Design Defense – Pirate Treasure Hunt Agent

Jaden Knutson

Southern New Hampshire University

CS 370: Current and Emerging Trends in Computer Science

Professor Travis Smith

June 2025

**Introduction**

In this project, I am working as an AI developer for a gaming company. The company is creating a treasure hunt game. For this, I developed an NPC pirate agent tasked with finding treasure in a maze before a human player. The agent uses a deep Q-learning algorithm to learn the optimal path through trial and error. Overall, this project is a pathfinding problem that is solved using reinforcement learning and neural networks (GeeksforGeeks, 2024b; DataHeroes, 2023). The purpose of this paper is to show how the agent functions, compare its behavior to how a human might solve the maze, and evaluate the reinforcement learning strategy used in the implementation.

**Human vs. Machine Problem-Solving**

**Steps for Humans Solving a Maze:**  
Humans go about solving mazes by trying potential paths, backtracking, and using trial and error, as well as identifying landmarks or memorable parts. Humans also use this with a mixture of intuition, and spatial memory/reasoning to create mental maps or possible strategies like trying to keep to the left wall until getting blocked or finding the end (de Cothi et al., 2022; Liu, 2021).

**Steps for an Agent Solving a Maze:**   
An agent like our pirate uses reinforcement learning, trial and error, repetition, actions, rewards, and adjustments to its neural network in order to enhance and improve its path-finding efficiency. From the start, the agent doesn’t necessarily understand the maze itself, it improves based on its rewards, and finding whatever the end goal is to increase its reward value at the end (GeeksforGeeks, 2024b; Liu, 2021).

**Similarities and Differences:**  
Both methods involved with human and machine learning involve trial and error and learning from past experiences. For differences, the agent doesn’t visualize or reason. Instead it statistically learns action values through repeated episodes and optimization (GeeksforGeeks, 2024b; Liu, 2021).

**Purpose of the Intelligent Agent and Exploration vs. Exploitation**

**Differences Between Exploitation and Exploration:**  
The goal of our pirate agents is to overall find the treasure with the greatest amount of points in the end. It explores in order to find good and bad paths, and then takes advantage of that by exploiting the known successful strategies. Exploration is when the agent explores, trying random moves to discover new strategies. Exploitation is when it uses the knowledge and strategies it knows to implement good moves to find the treasure and maximize rewards (GeeksforGeeks, 2024b; GeeksforGeeks, 2024c).

**Ideal Proportion of Exploitation vs. Exploration:**  
In my version, I used an epsilon greedy strategy in order to have it take random actions at the start but then choose the best actions as it advances. As its performance improves and reaches a 90% win rate, the epsilon is reduced along with the exploration and starts to focus harder on exploitation. This overall makes the agent start by exploring and discovering good paths and then use them to find the treasure later (GeeksforGeeks, 2024b; GeeksforGeeks, 2024c).

**Using Reinforcement Learning to Help the Agent Reach the Goal:**  
Using reinforcement learning with the agent allows it to learn from interactions with its environment by receiving rewards and penalties through each of its actions (GeeksforGeeks, 2024b). The neural network updates throughout the process to predict better future actions based on its rewards throughout past actions and playthroughs (GeeksforGeeks, 2024a; DataHeroes, 2023).

**Deep Q-Learning and Neural Networks for Complex Problems**

For this project, I implemented a deep Q-learning algorithm that uses a neural network to estimate the Q-Values (GeeksforGeeks, 2024b; GeeksforGeeks, 2024a). The training process consists of observing the current state of the maze, selecting an action for the agent to take, receiving a reward or penalty for the action, storing that experience, using samples from past experiences to help train the model, and then updating its predictions to enhance its decisions for future playthroughs and action. These steps allow the agent to generalize and improve its performance from past experiences (GeeksforGeeks, 2024a; DataHeroes, 2023).

**Conclusion**

Using a combination of memory, deep learning, exploration, and exploitation, the pirate has learned to solve the maze without having prior knowledge or logic. Reinforcement learning gave the agent the ability to adapt its behavior and learn which showcases an approach that is suited for this machine's learning and optimization in solving the maze (DataHeroes, 2023).

**References:**

DataHeroes. (2023, October 9). What is neural network. https://dataheroes.ai/glossary/neural-network/

GeeksforGeeks. (2024, July 19). Layers in artificial neural networks (ANN). https://www.geeksforgeeks.org/layers-in-artificial-neural-networks-ann/

GeeksforGeeks. (2024, September 29). Types of Reinforcement Learning. <https://www.geeksforgeeks.org/machine-learning/types-of-reinforcement-learning/>

GeeksforGeeks. (2024, May 18). *Exploitation and Exploration in Machine Learning*. GeeksforGeeks. https://www.geeksforgeeks.org/machine-learning/exploitation-and-exploration-in-machine-learning/#

de Cothi, W., Nyberg, N., Griesbauer, E.-M., Ghanamé, C., Zisch, F., Lefort, J. M., Fletcher, L., Newton, C., Renaudineau, S., Bendor, D., Grieves, R., Duvelle, É., Barry, C., & Spiers, H. J. (2022). Predictive maps in rats and humans for spatial navigation. *Current Biology*, *32*(17), 3676-3689.e5. <https://doi.org/10.1016/j.cub.2022.06.090>

Liu, J. (2021). *Comparing Human and AI Behavior in 3D Navigation Environments*. https://www2.eecs.berkeley.edu/Pubs/TechRpts/2021/EECS-2021-111.pdf