
COMP 579 Project Proposal: Reinforcement Learning for Ms. Pac-Man

Ivy Hu

nanqing.hu@mail.mcgill.ca
Department of Computer Science
McGill University

Simon Li

xi.yang.li@mcgill.ca
Department of Electrical Eng.
McGill University

Kenza Bellebouis

kenza.bellebouis@mail.mcgill.ca
Department of Computer Science
McGill University

1 Proposal

In this project, we propose to investigate reinforcement learning (RL) techniques in the context of the classic arcade game, Ms. Pac-Man, using the OpenAI Gym environment (MsPacman-v0). The game presents a challenging task for RL agents due to its complex state space, dynamic obstacles (ghosts), and the requirement for long-term planning to maximize score. Our project aims to implement and evaluate two RL algorithms: Rainbow DQN and Proximal Policy Optimization (PPO), and, if time permits, extend our investigation by incorporating additional variants such as the full Rainbow DQN architecture.

The motivation behind this project is to explore the strengths and limitations of value-based (Rainbow DQN) and policy-based (PPO) methods in a challenging, dynamic, and visually rich game environment. Rainbow DQN combines several improvements over the traditional DQN algorithm and is expected to provide robust performance with improved sample efficiency. On the other hand, PPO offers a more straightforward implementation with strong theoretical guarantees and ease of tuning. By comparing these two methods in Ms. Pac-Man, we seek to gain insights into how each algorithm handles high-dimensional visual inputs and complex reward structures.

Our methodology involves designing a common experimental framework in the MsPacman-v0 environment, preprocessing the game frames, and implementing both algorithms. We will tune hyperparameters and evaluate the agents using metrics such as average score, learning speed, and generalization across different game scenarios. The results will be presented through quantitative performance graphs and qualitative analysis of gameplay behaviors. This project not only aligns with our course objectives but also provides a practical exploration of modern RL techniques in a well-known benchmark game. The insights obtained may inform future research in RL algorithm selection and hybrid method development in complex environments. Our team of three will equally contribute to coding, experimentation, and analysis, ensuring a comprehensive study and detailed report. We look forward to successful outcomes.

References

References

- [1] M. Hessel, J. Modayil, H. van Hasselt, T. Schaul, G. Ostrovski, D. Silver, and R. S. Sutton, "Rainbow: Combining Improvements in Deep Reinforcement Learning," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2018.

- [2] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal Policy Optimization Algorithms,” arXiv preprint `arXiv:1707.06347`, 2017.
- [3] M. G. Bellemare, Y. Naddaf, J. Veness, and M. Bowling, “The Arcade Learning Environment: An Evaluation Platform for General Agents,” *Journal of Artificial Intelligence Research*, vol. 47, pp. 253–279, 2013.