Using deep reinforcment learning and self-play to master the game abalone without human knowledge

Ture Claußen, 1531067, ture.claussen@stud.hs-hannover.de

Hochschule Hannover Fakultät IV

Abstract. Perfect information games provide a great playing ground for the comparison of different algorithms, as they allow for direct comparison in a controlled environment. Whereas more classical approaches like minimax require human knowledge encoded in heuristics, modern approaches like (deep) reinforcement learning have excelled solely based on knowledge gained from self-play. To test how well this generalizes for other games, in the following, this is applied to the game abalone and compared with existing algorithms in a final tournament.

1 Introduction

Abalone is a fairly new game, that was devised in 1987 by Michel Lalet and Laurent Lévi. Nevertheless, with more than four million global sales it has established itself as a classic game [1]. Abalone is a two-player game consisting of a hexagonal board with 61 fields and 14 marbles for black and white respectively. The abstract nature of the game requires the player to plan ahead and find the right strategy in the plethora of moves.

As it is a two-player perfect information game it can be solved, in theory, by backward induction. However, in search of the optimal move it is not possible to expand all of the possible paths the game could take, even for modern computers. Hence, more sophisticated approaches for navigating the search space are needed. This simple yet complex nature of this type of game make it very suitable for theoretic analysis. [6, p. 1]

Ranging from game theory to artificial intelligence, games have sparked theoretical inquiry with many real world applications in economics, psychology, mathematics, neuroscience and many more. [12, pp. 46]

2 Background

The comparison between a human's and a computer's proficiency in a game is particularly interesting as it allows for very tangible conclusions about the current state of affairs. IBM's "Deep Blue" (DB) win against Gary Kasparov [7] in chess caused big public interest just as the win of "AlphaGo" (AG) against Lee

Sedol [5] in the game Go. These two milestones also represent a shift in technology. Whereas Deep Blue relied on minimax, AlphaGo relied on a combination of multiple deep neural nets trained in self-play. [14]

Building on this success DeepMind, the company behind AG, further improved the architecture. "AlphaGo Zero" and the generalization "AlphaZero" (AZ) learn *tabula rasa*, without the help of human knowledge and surpassed the performance of AG significantly. Since then the architecture has been applied to Chess, Shogi and Atari games by removing the last piece of human knowledge in the system: The rules of the game. [13]

There is a formidable body of work regarding classical game-playing agents and some based on learning algorithms. The most significant contributions are:

1. "Algorithmic fun-abalone" (2002) Considers foundational heuristics for the game and analyzes minimax and its refinements in the form of (heuristic) alpha-beta pruning. Furthermore it sheds light on the performance differences between those. [2]
2. "Abalearn: Efficient Self-Play Learning of the game Abalone" (2003) Utilizes TD-Learning to create a self-learning agent that is on par with intermediate players. [3]
3. "A Simple Intelligent Agent for Playing Abalone Game: ABLA" (2004) Implementation of a game-playing agent with minimax, alpha-beta pruning and some custom heuristics. The evaluation of the performance is done by comparing the agent to existing software in the form of ABA-PRO and RandomSoft. [11]
4. "Constructing an abalone game-playing agent" (2005) Provides a very thorough explanation and analysis of the game's fundamentals, such as the state space, rules and positions. In regards to the alpha-beta pruning it also explains strategies for ordering the nodes and performance concerns. [9]
5. "Implementing a computer player for abalone using alpha-beta and monte-carlo search" (2009) This master thesis is a very exhaustive analysis of the game, alpha-beta pruning and Monte Carlo tree search, conferring many of the previous results. [4]

Considering the date of the publications interest into the game has cooled off. However there are some smaller student publications of the more recent years that apply Q-Learning and TD-Learning. [10] [8]

