Project Two: X Marks the Spot

Department of Computer Science,

Southern New Hampshire University

CS-370: Current/Emerging Trends in CS

Brad Jackson

August 6, 2023

X Marks the Spot - A Pirate's Tale of Deep Q-Learning in 8x8 Maze Treasure Hunting

Pathfinding in mazes is a classic problem that has been studied both in the context of human cognition and artificial intelligence (AI). While humans rely on intuition and experience to navigate mazes, AI agents can use systematic algorithms like deep Q-learning. This paper explores how deep Q-learning, implemented with neural networks, can be used to solve an 8x8 maze pathfinding problem, comparing the AI approach with human problem-solving.

When a human being is faced with solving the given 8x8 maze, they would typically start by observing the entire maze and planning a path towards the treasure. They would rely on intuition, memory, and trial and error to navigate through free and occupied cells (Kearns, Mansour, & Ng, 2002). On the other hand, the intelligent agent, representing the pirate, uses deep Q-learning to solve the problem. It initializes Q-values, explores new paths, and learns from rewards and penalties to find the optimal path (Mnih et al., 2015). The agent's learning is systematic and data-driven, leveraging the computational power of algorithms, whereas humans rely more on instinct and experience.

The intelligent agent's purpose in pathfinding is to efficiently navigate the maze to reach the treasure. It balances exploration, where it takes random actions to discover new paths, with exploitation, where it uses known information to move towards the goal (Sutton, Bach, & Barto, 2018). The optimal balance between these two depends on the training stage and the complexity of the maze. Reinforcement learning, implemented through deep Q-learning, systematically guides the agent towards the treasure by adapting to the environment and refining the pathfinding strategy (Watkins & Dayan, 1992).

For this specific game, deep Q-learning was implemented using neural networks. The neural network's architecture was designed to approximate the Q-function, representing expected cumulative rewards for actions in different states (Mnih et al., 2015). Training data consisted of experiences collected during gameplay, and the network was trained to minimize the difference between predicted and target Q-values (Lillicrap et al., 2016). Action selection was guided by the Q-values, and continuous learning allowed the agent to adapt over time. This implementation effectively combined the power of neural networks with reinforcement learning to handle the complex 8x8 maze pathfinding problem.

The comparison between human and AI approaches to maze pathfinding highlights the unique strengths and methodologies of each. While humans leverage intuition and experience, AI agents like the one described in this paper use systematic, data-driven methods like deep Q-learning. The successful application of deep Q-learning to the 8x8 maze problem demonstrates the power and flexibility of modern AI techniques. It also underscores the potential for further research and development in the field of reinforcement learning, particularly in complex environments where human intuition may fall short.

References

Kearns, M., Mansour, Y., & Ng, A. Y. (2002). *Machine Learning*, 49(2/3), 193–208.

https://doi.org/10.1023/a:1017932429737

Lillicrap, T. P., Hunt, J. J., Pritzel, A., Heess, N., Erez, T., Tassa, Y., Silver, D., & Wierstra, D.

(2019, July 5). *Continuous control with deep reinforcement learning*. arXiv.org.

https://doi.org/10.48550/arXiv.1509.02971

Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). *Human-level control through deep*

reinforcement learning. Nature, 518(7540), 529–533. https://doi.org/10.1038/nature14236

Sutton, R. S., Bach, F., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT

Press Ltd.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine Learning*, 8(3–4), 279–292.

https://doi.org/10.1007/bf00992698