CS 370 Project Two: Design Defense

**Analyze the differences between human and machine approaches to solving problems.**

In the context of the maze, a human’s approach to solving a problem includes analyzing their environment, and randomly taking paths that either result in dead ends or continuing down the path that eventually leads you out of the maze. During this process, a human will rely on their memory with visual cues to successfully lead them down paths that lead them out of the maze. As a human continually engages with their environment, they will continue to identify specific landmarks and begin to understand the layout of the maze. It becomes an iterative process, where someone is continuously refining their strategy based on previous experiences of false paths and then determining which are the correct ones.

In contrast to how a human will solve the maze, the pirate-intelligent agent in this example, will rely on a Deep Q-learning algorithm. The agent begins with zero knowledge of its environment and begins interacting with it by gradually reevaluating its state. The maze is made up as a grid of states, where each state is represented as a position in the maze. As the agent is in a particular state, it considers all of the possible moves it can make, either up, down, left, or right and calculates the expected reward for each action, or the (Q-value) (GeeksforGeeks, 2023). The agent iteratively refines its policy, or strategy, by continually updating the Q-values based on the cumulative rewards from each move, ultimately discerning the optimal path through the maze.

While both approaches are uniquely tied to understanding and learning from their environment, what differs between each is how they determine their strategy. For instance, a human will rely on memorizing their environment, trial and error, basic human intuition, and logical decision-making. While this approach may help in solving complex problems, it can be time consuming and prone to errors, while also not taking into consideration the emotional factor of stress induced anxiety. On the other hand, a machine will systematically deduce the next logical decision through exploration, leading to an optimal decision-making process that yields successful completion of the maze. A machine also primarily relies on mathematical functions and equations to determine its output, which can be more efficient at determining what the correct solution is.

**Assess the purpose of the intelligent agent in pathfinding.**

In the context of the pirate agent, exploitation refers to the pirate’s ability to use its current knowledge to determine the best choice of action, while exploration involves an agent’s potential to discover new strategies (GeeksforGeeks, 2024e). For a pirate intelligent agent to solve a complex problem like the maze, it needs to combine and balance both strategies in order to effectively determine an optimal path. Over reliance on exploitation for example, may result in the agent missing out on paths that present better strategies at solving the maze. On the other hand, excessive exploration can cause the agent to wander excessively and not make significant progress.

To be able to have the pirate intelligent agent strike a balance between utilizing exploitation and exploration, it will prove crucial to have the agent focus on exploration initially until enough information is collected about its environment. As the agent is continually discovering paths through the maze, introducing exploitation would greatly benefit its ability to develop strategies that enhance its potential reward, since it relies on current knowledge. One of the strategies that is used to maintain a balance is the Epsilon-Greedy algorithm, which has the agent choose an action randomly with a probability of epsilon and “continuing to use the current best-known action with probability (1-epsilon) (GeeksforGeeks, 2024e). Essentially, the algorithm is used to encourage a balance between the two, then gradually reducing the epsilon to encourage higher focus on exploitation for maximum gain.

Reinforcement learning helps the pirate determine the path to the goal by learning from the outcomes of its actions. For instance, the pirate learns by associating a specific action with either a positive or negative reward, depending on the outcome. Eventually, it begins to learn from and associate these rewards and develops an optimal policy or strategy that helps it to find the treasure. As the pirate or agent continues through the maze, it continuously refines its policy and begins focusing on actions that lead to higher rewards.

**Evaluate the use of algorithms to solve complex problems.**

To enable the pirate agent to effectively reach the treasure and solve the maze, I implemented deep Q-learning, an advanced reinforcement learning technique. By also integrating neural networks to help approximate Q-values for each state-action pair, the agent can accurately assess the long-term rewards associated with its actions. The Q-value represents the cumulative reward expected from a specific action in a given state, guiding the agent toward the most optimal path through the maze. As the agent explores its environment, the Q-values are iteratively updated, allowing the agent to refine its strategy over time. I also introduced techniques like experience replay, critical to deep Q-learning. Experience replay allows the agent to store past experiences, including the “state, action, reward, and next state in a memory buffer” (GeeksforGeeks, 2023b). Which helps to improve the learning process by making it more stable and efficient by including random sampling during training.

In the context of the Treasure Hunt game, these combined methods enable the pirate agent to learn and adapt to the maze, and eventually develop a robust strategy for navigating towards the treasure. The pirate agent becomes increasingly more proficient at choosing the most rewarding paths and handling complex pathways to ultimately determine the best route. The use of deep Q-learning algorithms and neural networks allows the pirate agent to handle complex and challenging scenarios, while also demonstrating the effectiveness of advanced algorithms at solving problems. This ultimately showcases how the pirate intelligent agent can be trained to outperform traditional problem-solving methods.

References

GeeksforGeeks. (2023, January 23). *Deep Q-Learning*. GeeksforGeeks. <https://www.geeksforgeeks.org/deep-q-learning/>

GeeksforGeeks. (2023b, January 23). *Deep Q-Learning*. GeeksforGeeks. https://www.geeksforgeeks.org/deep-q-learning/

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