Project Two Design Defense

Jeffrey Sparks

CS 370

Prof Dewin Andujar

Southern New Hampshire University

First, we need to understand the base function and idea for the game that is being developed. The goal of the game is to find the treasure before the Pirate AI character, at the same time receiving the maximum reward available and winning the game. To do this a human would need to evaluate the environment, observe obstacles as they come into contact with them and make a decision on where to move next based on where they are currently located versus where they came from in hopes navigate the maze and obtaining the treasure before the pirate is able to.

To solve the maze both characters must make the decision to move in one of four directions, up, down, left, or right. If a player returns to a spot (retraces their steps) they will lose points, the goal is to move to the treasure using the quickest path possible without returning to previously visited cells within the maze. Much akin to the human progress, the system controlling the pirate learns using a reinforcement method; by mimicking the trial-and-error processes that a human goes through with learning through previous mistakes and building upon what it has experienced within past runs of the maze until it understands how to complete the maze with 100% certainty (*What Is Reinforcement Learning? - Reinforcement Learning Explained - AWS*, n.d.).

Within this type of machine learning, reinforcement learning relies on two main concepts in order to train the system for the game, exploitation and exploration. Exploration allows the pirate the opportunity to explore and navigate the maze freely, learning new things by exploring the environment in a new way than previously experienced; where as exploitation accesses the data from previous runs, analyzes and builds a better path forward based on what the system has experienced within the past (*What Is the Exploration vs Exploitation Trade off in Reinforcement Learning?*, n.d.). With this in mind the ideal within the code is to lower the exploration factor, by lowering the exploration factor within the pirate game, it allows the system to learn based on prior experience while not depriving the much needed exploration to further the progress within finding the path to the treasure, by focusing on exploitation, it allows the system to keep track of where it has been, what paths it has chosen to take in the past, and the end result of when it has reached the treasure with the maximum amount of reward.

In order to implement a deep Q-learning when using a neural network for this game, I used a few steps, the first of which involved importing the libraries needed to perform the operations required. Next, the creation of the training environments, followed by creating a reward system. Then the learning agent was instantiated and used enhanced algorithms for learning. Finally, the agent was tested within the environment. Going through the implementation steps for deep Q-learning when using a neural network helps to find the most optimized movement sequence to navigate the map and reach the treasure cell by maximizing the reward.

References:

What is Reinforcement Learning? - Reinforcement Learning Explained - AWS. (n.d.). Amazon Web Services, Inc. <https://aws.amazon.com/what-is/reinforcement-learning/#:~:text=Reinforcement%20learning%20(RL)%20is%20a>

What is the exploration vs exploitation trade off in reinforcement learning? (n.d.). Scribbr. https://www.scribbr.com/frequently-asked-questions/what-is-the-exploration-vs-exploitation-trade-off-in-reinforcement-learning/#:~:text=Exploration%20is%20any%20action%20that