DecisionHoldem: Safe Depth-Limited Solving With Diverse Opponents for Imperfect-Information Games

Qibin Zhou¹, Dongdong Bai¹*, Junge Zhang¹†, Fuqing Duan², Kaiqi Huang¹
¹Institute of Automation, Chinese Academy of Sciences, Beijing, China
²Beijing Normal University, Beijing, China
zqbagent@gmail.com, baidongdong@nudt.edu.cn, jgzhang@nlpr.ia.ac.cn, fqduan@bnu.edu.cn, kqhuang@nlpr.ia.ac.cn

Abstract

An imperfect-information game is a type of game with asymmetric information. It is more common in life than perfect-information game. Artificial intelligence (AI) in imperfect-information games, such like poker, has made considerable progress and success in recent years. The great success of superhuman poker AI, such as Libratus and Deepstack, attracts researchers to pay attention to poker research. However, the lack of open-source code limits the development of Texas hold’em AI to some extent. This article introduces DecisionHoldem, a high-level AI for heads-up no-limit Texas hold’em with safe depth-limited subgame solving by considering possible ranges of opponent’s private hands to reduce the exploitability of the strategy. Experimental results show that DecisionHoldem defeats the strongest openly available agent in heads-up no-limit Texas hold’em with safe depth-limited subgame solving by considering asymmetric information. It is more common in life than perfect-information game. Artificial intelligence (AI) in imperfect-information games, such like poker, has made considerable progress and success in recent years. The great success of superhuman poker AI, such as Libratus and Deepstack, attracts researchers to pay attention to poker research. However, the lack of open-source code limits the development of Texas hold’em AI to some extent. This article introduces DecisionHoldem, a high-level AI for heads-up no-limit Texas hold’em with safe depth-limited subgame solving by considering possible ranges of opponent’s private hands to reduce the exploitability of the strategy. Experimental results show that DecisionHoldem defeats the strongest openly available agent in heads-up no-limit Texas hold’em, namely Slumbot, and a high-level reproduction of Deepstack, viz, Openstack, by more than 730 mbb/h (one-thousandth big blind per round) and 700 mbb/h. Moreover, we release the source codes and tools of DecisionHoldem to promote AI development in imperfect-information games.

1 Introduction

The success of AlphaGo [Silver et al., 2016] has led to increasing attention to the study of game decision-making [Brown and Sandholm, 2018; Moravčík et al., 2017; Brown et al., 2018; Brown and Sandholm, 2019c]. Unlike perfect-information games, such as Go, real-world problems are mainly imperfect-information games. The hidden knowledge in poker games (i.e., private cards) corresponds to the real world’s imperfect-information. Research on poker artificial intelligence (AI) can provide means to deal with problems in life, such as financial market tracking and stock forecasting.

Research in imperfect-information games, particularly poker AI [Brown and Sandholm, 2018; Moravčík et al., 2017; Brown et al., 2017] [Hartley, 2017] [Brown and Sandholm, 2019a] [Schmid et al., 2019] [Farina et al., 2019c] [Farina et al., 2019a] [Farina et al., 2019b] [Li et al., 2020a] is only tested in games with small decision space, such as Leduc hold’em and Kuhn Poker. These algorithms may not work well when applied to large-scale games, such as Texas hold’em.

After decades of research, the poker AI DeepStack developed by Matej Moravčík et al. [Moravčík et al., 2017] and Libratus developed by Noam Brown and Tuomas Sandholm [Brown and Sandholm, 2018] successively defeat human professional players in 2017. This event confirms the breakthrough for HUNL. Subsequently, the poker AI Pluribus, also constructed by Noam Brown and Tuomas Sandholm [Brown and Sandholm, 2019c], defeats the human professional players in six-man no-limit Texas hold’em. Although Science magazine has published the poker AI mentioned above [Brown and Sandholm, 2018; Moravčík et al., 2017], the relevant code and main technical details have not been made public.