3 Aims and objectives

4 Methodology

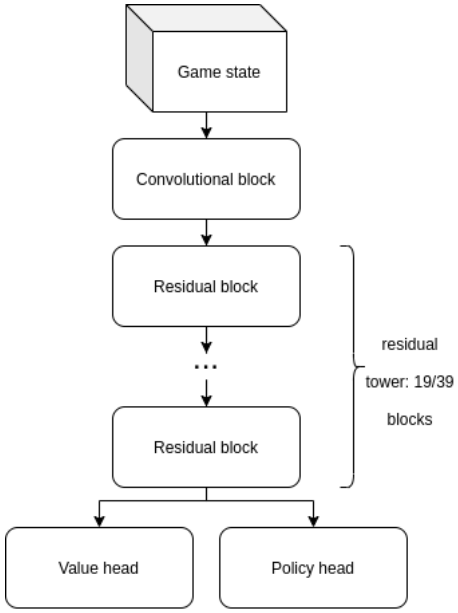


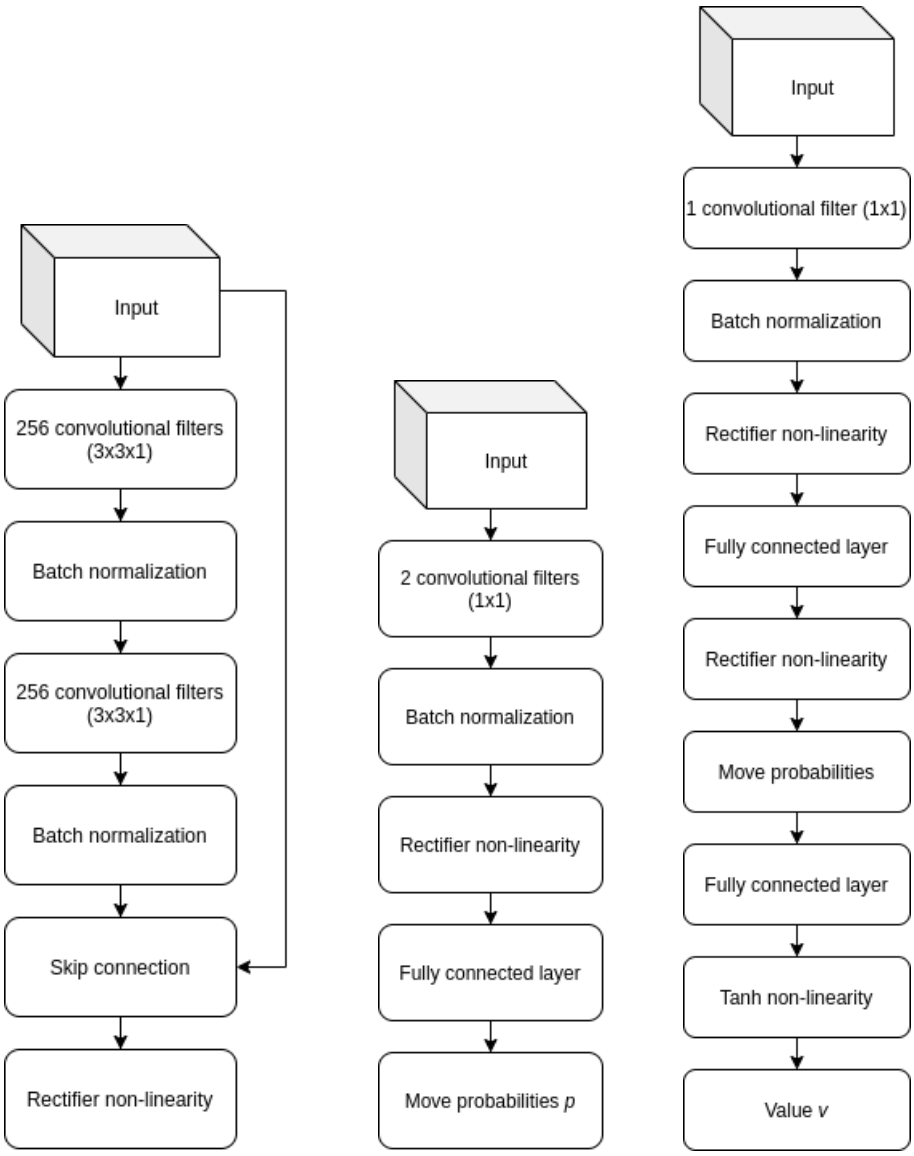
Fig. 1: The coarse architecture of AlphaZero [14]

5 Timeline

6 Preliminary table of contents

References

1. Abalone (board game). https://en.wikipedia.org/w/index.php?title=Abalone_(board_game)&oldid=994557581 (Dec 2020)
2. Aichholzer, O., Aurenhammer, F., Werner, T.: Algorithmic fun-abalone. Special Issue on Foundations of Information Processing of TELEMATIK **1**, 4–6 (2002)
3. Campos, P., Langlois, T.: Abalearn: Ecient Self-Play Learning of the game Abalone (2003)
4. Chorus, P.: Implementing a Computer Player for Abalone Using Alpha-Beta and Monte-Carlo Search. Master Thesis, Citeseer (2009)
5. DeepMind: Match 1 - Google DeepMind Challenge Match: Lee Sedol vs AlphaGo. https://www.youtube.com/watch?v=vFr3K2DORc8&t=7020s
6. Demichelis, S., Ritzberger, K., Swinkels, J.M.: The simple geometry of perfect information games. International Journal of Game Theory **32**(3), 315–338 (Jun 2004). https://doi.org/10.1007/s001820400169
7. Higgins, C.: A Brief History of Deep Blue, IBM's Chess Computer | Mental Floss. https://web.archive.org/web/20170803130439/https://www.mentalfloss.com/article/503178/brief-history-deep-blue-ibms-chess-computer (Jul 2017)
8. Lee, B., Noh, H.J.: Abalone –Final Project Report. Tech. rep. (2005)
9. Lemmens, N.: Constructing an abalone game-playing agent. In: Bachelor Conference Knowledge Engineering, Universiteit Maastricht. Citeseer (2005)



(a) Residual block (b) Policy head (c) Value head

Fig. 2: The different components of AlphaZero in detail [14]

10. Mizrachi, R., Golran, G., Jacobi, O., Zats, R.: Introduction to artificial intelligence Final Project. Tech. rep., The Hebrew University of Jerusalem (2017)
11. Özcan, E., Hulagu, B.: A Simple Intelligent Agent for Playing Abalone Game: ABLA (Jan 2004)

12. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach. Pearson Education, Inc, fourth edn. (2021)
13. Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., Lillicrap, T., Silver, D.: Mastering Atari, Go, chess and shogi by planning with a learned model. Nature **588**(7839), 604–609 (Dec 2020). https://doi.org/10.1038/s41586-020-03051-4
14. Silver, D., Schrittwieser, J., Simonyan, K., Antonoglou, I., Huang, A., Guez, A., Hubert, T., Baker, L., Lai, M., Bolton, A., Chen, Y., Lillicrap, T., Hui, F., Sifre, L., van den Driessche, G., Graepel, T., Hassabis, D.: Mastering the game of Go without human knowledge. Nature **550**(7676), 354–359 (Oct 2017). https://doi.org/10.1038/nature24270