**Thomas Martin**

**CS-370-11456-M01 Current/Emerging Trends in CS**

**7-3 Project Two Submission**

**Southern New Hampshire University**

**August 16, 2024**

**Design Defense for the Treasure Hunt Game**

The Treasure Hunt Game is a pathfinding AI project where players compete to find a treasure before an AI-controlled pirate agent. The pirate uses a deep Q-learning algorithm to learn the optimal path to the treasure from any starting position on the map. In Project Two, we were tasked with training this pirate agent to navigate a maze and reach the treasure chest at the end. This project underscores the differences between human and machine approaches to problem-solving.

When humans approach decision-making and problem-solving, they are often influenced by cognitive biases—systematic deviations from rationality that can lead to suboptimal choices, especially in complex situations where the solution isn't immediately apparent. When solving a maze, humans typically rely on visual cues, intuition, and past experiences. They may start by identifying potential paths, avoiding obvious obstacles, and using trial and error to refine their approach based on successes and failures.

In contrast, an intelligent system, such as the AI pirate agent, is guided by principles that shape its decision-making process. The agent balances exploration—testing new actions to uncover their effects—and exploitation—choosing actions that have previously led to success. Using a deep Q-learning algorithm, the agent refines its strategy over time, maximizing its chances of reaching the treasure. This approach ensures that the AI continuously learns from its experiences, leading to increasingly efficient pathfinding.

The key difference between human and machine learning processes is that machines learn systematically through reinforcement—updating their understanding based on rewards—while humans rely more on intuition and heuristics. Machine learning is incremental and driven by the reinforcement of positive outcomes, enabling continuous adaptation and improvement, especially in repetitive and data-intensive tasks. However, humans excel in applying prior knowledge and quickly adapting to new or unstructured situations, which gives them an edge in scenarios where flexibility and experience are crucial.

Reinforcement learning enables the agent to find the optimal path to the treasure by learning from its experiences. Each time the agent navigates the maze, it updates its Q-values based on the rewards received, gradually improving its decision-making. Over time, the agent becomes adept at predicting which actions are likely to lead to the treasure, allowing it to navigate the maze more efficiently and consistently reach the goal.

Balancing exploration and exploitation is crucial in reinforcement learning. In the treasure maze context, if the pirate agent only exploits known paths, it may get stuck in suboptimal routes, missing out on potentially better options. Therefore, the Q-learning algorithm, implemented using a neural network within the Keras environment, plays a vital role. The maze is represented as a grid, where each cell corresponds to a specific state, and the pirate can move in four directions: left, right, up, and down. The reward system provides positive feedback for reaching the treasure and negative feedback for hitting walls or losing. The Q-values for each state-action pair are approximated using a neural network, with the current state as the input and the Q-values as the output. During training, the agent plays multiple games (epochs), updating its Q-values based on the rewards received. Training continues until the agent achieves a consistently high success rate, such as a 90%-win rate.

This project successfully demonstrated how deep Q-learning can be used to train an intelligent agent to solve complex pathfinding problems. By effectively balancing exploration and exploitation and utilizing a neural network for Q-value approximation, the agent learned an optimal strategy for navigating the treasure maze.