**Project Two Design Defense**

Tyler Young

Southern New Hampshire University

CS-370: Current/Emerging Trends in CS

Dr. Obafemi Balogun

October 19th, 2025

In my implementation, I focused on how humans and machines learn and solve problems differently. A human would solve the maze by looking at it, thinking ahead, and using logic to decide which direction to move. They can often learn more from a single mistake or observation and can remember patterns easily, which makes their learning process faster and more flexible. Research from the University of Oxford shows that the human brain learns more efficiently than artificial intelligence because it uses a process called prospective configuration, where neurons predict outcomes instead of constantly adjusting connections like AI does through backpropagation (Stacy, 2024). The AI pirate agent, on the other hand, learns slowly through repetition. It must make many attempts, receive rewards or penalties, and adjust its behavior based on these results (Nuer, 2023). While both humans and the agent learn from trial and error, humans can adapt faster and retain knowledge better, whereas the agent depends entirely on repeated feedback to form an effective strategy.

The goal of the intelligent agent is to find the treasure on its own without being told how to do it. To learn effectively, it must balance exploration, which is trying new paths, and exploitation, which is using the paths that have worked well before. At first, the agent explores often to understand the maze, but as it learns more, it shifts toward exploitation, focusing on moves that lead to success, or finding the treasure in this case. This balance is managed using the epsilon-greedy strategy, which starts with random actions and slowly becomes more selective over time (Nuer, 2023). The best balance allows the agent to learn from mistakes early on while later relying on what it has discovered. Reinforcement learning supports this by giving the agent feedback after each move, rewarding progress and punishing poor decisions, which helps it find the most efficient path to the treasure.

I implemented deep Q-learning by using a neural network that takes the maze as input and outputs values for each direction the pirate can move. These values show how good each move is in the current situation. The model uses experience replay to train on past experiences, which helps prevent bias, and a target network that updates less often to keep training stable (Nuer, 2023). Over many training rounds, this allows the pirate to learn how to reach the treasure efficiently. While this method is much slower and less adaptive than how the human brain learns, it still shows how artificial intelligence can gradually solve complex problems through feedback and repetition (Stacy, 2024).

**References**

Nuer, J. (2022, May 20). A Comprehensive Guide to Deep Q-Learning. Medium. <https://medium.com/@jereminuerofficial/a-comprehensive-guide-to-deep-q-learning-8aeed632f52f>

Stacy, K. (2024, January 5). Human brain learning is more efficient, superior compared to artificial intelligence, scientists say. *Science Times*. <https://www.sciencetimes.com/articles/48017/20240105/human-brain-learning-more-efficient-superior-compared-artificial-intelligence-scientists.htm>