**Kain Mason**  
**Southern New Hampshire University**  
**CS 370: Project Two Design Defense**  
**June 22, 2025**

**Design Defense: Pirate Intelligent Agent**

**Human Versus Machine Problem-Solving in Maze Navigation**

When a person tries to solve a maze, they often go with instincts, memory, and general problem-solving habits—like keeping a hand on one wall or following visual clues. People can adjust on the fly, take shortcuts, or guess based on previous knowledge or emotions. According to Pew Research Center (2023), human thinking mixes logic with feelings and life experience, which machines don’t really understand or replicate yet.

On the other hand, the pirate agent in this project works purely off data and logic. It doesn’t guess or feel; it learns by doing repeatedly. Through deep Q-learning, the pirate moves around the maze, tries different paths, and gradually figures out which moves lead to success based on rewards. This trial-and-error learning is key to how intelligent agents like this improve over time (OpenAI, n.d.). Everything in the learning process came from the framework built into the course files provided by Southern New Hampshire University (2025).

**Why the Pirate Agent Was Built and How It Works**

The main point of building this pirate agent was to give it a way to reach the treasure on its own—without help from a human. The tricky part is finding the best path through the maze. Reinforcement learning fits perfectly here because it teaches the agent to learn from its own actions and mistakes.

At the start, the pirate explores the maze randomly some of the time (10%) using an epsilon value of 0.1. This gives it a chance to find new routes. As it starts learning, the epsilon drops to 0.05, so the pirate relies more on what it has learned already. This method, starting with more exploring and slowly shifting to smarter choices, is one of the go-to strategies in reinforcement learning (OpenAI, n.d.).

**Using Deep Q-Learning to Train the Pirate**

The way this pirate figures out what to do is by using deep Q-learning. A neural network helps it decide which move to take by assigning values (Q-values) to every possible action from its current spot in the maze. The higher the value, the better the move.

Instead of learning only from its most recent step, the agent also remembers past experiences using something called experience replay. It randomly pulls from this memory to train, which helps avoid bias and makes learning smoother and more stable. Jensen et al. (2019) explain that having a mix of different past experiences—like how humans reflect—leads to smarter decisions and better learning overall.

Once trained, the pirate was able to get through the maze from different starting points, proving that the model worked well and the learning process paid off.

**References**

Jensen, K. T. J., Alayrac, J.-B., Veličković, P., Zisserman, A., & Blundell, C. (2019). Neural scene representation and rendering. Trends in Cognitive Sciences, 23(10), 721–734. https://www.cell.com/trends/cognitive-sciences/fulltext/S1364-6613(19)30061-0

OpenAI. (n.d.). Research. https://openai.com/research/index/

Pew Research Center. (2023). Internet & Technology. https://www.pewresearch.org/topic/internet-technology/

Southern New Hampshire University. (2025). Treasure Hunt Game project files: TreasureHuntGame.ipynb, TreasureMaze.py, GameExperience.py [Course materials]. CS-370: Artificial Intelligence.