

Title	初級者の教育を目的とした状況に応じた着手モデル選択
Author(s)	田中, 悠
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	<a href="http://hdl.handle.net/10119/12043">http://hdl.handle.net/10119/12043</a>
Rights	
Description	Supervisor:池田 心, 情報科学研究科, 修士

# Selecting of action models, for educating beginner

Yu Tanaka (1110201)

School of Information Science,  
Japan Advanced Institute of Science and Technology

February 12, 2014

**Keywords:** Computer mahjong, machine learning, classification.

The “strength” of computer players has remarkably developed with the advancement of computer performance and the technology of artificial intelligence. Now, their strength exceeds any human player’s strength in many board games. For example, in the case of Chess, Deep Blue defeated the world champion Kasparov in 1997, and in the case of Shogi (Japanese Chess) many professionals have lost recently to computer players. So we can say that the “strength” of computer players is quite enough for most human players.

Research of more complex games is also making advances as well. For example, new target games include Puyopuyo, where the state transition is probabilistic, Poker, whose information is imperfect for the players, Starcraft, which is played by many players at once and “Mahjong”, which has all of those difficulties.

As the research of computer players makes advances, many methods for constructing computer players have been proposed. In the past, state evaluation functions were designed manually. But nowadays, they can be automatically learned by using machine learning and optimization techniques. For example, the Bonanza method is a method to learn a state evaluation function from game records and the Bradley-Terry method is a method to learn an action evaluation function. Neural networks (ANN), and Support Vector Machines (SVM) are also commonly used.

Those methods contribute to the development of computer players. On the other hand, understanding the reason why the program chooses an

action became more difficult for humans. Even if their output is strong we cannot understand “why” it is strong. As we mentioned, the “strength” of computer players is enough for most players. So we forecast that aspects other than “strength” will be more important in the future for computer players. In this research, we focused on the “education” aspect and propose a method to construct a computer player which can output not only the move but also the strategy that explains the move.

“Mahjong”, the target of this research, is very famous in Japan. This game has an element of chance. Therefore beginner and advanced players can play together. But because of the complexity of the rules, it is difficult for a beginner player to learn the rules and become strong by himself.

Basically, one game of Mahjong is made of about 10 independent rounds, and so in the case of Mahjong, choosing a suitable strategy according to the current state of the round and also the current score after the previous rounds is important to become a strong player. We guess that the biggest difference between beginner players and advanced players is that choice of strategy. Beginners always think about completing their hand. As a result, their loss of score increases and they suffer a defeat.

In this paper, we propose a method to construct an educational computer player, by the view of strategy choosing. At first, we found from the analysis of advanced players game records that most of Mahjong hands can be explained by five strategies, and we picked up 3 of them “win quickly”, “win with a high score” and “prevent losing score”. Then we constructed 3 single-target action models, each of which evaluates one of the picked-up strategies.

The efficiency of the single-target action models was tested in a game of “single player Mahjong”, and we obtained that the “win quickly” single-target action model completed a winning hand 1.5 times more frequently than the “win with a high score” single-target action model, whereas the average score of completed hand of the “win with a high score” single-target action model is 1.7 times higher than that of the “win quickly” single-target action model.

For the next step, we labeled in game records the strategy of advanced player’s hand, according to the evaluation from the three single-target action models. Then we made a decision tree by using the labeled game

records as a training data. From the input of the current game state, this decision tree outputs a list of weights for the single-target action models. This list of weights can be interpreted as an advice like “ which strategy should be taken ” and “ therefore this hand is the best ” .

The complete system to find the best moves from a current game state is made from a first step where the decision tree outputs the best strategies in the form of a list of weights, and a second step where these weights are used to balance the three single-target action models. The generalization performance of the complete computer player system was tested by computing the probability that the move of advanced players is found in the top 3 moves of the computer player. We found a probability of 86%.

By using this system, we are able to make a new kind of advice by the computer player. Most of conventional computer players could not explain “ why the hand is good ” . On the other hand the proposed system succeeds to output good hands with information about the appropriate strategy. For example, “ In this situation, you should regard to prevent losing score, rather than completing your hand. Hand A is a good choice to complete a high score hand, but it may be useful for the other player, so you should choose hand B ” . We believe that this new kind of advice is more educational than the past attempts.

We hope that this research will help to increase the number of Mahjong players and to lighten the burden of educating beginner players. Also, a similar approach will probably be possible for constructing computer players in other strategic games.