Humans and machines have many similarities and differences in how they learn. Humans learn based off experiences, emotions, and interactions. A human then takes what it has learned and uses that knowledge to make decisions. An artificial intelligence uses a preset data set to learn from. An AI learns from this data set and does not have any knowledge outside of this data. AI does not use emotion to interact with a problem. It strictly uses logic to solve problems. That leads to the larges difference between AI and humans. Humans use emotions to drive decisions just as much as they do logic.

In this experiment, the goal is to get a pirate to the treasure by navigation through a maze. The maze has a starting point, an ending point, black squares that cannot be passed through and white squares that can. The maze can be altered to have a different starting point, ending point, and order of black and white squares. The AI agent’s goal is to solve the maze in as few steps as possible and be able to solve the maze no matter how the different aspects of it are arranged. To do this, the AI will be using reinforcement learning.

A human would view the maze and preplan a route through the maze to reach the goal due to prior knowledge about how mazes are to be solved. An AI goes through a point-based trial and error process. The AI test different possible direction choices and is rewarded points based on the direction they choose. The AI uses this system to know when it gets to the goal and when it hits a black square in the maze. It uses this information to be able to solve the maze and get the pirate to the treasure. These two processes both lead to the same goal but have vastly different methods to accomplish the task. Humans can look at the whole picture, but AI must go task by task.

The purpose of the intelligent agent in pathfinding is to find a balance between exploration and exploitation. Exploration is the idea that a riskier decision now can lead to a greater reward in the future and exploitation is the idea that instant satisfaction is the way to go. In AI, the agent must decide if it wants to choose a direction based on its previous knowledge of what the correct decision is or if it wants to take a risk and go a way its knowledge does not agree with. This decision-making process is the basis for reinforcement learning. Reinforcement learning provides an agent with the ability to tell if the risk was worth the reward and allows it to make a more logical decision in the future when faced with similar circumstances.

Deep Q learning uses a neural network as the memory for the AI agent. The AI agent stores q values in a neural network and updates the values as it finds a better, more efficient path to the goal, or treasure in this experiment. This memory process allows the agent to make decisions based on past experiences in the maze and allows it to make a more logical approach to solving the maze.

Works Cited

Rocca, J. (2021, May 23). *The exploration-exploitation trade-off: Intuitions and strategies*. Medium. https://towardsdatascience.com/the-exploration-exploitation-dilemma-f5622fbe1e82