Abstract

In recent years, AI technology has advanced rapidly. Technologies such as image recognition and speech recognition are being used across various fields, including healthcare, manufacturing, transportation, distribution, education, and entertainment. Recently, AI has evolved not only in recognizing text, images, and audio but also in naturally generating them, allowing it to increasingly support or even replace human creators and operators.

Games, with their clear rules, accessibility, and ease of evaluation, also require intellectual thinking. Due to these characteristics, games have frequently served as testbeds for artificial intelligence. Indeed, a variety of AI technologies have been developed or evaluated with games as their focus.

When applied to games, the most straightforward goal is to create a strong computer player. This goal has been pursued in academic research for many years, leading to the development of various methods and notable achievements, such as DeepBlue in chess and AlphaGo Zero in Go, which have outperformed top human players. Although research continues to focus on building stronger AI players, interest is gradually shifting toward making strong computer players beneficial for humans. Examples of this include creating human-like computer players, implementing AI that can be played to entertain, automatically generating puzzles and game content, and supporting skill improvement.

Supporting the improvement of beginner and intermediate players is crucial for maintaining or increasing the player base of a game. A coach who provides this support needs more than just "skill in the game." Advanced abilities are required, including explaining the core of the game, identifying the current issues a student faces, guiding them along the path to improvement, and maintaining their motivation. For this reason, many sports and well-known games have established roles for coaches focused on player improvement, where students typically need to pay a fee for quality instruction. If part of this coaching role could be replaced by a computer coach, it could not only promote the development of the game itself but also serve as a meaningful step toward building a better relationship between humans and AI.

Ikeda et al. conducted interviews with Go coaches to investigate the types of guidance provided. They reported that among the various coaching methods, one of the core tasks is reviewing a match after it concludes, pointing out the student's bad moves, and explaining the reasons behind them. The moves considered "bad" are not merely "loss-making" moves. Some moves that incur minor losses are allowed if they serve as bold moves to take control or as safe moves to secure victory. On the other hand, even if a loss is minor, a move that reveals flawed thinking on the student's part is often pointed out. To make the selection and explanation of bad moves replicable by a computer, Ikeda et al. asked Go coaches to annotate game records of intermediate-level players. They requested that for each move, the coach decide "whether to point it out as a bad move" and, if so, "to select the reason from ten categories." Using the collected training data, they applied supervised learning to develop a model for detecting bad moves and a model for reasoning behind them. Although these models perform slightly below the level of a human expert coach, professional players have judged them to be sufficiently practical.

In this study, following the work of Ikeda et al., we undertook a similar approach using Xiangqi (Chinese Chess) as our focus. Xiangqi is a very popular game in China. While its nature shares similarities with Chess and Shogi, it is a fairly different game from Go.

Therefore, to build a model, it was necessary to carry out tasks such as game record collection, reason label selection, annotation, and feature engineering.

We first conducted interviews with several Xiangqi coaches and observed their actual coaching sessions. As a result, we found that "pointing out bad moves and explaining the reasons" is one of the main tasks, similar to Go. However, the nature of the reason explanations was found to be completely different from Go. We first accumulated much natural-language explanations from the coaches regarding their reasons and, through discussions with the coaches, grouped them into about 16 categories. We then requested annotations on 100 game records from intermediate-level players. The collected training data consisted of 4,666 annotations, of which only 798 were marked as bad moves. If this were to be treated as a 16-class problem as-is, some groups would have too few samples, making high-accuracy predictions unlikely. Therefore, we further grouped these into five reason groups for supervised learning.

In supervised learning, when there is a significant imbalance in the number of samples across classes, it often limits the learning performance. In fact, the number of moves identified as bad moves is only about one-fifth of those identified as good moves. Therefore, we decided to augment the training data for bad moves through oversampling.

One critical factor that affects the performance of supervised learning is what input to use. In cases where new board states can be generated indefinitely through self-play, as with AlphaZero, or where an extremely large database is available, as with Maia, high performance can be expected even if the board state itself is used as input. However, the training data we used consists of only 4,666 samples for the detection model and 798 for the reasoning model. For this reason, rather than directly inputting the board state and moves, we need to provide (as input) the "meaning and features in Xiangqi" that they have.

We held discussions with Xiangqi coaches to clarify which aspects of the board state and moves are considered important when detecting bad moves and providing reasoning. These aspects were then quantified in a definable way. Next, we conducted preliminary experiments to investigate how combining these aspects could improve performance—or, alternatively, decrease it due to overfitting.

Additionally, many different models have been proposed for supervised learning, and it is essential to choose the appropriate model based on the characteristics of the input space, output space, and the volume and distribution of the data. After comparing and evaluating numerous models available in the supervised learning tool Weka, we selected RandomForest for the bad move detection model and AdaboostM1 for the reasoning model.

Based on these preparations, the training yielded an accuracy for the bad move detection model (F1-score for bad moves) of 0.825. Although this is slightly lower than the level of agreement among the three coaches (average F1-score of 0.837), it is quite close in performance. The reasoning model achieved an accuracy of 0.501, which slightly exceeded the human average of 0.497. While this result is not statistically significant due to the limited number of samples, it can be considered promising.

Finally, a coach with a professional Xiangqi coaching license evaluated the bad move detection and reasoning results for 10 games. The results from the three coaches involved in training, as well as those from the proposed RandomForest/AdaboostM1 method, were presented in a blind manner and rated on a five-point scale for suitability. The evaluation for bad move detection showed scores of 3.988 for the human coaches and 4.074 for the proposed method. For reasoning, the human coaches scored 4.074, while the proposed method scored 4.048. These results indicate that, in addition to statistical agreement, the quality of our approach is at a level comparable to human coaches.