PROJECT 2: DESIGN DEFENSE

When you go to solve a maze, there are multiple ways to approach it. Some may simply put a hand on the outside wall and just walk while keeping contact with the wall and eventually make their way to the exit, another may deploy a strategy of only making left turns. Still others may just pick random paths using a system of trial and error. “[T]rial-and-error learning connotes attempts at meeting the situation in various ways until the correct responses are found more or less accidentally.” (Mondal, 2013) With the intelligent agent, they are following a laid out algorithm that allows for little trial-and-error, and is more concentrating on a set of given steps to follow. While both a human and the intelligent agent can use multiple ways to solve a maze, the intelligent agent is definitely using a more methodical approach where it is able to assess it’s position on the board and through the steps find the best possible solution to solving the maze.

When you look at the possible ways to solve the maze using the intelligent agent, you also want to consider the differences between exploitation and exploration. First, lets define each. Exploitation is making decisions based on what has already proven to work in the past, where exploration is making decisions based on the possibility of a greater reward. (Lindwurm, 2021) Although both exploration and exploitation were used for the intelligent agent, finding that balance is definitely necessary. Allowing to much exploration can take your agent more tries to get through the maze successfully with a higher chance of failure. However, to much exploitation will see your agent finishing the maze on a less optimal path. A good balance would use mostly exploitation (I mean we are having the machine learn for a reason) with a little exploration to make sure that our path is as optimal as possible, while completing the task. Reinforcement learning can help our agent reach the goal by providing the maximum reward for finding the shortest (optimal) path to the goal. When an agent is learning, it can test the possibilities and find the shortest path to the goal, and will know it is the optimal path when it has maximized its rewards.

Deep Q-learning uses the idea of Q-learning to find a value of an action given a particular state, but takes it a little further and implements neural networks to estimate the value of an action based on the state. (Loeber, 2022) For this project, deep Q-learning was used to provide the value approximations based on past actions to the agent to be able to choose the highest value action and find the most optimal path to the goal.

References:

Lindwurm, E. (2021, December 12). Intuition: Exploration vs Exploitation - Towards Data Science. *Medium*. https://towardsdatascience.com/intuition-exploration-vs-exploitation-c645a1d37c7a

Loeber, P. (2022). Reinforcement learning with (Deep) Q-Learning explained. *News, Tutorials, AI Research*. https://www.assemblyai.com/blog/reinforcement-learning-with-deep-q-learning-explained/#:~:text=Deep%20Q%20Learning%20uses%20the,action%20based%20on%20that%20state.

Mondal, P. (2013, December 11). *Trial And Error System of Learning | Principles of Teaching*. Your Article Library. https://www.yourarticlelibrary.com/learning/trial-and-error-system-of-learning-principles-of-teaching/6052