John Lopes

June 23, 2024

### Design Defense for Project Two

**The Differences between Human and Machine Approaches to Solving Problems**

In developing a solution for the pirate maze problem, humans use intuition, experience, and foresight to navigate challenges. They plan multiple steps ahead, visualize potential outcomes, and adapt to new situations creatively. In contrast, machines follow predefined algorithms strictly, relying heavily on trial and error, large-scale data processing, and precise memory recall. Machines maintain consistency in their actions, free from fatigue or emotional interference, which is a significant advantage in computational tasks (Vartak, 2023). To solve a maze, a human would typically start with initial exploration to understand the layout, noting paths, obstacles, and dead ends. They would then plan a route from the start to the goal, considering potential shortcuts and obstacles. During execution, a human would follow the planned path but adjust as needed based on new information or unforeseen obstacles. If the initial plan fails, reflection and adjustment come into play, allowing the human to devise a new strategy and try again (Wentworth, 2019).

The intelligent agent I developed takes a systematic approach to solving the pathfinding problem. Initially, the agent starts with a random policy and initializes Q-values arbitrarily. It explores the environment by taking random actions to gather information about rewards and penalties. As the agent learns from these experiences, it updates its Q-values using the Bellman equation. Over time, it transitions from exploration to exploitation, choosing actions that maximize expected rewards based on its learned values. This learning process eventually leads to the agent consistently finding the optimal path to the treasure.

Both approaches, human and machine, start with exploration to gather information about the environment and adjust strategies based on feedback. However, humans use foresight and planning, considering multiple future steps simultaneously, while the agent makes decisions based on current state and learned values. Humans adapt creatively to new challenges, whereas the agent relies on a strictly defined policy. Moreover, machines can process and remember vast amounts of data accurately, unlike humans who might forget or overlook details.

### The Purpose of the Intelligent Agent in Pathfinding

The intelligent agent in pathfinding serves an important role in autonomously navigating complex environments to achieve a specific goal, such as finding a treasure in a maze. Its primary purpose spans various applications from gaming to real-world scenarios like robotics and logistics. In gaming contexts, the agent represents a non-player character (NPC) tasked with reaching the treasure before a human player. Beyond entertainment, in fields like robotics, these agents facilitate efficient route planning and navigation through unknown or dynamic environments. The agent's objectives include optimizing pathfinding efficiency, adapting to changes in the environment, and continuously improving its navigation strategy through iterative learning. The agent employs Deep Q-learning reinforcement learning technique to learn from its interactions with the environment, adjusts its decision-making based on rewards and penalties, and ultimately develops an effective strategy to solve the pathfinding problem.

Exploitation involves choosing actions based on the agent's current knowledge to maximize immediate rewards, leveraging learned Q-values to make optimal decisions. On the other hand, exploration entails selecting actions that the agent has not yet extensively tried, aiming to gather more information about the environment and potentially discover better strategies in the long term. In the context of the pathfinding problem, an initial emphasis on exploration is important. This allows the agent to thoroughly explore the maze, gather comprehensive data on various paths and their associated rewards, and build an accurate model of the environment. As the agent accumulates experience and refines its Q-values through learning, the balance shifts towards exploitation. The ideal proportion between exploitation and exploration is typically managed using an epsilon-greedy strategy, where epsilon (ε) starts high to encourage exploration and gradually decays over time. This approach ensures that the agent explores sufficiently to gather knowledge early on while exploiting its learned knowledge effectively to optimize pathfinding efficiency as it gains experience.

In the treasure hunt game scenario, reinforcement learning allows the agent to interact with the maze environment, take actions such as moving in different directions, and receive feedback in the form of rewards or penalties. Rewards are positive for actions leading towards the treasure and negative for hitting obstacles or making inefficient moves. By updating its action-value function (Q-values) using Deep Q-learning algorithms, the agent learns the expected cumulative reward for each action in every state of the maze. Over successive learning episodes, the agent chooses the optimal action for each state to maximize cumulative rewards thus guiding it efficiently towards the treasure. Reinforcement learning also provides adaptability, enabling the agent to adjust its pathfinding strategy in response to changes in the maze layout or new learning experiences, ensuring robust performance in complex and dynamic environments (Sutton & Barto, 2015).

**The use of Algorithms to Solve Complex Problems**

I implemented Deep Q-learning using neural networks to facilitate the intelligent agent's pathfinding in this game. Initially, I defined the game environment, incorporating the maze layout and determining feasible actions such as movement in multiple directions. The core of the implementation involved utilizing a neural network architecture to manage the action-state pairs. This neural network processed the current state of the game environment and predicted Q-values for each potential action. During the training phase, the agent interacted with the environment by taking actions based on an epsilon-greedy policy. This policy balanced between exploration, where the agent tried new actions to gather information about the maze, and exploitation, where it utilized its learned knowledge to make decisions aimed at maximizing rewards, particularly reaching the treasure efficiently.

### References

Sutton, R. S. & Barto, A. G. (2015). Reinforcement Learning: An Introduction. <https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf>

Vartak, V. U. (2023). Maze Solving Robot using A\* Algorithm. <https://scholarworks.calstate.edu/downloads/n296x582q>

Wentworth, J. (2019). Problem Solving with Mazes and Crayon. https://medium.com/@johnwentworth/problem-solving-with-mazes-and-crayon-f262f957e3a0