Article

Chinese Standard Mahjong AI Based on Supervised learning

Yizhe Huang [1,†]**, Kangcheng Luo** [2,†]

1   Peking University; szhyz@pku.edu.cn
2   Peking University; 2000012940@stu.pku.edu.cn
†   These authors contributed equally to this work.

Abstract: In reccent years, Artificial Intelligence (AI) has achieved great success in many fields. Based on deep-learning methods, Game AI has achieved human level performance in many challenging tasks such as Go, chess, and Texas Hold'em. Chinese Standard Mahjong is a popular multi-player imperfect-information game originated in China. The large amount of hidden information and the complex playing/scoring rules make it very challenging for AI research and attract many researchers. In IJCAI 2022 Mahjong AI Competition, We designed an AI for Mahjong based on supervised learning, which got a good performance in the competition.

1. Introduction

Game AIs developed rapidly in the past 30 years. In 1994, checkers program Chinook beat World Checker Champion Tinsley and this is the first time in history that a computer has won a human World Championship.[1].After that, AI has surpassed professional human players in succession, from the Atari[2], AlphaGo[3], AlphaStar[4] to JueWu[5]. Using deep reinforcement learning, AI can handle not only simple perfect-information two-player games but also complex imperfect-imformation multi-player games.

Chinese Standard Mahjong is a four-player game with imperfect information. In each round of a Mahjong game, four players compete with each other to firstly complete a winning hand. Building a intelligent Mahjong AI is a great challenge.

First, Mahjong has complex playing and scoring rules. There are 144 tiles in Mahjong, including 36 characters, 36 dots, 36 bamboos, 16 winds, 12 dragons, and 8 flowers which lead to a huge number of winning hands. The combinations of winning hands are various and can be very different from each other. The rules of Mahjong are different in different places. In Chinese Standard Mahjong, a winning hand must have at least 8 points, which increases the difficulty of a winning hand. There are 81 types of faan with different points, including specific combinations of tiles and special winning approaches such as *from a discard* and *from the wall*. Such scoring rules are far more complicated than previously studied games like Go and Texas hold'em. Besides, the regular order of plays in Mahjong can be interrupted by Chow, Pong, Kong, etc.. This leads the successful MCTS based technique for Go isn't suitable for Mahjong, for we can't build a regular game tree. There are too many paths between the consecutive actions of one player.

Second, the large number of hidden information makes Mahjong a much more difficult imperfect-information game than card games like Texas hold'em poker and DouDiZhu. Every player has 13 privite tiles in hand and 21 tiles in the tile wall. These will become visible only if they're discarded by the players. On average, there are nearly 10^{48} hidden states[6] that are indistinguishable to a player for one information set. Based on the visible information, it's hard to determine an action to be good or not.

In this competition, we built a Chinese Standard Mahjong bot, *Behold*. It adopts deep convolutional neural networks as its model and is trained through supervised learning

using game datasets from strong AIs provided by the competition. *Behold* ranked the first　⁴⁰
in Final Round Stage One and the second in Stage Two.　⁴¹

References　⁴²

1.　Schaeffer, J.; Lake, R.; Lu, P.; Bryant, M. CHINOOK The world man-machine checkers champion. *AI Magazine* **1996**, *17*, 21.　⁴³
2.　Mnih, V.; Kavukcuoglu, K.; Silver, D.; Rusu, A.A.; Veness, J. Human-level control through deep reinforcement learning. *Nature*　⁴⁴
　　2015, *518*, 529–533.　⁴⁵
3.　Silver, D.; Huang, A.; Maddison, C.J.; Guez, A.; Sifre, L.; et al. Mastering the game of Go with deep neural networks and tree　⁴⁶
　　search. *Nature* **2016**, *529*, 484–489.　⁴⁷
4.　Vinyals, O.; Babuschkin, I.; Czarnecki, W.M.; Mathieu, M.; Dudzik, A.; et al. Grandmaster level in StarCraft II using multi-agent　⁴⁸
　　reinforcement learning. *Nature* **2019**, *575*, 350–354.　⁴⁹
5.　Ye, D.; Chen, G.; Zhang, W.; Chen, S.; Yuan, B.; Liu, B.; Chen, J.; Liu, Z.; Qiu, F.; Yu, H.; et al. Towards Playing Full MOBA Games　⁵⁰
　　with Deep Reinforcement Learning. *arXiv e-prints* **2020**, p. arXiv:2011.12692, [arXiv:cs.AI/2011.12692].　⁵¹
6.　Li, J.; Koyamada, S.; Ye, Q.; Liu, G.; Wang, C.; et al. Suphx: mastering mahjong with deep reinforcement learning.　⁵²
　　arXiv:2003.13590v2 **2020**.　⁵³