Taylor Stapus

CS 370

Module 7

04/17/2025

**Project Two: Design Defense**

The treasure hunt game uses AI to represent a non-player character to play a pirate that has the goal to find the treasure before the human player does. This is commonly known as a path-finding problem as the pirate created will need to find the most optimal path towards its end goal. If a human were playing, the steps they would take to solve this maze would include making a choice to go left, right, up or down. They would determine their potential steps by getting a sense of the maps layout and determining if there are any obstacles or dead ends. For every step that they take, they would record where they have already been and use that information to progress forward. AI would approach the situation a little differently by taking steps that, for the most part, are already predetermined through an algorithm. Learning-based algorithms such as Deep Q-Learning use neural networks to learn the value of an action through trial and error. It would then react to the feedback acquired by taking such an action and update its internal model to continue. While both ‘players’ are essentially ‘making a decision’, it is important to see that humans make decisions based on intuition and visual knowledge while AI uses explicitly coded goals and algorithms.

Exploration and exploitation are key concepts in machine learning. Exploitation uses information that is already known, and chooses the best-known option based on the current knowledge. The goal is to maximize immediate performance or reward. Exploration means trying out new options to discover better choices for the future. When applied to a model, the algorithm will choose an option that is less certain about improving the long-term learning potential that will hopefully lead to better outcomes. Balancing exploration and exploitation are crucial for the learning functions to excel. If one option is prioritized too much over the other, there runs a risk of missing out on a better option or not making use of knowledge already acquired. The ideal balance between the two for Treasure Hunt Game is that for every tenth step, the AI will explore a new path. The goal of this balance is to remain consist in utilizing the information already available while also allowing brief exploration encounters to maximize long term success. The goal of reinforcement learning is to determine a policy by recording the actions, rewards and penalties, taken in an environment through trial and error. Over time, that knowledge is used to enhance the decision-making of the model to complete its goal faster and more efficiently.

The Treasure Hunt Game implements deep Q-leaning to find the best possible navigation sequence that will result in finding the treasure while also maximizing the reward. Deep Q-learning uses deep neural networks to approximate ethe q value. The idea is that the algorithm will store the pirates’ experience and then sample them randomly to break correlation and improve learning stability. Every step that the pirate takes is recorded and then utilized to navigate the maze more efficiently than next time playing. With experience, the AI should be able to find the treasure in the least number of steps possible by avoiding obstacles and making unnecessary moves.