**Armando Gomez**

**CS-370-10993-M01 Current/emerging Trends in CS**

**7-3 Project Two:**

**Design Defense**

**Professor Jason Richard**

**August 17, 2024**

* **Analyze the differences between human and machine approaches to solving problems.**
  + **Describe the steps a human being would take to solve this maze.**
* When a human faces a maze, they typically start by looking at the layout to figure out where they can go and what might block their way. They would then. Remember to choose the path that seems most promising for finding the treasure, considering what they see in front of them. If they hit an end or obstacle, they would go back and try another way. This approach involves using skills like recognizing patterns, understanding space and making decisions based on encounters.
  + **Describe the steps your intelligent agent is taking to solve this pathfinding problem.**
* The agent navigates through the maze by employing the Q learning method, a form of reinforcement learning. At the start, the agent has no understanding of the surroundings and begins exploring the maze by making random moves. Each move leads to either a reward or a penalty. The agent keeps track of these outcomes to enhance its understanding of the environment. With time, the agent grasps how to anticipate moves based on past encounters, refining its strategy to maximize overall rewards. This procedure entails utilizing a network to estimate Q values, which indicate the anticipated future rewards for every potential action from any given state within the maze.
  + **What are the similarities and differences between these two approaches?**
* Both human and machine tackle challenges by exploring and drawing insights from previous encounters. Yet humans rely on intuition, logic and mental shortcuts to make decisions with limited information. On the other hand, the intelligent agent strictly follows systematic exploration, mathematical computations and policy updates guided by reinforcement learning principles. While humans may navigate mazes quicker through shortcuts, the agent method guarantees discovering the best solution by thoroughly examining all options.
* **Assess the purpose of the intelligent agent in pathfinding.**
  + What is the difference between exploitation and exploration? What is the ideal proportion of exploitation and exploration for this pathfinding problem? Explain your reasoning.
* When an agent is exploiting, it uses what it knows to make choices that bring rewards. On the other hand, exploration involves trying out new actions that haven't been tested yet in the hopes of achieving better long-term results. In a maze scenario, finding the balance between exploitation and exploration depends on how complex the maze is. Initially more exploration is needed to learn about the environment. As the agent becomes more knowledgeable, it should lean towards exploitation to improve its strategy based on what it has learned. A common approach is to start with exploration and then gradually reduce it, allowing the agent to explore new paths while also refining its existing ones.
  + **How can reinforcement learning help to determine the path to the goal (the treasure) by the agent (the pirate)?**
* Reinforcement learning assists the agent in finding the way to the treasure by enabling it to gain knowledge from its interactions with its surroundings. The agent gets responses through rewards or punishments depending on its decisions, which it utilizes to adjust its strategy—a connection between states and actions. Gradually, by maximizing rewards, the agent grasps the best series of actions that guide it to the treasure. Through a process of trial and error, the agent can uncover the effective route even in a challenging setting such as a maze.
* **Evaluate the use of algorithms to solve complex problems.**
  + How did you implement deep Q-learning using neural networks for this game?
* A neural network was used to estimate the Q values for actions in different states through deep Q learning. The network underwent training with experience replay, a method that involves storing experiences and randomly selecting them to update the network. This process helps reduce the correlation between experiences, thus enhancing the stability of learning. The Q learning algorithm utilizes the network to forecast Q values, adjusting these forecasts based on received rewards and anticipated future rewards. Through training sessions, the network becomes proficient at accurately predicting Q values, enabling the agent to make decisions that optimize its total expected rewards and ultimately find the best route to uncovering the treasure.

**References**

**Achieving Human Level Competitive Robot Table Tennis. (2024, August 5). Google DeepMind. https://deepmind.google/research/publications/107741/**

**Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement Learning, Fast and Slow. *Trends in Cognitive Sciences, 23*(5), 408–422.** [**https://doi.org/10.1016/j.tics.2019.02.006**](https://doi.org/10.1016/j.tics.2019.02.006)

**Internet & Technology. (2024, June 12). Pew Research Center. https://www.pewresearch.org/topic/internet-technology/**