In addition, considerable poker AI progress [Brown et al., 2017] [Hartley, 2017] [Brown and Sandholm, 2019a] [Farina et al., 2019c] [Farina et al., 2019a] [Farina et al., 2019b] [Li et al., 2020a] is only tested in games with small decision space, such as Leduc hold’em and Kuhn Poker. These algorithms may not work well when applied to large-scale games, such as Texas hold’em.

In this paper, we propose a safe depth-limited subgame solving algorithm with diverse opponents. To evaluate the algorithm’s performance, we achieve a high-performance and high-efficiency poker AI based on it, namely DecisionHoldem. Experiments show that DecisionHoldem defeats the
strongest public poker AI, such as Slumbot\(^1\) (champion of 2018 Annual Computer Poker Competition [ACPC]) and OpenStack (a reproduction of DeepStack built-in OpenHoldem [Li et al., 2020b]\(^2\), by a big margin. Meanwhile, we release DecisionHoldem’s source code, and tools for playing against the Slumbot and OpenHoldem [Li et al., 2020b]. In addition, we also provide a platform to play DecisionHoldem with humans (as in Figure 1\(^3\)). Our code is available at https://github.com/AI-Decision/DecisionHoldem.

| Round       | Number of Abstract Hands | 1st ~ 2nd Actions | 3rd ~ 5th Actions | Remaining Actions |
|-------------|--------------------------|-------------------|-------------------|-------------------|
| Pre-Flop    | 169                      | F, C, 0.5P, P, 2P, 4P, A | F, C, P, 2P, 4P, A | F, C, A          |
| Flop        | 50,000                   | F, C, 0.5P, P, 2P, 4P, A | F, C, P, 2P, 4P, A | F, C, A          |
| Turn        | 5,000                    | F, C, 0.5P, P, 2P, 4P, A | F, C, P, 2P, 4P, A | F, C, A          |
| River       | 1,000                    | F, C, 0.5P, P, 2P, 4P, A | F, C, P, 2P, 4P, A | F, C, A          |

Table 1: The number of abstract hands and actions available for each round (pre-flop, flop, turn, and, river) of DecisionHoldem on HUNL. F, C, 0.5P, P, 2P, 4P, and A represent Fold, Call, 0.5 Pot size, 1.0 Pot size, 2.0 Pot size, 4.0 pot size, and all-in, respectively.

In the real-time search part, we propose a safer depth-limited subgame solving algorithm than modicum’s [Brown et al., 2018] on subgame solving by considering diverse opponents for off-tree nodes. Since the opponent’s private hand range reflects the opponent’s play style and strategy, we propose a safe depth-limited subgame solving method by explicitly modeling diverse opponents with different ranges. This algorithm can refine the degree of subgame strategy without worsening the exploitability compared with the blueprint strategy. That is to say, safe depth-limited solving with diverse opponents can significantly enhance the AI decision-making level and ability with changeable challenges. Our subsequent articles will introduce the details of the algorithm.

3 Experiments and Results
DecisionHoldem plays against Slumbot and OpenStack [Li et al., 2020b] to test its capability. Slumbot is the champion of the 2018 ACPC and the strongest openly available agent in HUNL. OpenStack is a high-level poker AI integrated in OpenHoldem, a replica AI version of DeepStack. The experimental configurations are as follows.

For the first three rounds of the game, DecisionHoldem prioritizes using blueprint strategies when making decisions. For off-tree nodes, DecisionHoldem starts a real-time search. For the first two rounds of poker (preflop, flop), the real-time search iterations are 6,000 times; for the third round (turn), the real-time search iterations are 10,000 times.

While for the last round (river), DecisionHoldem employs the safe depth-limited subgame solving algorithm for real-time search with 10,000 iterations directly.

In approximately 20,000 games against Slumbot, DecisionHoldem’s average profit is more remarkable than 730 mbb/h (one-thousandth big blind per round). It ranked first in statistics (DecisionHoldem’s name on the leaderboard is zqbAgent\(^4\), as the Figure 2 and 3. With approximately 2,000 games against OpenStack, DecisionHoldem’s average profit is greater than 700 mbb/h, and the competition records are available in the Github repository of DecisionHoldem.

