**Module 7-3 Project**

Mitchel Harmon

Online Campus, Southern New Hampshire University

CS-370-15800-M01 Current/Emerging Trends

Instructor Roland Morales

December 11, 2024

**Module 7-3 Project**

Humans and machines approach problem-solving in fundamentally different ways due to their unique capacities for reasoning, learning, and decision-making. Humans typically solve mazes by visually analyzing the layout, identifying starting and ending points, and noting obstacles. They plan a potential route based on spatial reasoning and intuition, adjusting their path when encountering dead ends or obstacles. In contrast, machines follow an algorithmic process, iteratively exploring paths using predefined rules or learned behaviors. Machines excel at systematically navigating large solution spaces, which would be tedious and error-prone for humans (Sutton & Barto, 2018).

To solve a maze, a human would start by visualizing the layout and planning a path toward the goal. They would navigate step by step, making adjustments based on immediate feedback from the environment. When encountering a dead end, they would backtrack and choose an alternative route, repeating this process until reaching the goal. The intelligent agent, however, approaches the problem differently. It begins by observing the maze environment, represented as a state space where each cell is a state. Based on this observation, the agent uses an epsilon-greedy strategy to decide its next action. When epsilon is high, it selects random actions to explore new possibilities. When epsilon is low, it exploits learned knowledge by selecting the action with the highest predicted reward. After taking an action, the agent receives feedback in the form of a reward or penalty and updates its knowledge by training its neural network with this experience (Mnih et al., 2015). This cycle of observation, action, and learning continues until the maze is solved or the game ends.

Although humans and machines share some similarities in their approaches, such as relying on feedback to adjust their strategies, significant differences exist. Humans depend on intuition and heuristics, while machines rely on systematic exploration and reinforcement learning, which enables them to optimize paths without preprogrammed rules (Sutton & Barto, 2018).

The purpose of the intelligent agent in this pathfinding problem is to automate the process of finding the optimal route to the goal. Reinforcement learning enables the agent to learn this optimal path by associating actions with outcomes. Actions that move the agent closer to the treasure are reinforced with positive rewards, while those that lead to undesirable outcomes are penalized. Over time, the agent learns to maximize cumulative rewards, ensuring it consistently selects the most efficient path to the goal (Mnih et al., 2015).

Exploitation and exploration are key aspects of this learning process. Exploration involves testing random paths to gather diverse experiences, while exploitation uses accumulated knowledge to select the most promising action. A balance between the two is necessary, with an ideal strategy starting with high exploration to discover new paths and gradually shifting to exploitation as the agent becomes more knowledgeable. In this maze problem, beginning with an epsilon value of 1.0 (pure exploration) and decaying it over time ensures that the agent gathers sufficient data early on and leverages its learning effectively in later stages (Sutton & Barto, 2018).

Reinforcement learning plays a crucial role in guiding the agent to the treasure. By assigning rewards to successful actions and penalties to failures, the agent learns to predict which actions will yield the highest rewards. Using deep Q-learning, this process is further enhanced by a neural network that approximates the Q-value function, mapping states to predicted rewards for each action. The agent stores its experiences in memory and uses them to train the network by minimizing the error between predicted Q-values and target Q-values, calculated using the Bellman equation (Mnih et al., 2015).

In implementing deep Q-learning, the neural network learns to associate input states, represented as the maze cells, with Q-values for all possible actions. Experience replay ensures that the training data is diverse and unbiased by randomly sampling stored experiences. During training, the model minimizes the mean squared error between predicted and target Q-values. This iterative process enables the agent to continually improve its predictions and adapt to the environment.

By combining systematic exploration, exploitation of learned knowledge, and neural network-based learning, the intelligent agent effectively solves the maze and identifies the optimal path to the treasure.

**Citations**

Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. Nature.  
Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction.