Deep neural network design for learning Kriegspiel, an imperfect information game

Yasar Mahomed Abbas[1[0000-0003-0445-6956]], Anban Pillay[1[0000-0001-7160-6972]], Brett van Niekerk[1[0000-0003-1050-4256]] and Franziska Pannach[2[0000-0003-4216-8410]]

[1] University of KwaZulu-Natal, Westville Durban 3630, South Africa
[2] University of Göttingen, 37073 Göttingen, Germany
Yasar.tm44@gmail.com

Abstract. State of the art Artificial Intelligence (AI) systems perform well on various types of games with perfect information. However, in many real-life settings only limited information about opponents is provided. In this paper, an architecture for an agent for an imperfect information game, Kriegspiel chess, is proposed. The architecture uses Deep Reinforcement Learning and learning by self play. We encode the state of the board and information on previous moves in an 8x8x27 layered Neural-Network. In order to select the best possible action, a Deep Counterfactual Regret Value minimization algorithm is used. Neural Networks are trained using self play in a tournament setting.

Keywords: Kriegspiel, Imperfect Information, Self Play, Information State, Counter Factual Regret.

1 Introduction

In imperfect information games players only observe their information state, U_t, and generally do not know which exact game state, S_t, they are in. While the game state is the state of the entire game, e.g. all pieces on a chessboard at a particular time, the information state contains only the information that is available to a player at a given time. The player may form beliefs, $P(S_t|U_t)$, which are generally affected by fellow players' strategies at preceding states, S_k, $k<t$ [1]. Game solving agents take advantage of a concept called subgame solving. A subgame is a potential state of the game that is initiated by a specific action, thus creating a new game to be solved. Building artificial intelligence agents to play games with imperfect information is particularly difficult because an optimal strategy for a subgame cannot be determined from that subgame alone as the AI is unable to determine exactly which subgame it is in [2].

The game of Kriegspiel (German: war game) is a chess variation that limits the players' vision to their own pieces while following the same set of rules. If a player captures an opponent's piece, the opponent will learn that he or she lost a piece, but not by which piece it was captured. A referee is used to determine whether a move is allowed or not [3]. Gathering real-life data for training agents to play chess-like games is costly and prone to passing on human mistakes. A solution to these problems is using self play to gather training data by allowing the agent to play many games against itself and learning using reward and punishment. By removing the human element, the AI can achieve near perfect results as shown [4]. Libratus is an AI that can compete against professional poker players which combines game abstraction and the Counter Factual Regret (CFR) Algorithm. [5]

Developing agents that play games with imperfect information is particularly challenging as the state of the game and effect of moves cannot be predicted. This project will seek to develop artificial intelligence techniques to train agents to play such games.

2 Design

Within the Kriegspiel environment the observation state is the source of information from which the agent will determine how to interact with the board. The observation uses an 8x8x27 stack with each 8x8 layer representing positions on the game board. The first 8 layers shows the agents current information such as pieces on their respective positions. The next 20 layers are used to show previous actions played. The output layer of the Neural Network will yield a list of Counterfactual Regret Values associated with all possible moves in a given game state.

A Deep CFR algorithm will be used to perform action selection. CFR performs two main actions. First, it defines a new notion of state-action value, the counterfactual value. Then it performs a decomposed regret minimization procedure (based on these Counterfactual values) at every information state that works together to minimize an average regret, the total regret is stored in a table and the policy is constantly updated [5]. Deep Counterfactual Regret minimization uses a deep Neural Network to store the total regret, replacing the table used in the traditional CFR algorithm. This will produce the total average regret at the given state of the game. This value is passed into the CFR algorithm to produce an action, the action is played out and a new observation is read. This process is repeated until a terminal state is reached providing the players with an appropriate reward and terminating the game. Training data will be collected using Self Play on the best network in numerous iterations. After each action of each game, the observation state, action played and winner will be stored and used for retraining.

3 Conclusion

The idea behind creating game playing agents is to test the effectiveness of certain techniques in simulate environments. Since games like chess include a lot of strategic decisions, gathering training data from human players is challenging and costly. In many real-world situations there is imperfect information. Important strategic decisions need to be made without having full real-life knowledge. Regret algorithms have been used to deal with this problem. The large search space of Kriegspiel have made the standard tabular version of Counterfactual Regret minimizations impractical. This is dealt with by using a deep neural network with the architecture inspired by AlphaZero. In this project, we aim to solve Kriegspiel chess by using the OpenSpiel framework [6] to build a deep reinforcement learning architecture.

References

1. Heinrich, J.: Reinforcement Learning from Self-Play in Imperfect-Information Games. PhD thesis. University College London (2017)
2. Brown, N., Sandholm, T.: Safe and Nested Subgame Solving for Imperfect Information Games. In: Guyon, I., Luxburg U.V., Bengio S., Wallach H., Fergus R.,Vishwanathan S., Garnett R. (eds.) Advances in Neural Information Processing Systems, NIPS 2017, vol. 30 pp. 689—699. Curran Associates, Inc, San Jose, California, United States (2017).
3. Ciancarini, P., Dalla, L., Maran. F.: Decision Making Under Uncertainty: A Rational Approach to Kriegspiel. In: van den Herik, H.J., Uiterwijk, J.W.H.M. (eds.) Advances in Computer Chess, Conference Advances in Computer Chess, vol. 8, pp. 277-298 (1997)
4. Silver, D.: Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm. arXiv preprint arXiv:1712.01815 (2017).
5. Brown, N.: Superhuman AI for heads-up no-limit poker: Libratus beats top professionals, Science 359(6374), 418-424 (2019).
6. Lanctot, M., Edward, L., Jean-Baptiste, L., Vincius, Z., et al.: OpenSpiel: A Framework for Reinforcement Learning in Games, preprint, https://arxiv.org/abs/1908.09453 (2019).