4 Conclusions
This paper introduces the safe depth-limited subgame solving algorithm with the exploitability guarantee. It achieves

\(^1\)www.slumbot.com
\(^2\)holdem.ia.ac.cn
\(^3\)https://github.com/ishikota/PyPokerGUI
\(^4\)https://github.com/ericgjackson/slumbot2017/issues/11
the outstanding AI DecisionHoldem for HUNL with the proposed subgame solving algorithm for real-time search and suitable abstraction methods for blueprint strategy. DecisionHoldem defeats the current typical public high-level poker AI, namely Slumbot and OpenStack. To our best knowledge, DecisionHoldem is the very first open-source high-level AI for HUNL. Meanwhile, we also provide toolkits against Slumbot and OpenStack, and a platform to play DecisionHoldem with humans to assist researchers in conducting further research.

References

[Brown and Sandholm, 2018] Noam Brown and Tuomas Sandholm. Superhuman ai for heads-up no-limit poker: Libratus beats top professionals. Science, 359(6374):418–424, 2018.

[Brown and Sandholm, 2019a] Noam Brown and T. Sandholm. Solving imperfect-information games via discounted regret minimization. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, pages 1829–1836, 2019.

[Brown and Sandholm, 2019b] Noam Brown and Tuomas Sandholm. Solving imperfect-information games via discounted regret minimization. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, pages 1829–1836, 2019.

[Brown et al., 2017] Noam Brown, Christian Kroer, and T. Sandholm. Dynamic thresholding and pruning for regret minimization. In AAAI, 2017.

[Brown et al., 2018] Noam Brown, Tuomas Sandholm, and Brandon Amos. Depth-limited solving for imperfect-information games. arXiv preprint arXiv:1805.08195, 2018.

[Farina et al., 2019a] Gabriele Farina, Christian Kroer, and T. Sandholm. Stable-predictive optimistic counterfactual regret minimization. In NeurIPS, 2019.

[Farina et al., 2019b] Gabriele Farina, Christian Kroer, and T. Sandholm. Optimal regret minimization for extensive-form games via dilated distance-generating functions. In NeurIPS, 2019.

[Farina et al., 2019c] Gabriele Farina, Christian Kroer, and T. Sandholm. Regret circuits: Composability of regret minimizers. In ICML, 2019.

[Hartley, 2017] M. Hartley. Multi-agent counterfactual regret minimization for partial-information collaborative games. 2017.
[Johanson, 2013] Michael Johanson. Measuring the size of large no-limit poker games. *arXiv preprint arXiv:1302.7008*, 2013.

[Li et al., 2020a] Hui Li, Kailiang Hu, Zhibang Ge, Tao Jiang, Yuan Qi, and L. Song. Double neural counterfactual regret minimization. *ArXiv*, abs/1812.10607, 2020.

[Li et al., 2020b] Kai Li, Hang Xu, Meng Zhang, Enmin Zhao, Zhe Wu, Junliang Xing, and Kaiqi Huang. Openholdem: An open toolkit for large-scale imperfect-information game research. *arXiv preprint arXiv:2012.06168*, 2020.

[Moravčík et al., 2017] Matej Moravčík, Martin Schmid, Neil Burch, Viliam Lisý, Dustin Morrill, Nolan Bard, Trevor Davis, Kevin Waugh, Michael Johanson, and Michael Bowling. Deepstack: Expert-level artificial intelligence in heads-up no-limit poker. *Science*, 356(6337):508–513, 2017.

[Schmid et al., 2019] Martin Schmid, Neil Burch, Marc Lanctot, Matej Moravcik, Rudolf Kadlec, and Michael H. Bowling. Variance reduction in monte carlo counterfactual regret minimization (vr-mccfr) for extensive form games using baselines. *ArXiv*, abs/1809.03057, 2019.

[Silver et al., 2016] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *nature*, 529(7587):484–489, 2016.

[Zinkevich et al., 2007] Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information. In *Advances in Neural Information Processing Systems 20*, volume 20, pages 1729–1736, 2007.