the support for this work from the National Science Foundation of China (Grant Nos. 61035003, 61170180, 61175042, 61021062), and Program for New Century Excellent Talents in University (Grant No. NCET-10-0476).

References

1. Billings, D., Papp, D., Schaeffer, J., Szafron, D.: Opponent modeling in poker. In: Proceedings of the Fifteenth National AAAI Conference, pp. 493–499 (1998)
2. Davidson, A.: Opponent modeling in poker: Learning and acting in a hostile and uncertain environment. Master's thesis, University of Alberta (2002)
3. Machado, M., Fantini, E., Chaimowicz, L.: Player Modeling: Towards a Common Taxonomy. In: 16th International Conference on Computer Games, pp. 50–57 (2011)
4. Lockett, A.J., Mikkulainen, R.: Evolving opponent models for Texas Hold'Em. In: IEEE Symposium on Computational Intelligence and Games, pp. 31–38 (2008)
5. Barone, L., While, L.: Adaptive learning for poker. In: Proceedings of the Genetic and Evolutionary Computation Conference, pp. 566–573 (2000)
6. Baker, R.J.S., Cowling, P.I.: Bayesian Opponent Modeling in a Simple Poker Environment. In: IEEE Symposium on Computational Intelligence and Games, pp. 125–131 (2007)
7. Davidson, A., Billings, D., Schaeffer, J., Szafron, D.: Improved opponent modeling in poker. In: International Conference on Artificial Intelligence, pp. 1467–1473 (2000)
8. Billings, D., Davidson, A., Schaeffer, J., Szafron, D.: The challenge of poker. *Artificial Intelligence* 134, 201–240 (2002)
9. Korb, K.B., Nicholson, A.E., Jitnah, N.: Bayesian poker. In: Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence, pp. 343–350 (1999)
10. Billings, D., Castillo, L., Schaeffer, J., Szafron, D.: Using probabilistic knowledge and simulation to play poker. In: Proceedings of the 16th National Conference on Artificial Intelligence, pp. 697–703 (1999)
11. Russell, S., Norvig, P.: *Artificial Intelligence: A Modern Approach*. Prentice Hall (2006)
12. Korb, K.B., Nicholson, A.E.: *Bayesian Artificial Intelligence*. Chapman and Hall/CRC, Boca Raton (2004)
13. Tretyakov, K., Kamm, L.: Modeling texas hold'em poker strategies with bayesian networks (2009)
14. Annual Computer Poker Competition,
<http://www.computerpokercompetition.org/>