7-3 Project Two

**1. Differences between human and machine approaches to solving problems**

A human navigating the pirate maze would rely on trial and error, memory, and pattern recognition to explore the environment. Humans typically identify possible paths, backtrack when they get stuck, and use prior knowledge of similar environments. This aligns with the concept that “humans are great ‘one-shot learners’” (Fernandez, 2019), able to make decisions based on minimal input, such as seeing a path once and recalling its pattern.

The intelligent agent uses deep Q-learning to solve the maze. It “learns through trial and error” by interacting with the environment, receiving feedback, and updating its understanding of the best paths (Surma, 2018). The agent uses a combination of exploitation, learned information, and exploration, testing new paths, to improve its navigation strategy over time.

The humans and the intelligent agent rely on feedback loops to refine their approaches. However, while a human might quickly generalize from a few experiences and apply pattern recognition, the agent requires thousands of experiences to form a robust strategy; meanwhile, the agent might be “myopic,” focusing on immediate gains without fully understanding the bigger picture (Fernandez, 2019; Surma, 2018).

**2. The purpose of the intelligent agent in pathfinding**

Exploration involves the agent “randomly choosing actions” to discover new paths and gather more information about the environment, while exploitation relies on using the already known information to select the best possible action (Lamba, 2018). The ideal balance for this pathfinding problem is to start with more exploration (higher epsilon) and shift towards exploitation as the agent gains confidence in estimating the best actions. This approach allows the pirate to learn the maze effectively without getting stuck in suboptimal paths.

Reinforcement learning helps the pirate determine the path to the treasure by continuously updating the Q-values in the Q-table, using the Bellman equation. Each action is based on the immediate reward and expected future rewards, allowing the agent to “maximize the expected reward by selecting the best of all possible actions” (Lamba, 2018). Over time, as the Q-table is refined, the agent becomes better at navigating the maze and reaching the goal efficiently.

**3. The use of algorithms to solve complex problems**

In the pirate game, deep Q-learning was implemented using a neural network model that is trained to predict the optimal action at each state based on past experiences stored in a memory. The neural network learns by updating Q-values repeatedly, where the expected future rewards for each state-action pair are adjusted based on the outcomes (Lamba, 2018). This process uses exploration and exploitation, gradually refining the Q-values in the table until the agent can consistently find the shortest path to the treasure. As training progresses, the agent’s strategy shifts from exploration to exploitation, allowing it to “exploit the environment and start taking better actions” which guides it towards the target while minimizing penalties (Lamba, 2018).

References

Fernandez, E. (2019, November 30). AI is not similar to human intelligence. Thinking so could be dangerous. Forbes. <https://www.forbes.com/sites/fernandezelizabeth/2019/11/30/ai-is-not-similar-to-human-intelligence-thinking-so-could-be-dangerous/#3fbd0b846c22>

Lamba, A. (2018, September 3). An introduction to Q-Learning: Reinforcement learning. Medium. <https://medium.com/free-code-camp/an-introduction-to-q-learning-reinforcement-learning-14ac0b4493cc>

Surma, G. (2018, September 26). Cartpole - Introduction to Reinforcement Learning (DQN - Deep Q-Learning). Medium. <https://gsurma.medium.com/cartpole-introduction-to-reinforcement-learning-ed0eb5b58288>