**Roderick Fisher | CS-370 | Project Two | Design Defense**

This project utilizes a cutting-edge Deep Q-Learning framework to traverse a labyrinth and achieve a highly sought-after "treasure" objective. The problem-solving approach used by an intelligent agent diverges from the methods used by human agents, yet intriguing parallels exist.

When confronted with a maze, human agents tend to adopt a comprehensive outlook that encompasses the entire labyrinth, including the initial point, the endpoint, and all the intermediary passages. This broad perspective empowers them to make well-informed decisions regarding the optimal direction to take. They possess the ability to gauge their advancement toward the goal and instinctively select the most favorable route. Furthermore, humans possess an inherent inclination to evade obstacles without viewing them as detrimental moves, given that the maze's rules explicitly forbid passage through walls. Consequently, they do not expend unnecessary effort contemplating such hindrances.

In contrast, when it comes to solving the pirate treasure maze, the intelligent agent lacks the innate comprehension of concepts such as "progressing towards the finish" or "moving away from the finish" that a human agent possesses. Despite the consistency of maze rules, the intelligent agent must undergo a simplified instructional process, focusing on avoiding boundaries and recognizing specific spatial patterns that facilitate reaching the destination more efficiently. To achieve this, the intelligent agent relies on a specialized table known as a "Q-Table," which contains a compilation of values associated with each feasible move. Moves that contribute to the agent's successful arrival at the goal are assigned positive values, while moves that lead the agent into prohibited areas receive negative values. Consequently, the agent is equipped to make well-reasoned decisions regarding its course of action.

In contrast, when it comes to tackling a maze, the intelligent agent's approach diverges from that of a human agent. While the human agent possesses a comprehensive view of the maze, including the starting point, destination, and all paths in between, the intelligent agent lacks this inherent understanding. Instead, the intelligent agent must be taught in simpler terms, emphasizing the avoidance of boundaries and recognizing specific spatial patterns that expedite reaching the goal. To accomplish this, the intelligent agent relies on a table known as a "Q-Table," which lists the values assigned to each available move. Positive values are associated with moves that bring the agent closer to the goal, while negative values indicate moves that lead the agent astray. By considering these values, the agent can make informed decisions regarding its approach.

Both the human agent and the intelligent agent employ a similar strategy of "looking ahead" to identify the fastest and most reliable path to the finish, avoiding unnecessary backtracking. Although the human agent cannot systematically evaluate all possible moves and their values, they intuitively assess the perceived value of a move based on other factors. They naturally gravitate towards directions that bring them closer to the end goal. Similarly, the intelligent agent selects moves that indicate progress in the desired direction. Unlike the human agent, the intelligent agent does not require negative values for moving into barriers in the maze since it learns the rules regarding barriers and adeptly avoids them.

In decision-making, there is a balance between exploiting existing knowledge and taking calculated risks. This balance, known as exploitation and exploration in artificial intelligence learning, plays a vital role. During exploitation, an agent chooses moves that promise immediate rewards based on available information. These moves, often referred to as "greedy actions," prioritize short-term gains without considering long-term benefits. On the other hand, exploration involves actions that may not yield immediate substantial benefits but offer the potential for greater rewards in the long run. Agents gather additional information about the environment and outcomes of specific actions to make optimal decisions.

Finding the right balance between exploitation and exploration is crucial for effective learning, with the optimal ratio depending on the learning objective. Setting the exploration rate too high can impede learning progress as the agent exhaustively explores all possible moves and scenarios. Conversely, relying solely on exploitation can lead to suboptimal solutions, even if the agent reaches the goal. Therefore, it is essential to allow the agent to employ both exploitation and exploration methods based on the specific requirements of the problem.

Reinforcement learning algorithms aim to maximize an agent's cumulative reward in an unknown environment. In the case of maze-solving, such as the pirate treasure hunt, these algorithms teach the agent which sequences of squares to follow to accumulate the maximum number of points by the end. As the agent learns which moves or squares result in point gains or losses, it gradually determines the most efficient path toward the end goal, marking the completion of the game.

While the standard Q-Learning algorithm relies on a Q-Table encompassing all possible states and actions, it becomes impractical for larger environments with millions of Q-values. In such cases, a Deep Q-Network (DQN) is preferable. The DQN comprises two neural networks that collaborate to approximate the values stored in the Q-Table. As one network adjusts its weights during the learning process, the updated weights are periodically transferred to the second network, improving the value approximations computed by both networks.

The "Pirate Treasure Maze" project implements a Deep Q-Network to solve the maze. The designed model utilizes the previous environmental state to predict future actions' q-values. Subsequently, the agent selects an action, updates the environment, receives rewards for the action, and evaluates the resulting state. As mentioned earlier, the model incorporates a certain level of exploitation, allowing the agent to pursue immediate high rewards rather than uncertain long-term gains. However, the model also embraces exploration, resulting in variations in chosen paths to the same treasure square across different runs. This variability stems from the exploration component, as pure exploitation would theoretically produce consistent results.

While the intelligent agent may differ from a human agent in solving a maze, such as the one in this project, its learning capabilities can be preserved and applied to future projects. Additionally, the methods employed can be adapted and modified to tackle different problems with distinct structures. The Deep Q-Learning network has a broad range of applications in problem-solving within the field of artificial intelligence, and its implementation can be relatively straightforward, as demonstrated by this project.