RL-reversi

Cameron Humphreys
Student: 101162528 Email: CameronHumphreys@cmail.carleton.ca
Lauris Petlah
Student: 101156789 Email: laurispetlah@cmail.carleton.ca
Sukhrobjon Eshmirzaev
Student: 101169793 Email: SukhrobjonEshmirzaev@cmail.carleton.ca
Awwab Mahdi
Student: 101225637 Email: awwabmahdi@cmail.carleton.ca
Vivek Chand
Student: 101119792 Email: VivekChand@cmail.carleton.ca

1 Problem Statement

We are working on using Reinforcement Learning to teach an agent to play the game Reversi/Othello through self-play.

2 Feasibility

In this section, we will explore how the fundamental concepts of reinforcement learning (RL) can be applied to Othello, with a focus on connecting the game to abstract ideas such as agents, an environment, actions, and short-term as well as long-term rewards. This projects aims to employ RL algorithms to create intelligent agents that mimic human players capable of making strategic decisions. These agents will interact with the game's environment, an 8×8 grid by selecting legal actions governed by an adaptive algorithm, learning optimal policies and strategies that maximize cumalitive rewards. These rewards can be in the form of,

- Short-term rewards, such as capturing opponent discs, and
- Long-term rewards, such as the desired end-state of dominating the board, and winning the game.

This project's goal focuses on the different methods of using RL, including using reinforcement learning techniques such as Deep Q-Learning, Q-tables, and Self-play to train agents for Othello. The challenges include optmizing against the fact that the 8×8 sized board contains approximately 3^{64} possible states, each with an set of actions ≥ 0, which prevent the use of simple enumerative techniques and requires more sophisticated algorithms. The potential of the use of RL algorithms appears to be worthwile to explore, allowing us to master complex strategy games like Othello.

3 Milestones

In accordance with the bi-weekly schedule specified in the project requirements, we have constructed a series of milestones, interleaved with the actual project deadlines, that will allow us to consistently produce the necessary amount of work and remain on track to complete.

The first deadline has already passed, the initial project proposal.

The first milestone is the initial implementation of the OpenAI Gym environment, along with the creation of the Project Report structure and completion of the Introduction section.

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

Table 1: Milestone Dates

Part	
Date	Milestone
01/10/2023	Project Proposal
30/10/2023	Milestone 1 (Gym Implementation and Report - Introduction)
30/10/2023	Enviroment Demo
30/10/2023	Milestone 2 (Algorithm Selection and Report - Approaches)
30/10/2023	Milestone 3 (Data Collection and Report - Empirical Studies)
06/12/2023	Result Demo
10/12/2023	Project Report

The second deadline is the Environment Demo, in which we will be presenting the completed Gym environment with action and state spaces, reward structure, and step functions implemented.

The second milestone is the selection of algorithms for the Result Demo, as well as their preliminary implementation in the Gym environment. Additionally, we plan to complete the Project Report's Approaches section.

The third milestone is the experimentation and data collection after running RL agents with various algorithms, along with the completion of the Empirical Studies section of the Project Report.

The third deadline is the Result Demo, in which we will present our analysis of the results of our experiments, and how our implementation was able to solve the problem. We also aim to have the Project Report's Conclusion section done by this point.

The fourth and final deadline is for the Project Report, which at this point should be finished, with earlier milestones having taken care of each section.

References

The References are in APA style

[1] Jan van Eck, N., & van Wezel,M. (2008). Application of reinforcement learning to the game of othello. *Computers & Operations Research*, 35(6), 1999–2017. https://doi.org/10.1016/j.cor.2006.10.004

[2] van der Ree, M., & Wiering, M. (2013). Reinforcement learning in the game of Othello: Learning against a fixed opponent and learning from self-play. *2013 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*, 108–115. https://doi.org/10.1109/adprl.2013.6614996

[3] Kim, K. J., Choi, H., & Cho, S. B. (2007, April). Hybrid of evolution and reinforcement learning for othello players. *In 2007 IEEE Symposium on Computational Intelligence and Games* (pp. 203-209). IEEE. https://ieeexplore.ieee.org/abstract/document/4219044

[4] Skoulakis, I. E., & Lagoudakis, M. G. (2012, November). Efficient reinforcement learning in adversarial games. *In 2012 IEEE 24th International Conference on Tools with Artificial Intelligence* (Vol. 1, pp. 704-711). IEEE. https://ieeexplore.ieee.org/abstract/document/6495112