**Project Two**

Michael Kinful

Southern New Hampshire University

CS-370-15800

Roland Morales

12/13/2024

**Project Two**

As humans, we solve mazes using intuition, trial-and-error, and memory. We observe, plan, explore, and adjust based on obstacles. Machines, however, use systematic approaches like algorithms to evaluate paths, relying on structured data rather than intuition. Machines are efficient but lack flexibility outside programmed behaviors.

Humans observe the maze, plan a route, and adjust based on trial and error or heuristics. Intelligent agents represent the maze as a structured grid or graph, set rewards for goals, explore paths using algorithms, and optimize decisions by learning from feedback.

Both methods evaluate paths, learn from failures, and try to be efficient. Humans use intuition and memory to adapt flexibly, while machines rely on algorithms and systematic exploration to optimize.

The agent automates the navigation and effectively finds the best path by exploring, learning, and refining his strategies. Its purpose is to minimize human intervention and handle dynamic or unknown environments effectively.

Exploitation leverages existing knowledge to make decisions, while exploration seeks to discover new information. Early in the task, more exploration ensures a comprehensive understanding of the maze. Later, exploitation uses this knowledge to optimize pathfinding, balancing exploration and exploitation to achieve efficiency.

Reinforcement learning teaches the agent to maximize rewards by navigating toward the goal and avoiding penalties for undesirable actions. By updating Q-values based on feedback, the agent iteratively refines its strategy, discovering the shortest path to the goal.

Deep Q-learning approximates the Q-values of state-action pairs using neural networks. The maze is encoded in a grid, and the agent chooses its actions by following an epsilon-greedy strategy. Experience replay makes sure that effective learning is achieved by training on previous experiences, and Q-values are updated using the Bellman equation. After numerous iterations, the agent finds the optimal policy to reach the path.

Intelligent agents solve pathfinding problems by systematically exploring and optimizing decisions. While humans rely on intuition, agents use algorithms and reinforcement learning, with deep Q-learning providing an efficient, scalable solution for complex mazes.

**Reference:**

Cook, A. (2018, July 9). *Deep Reinforcement Learning Nanodegree Program: What You’ll Learn*. Medium.com. <https://medium.com/udacity/deep-reinforcement-learning-nanodegree-program-what-youll-learn-e424d0d66c1e>

Leon. (2024, April 29). *Python maze solver*. sqlpad.io.

<https://sqlpad.io/tutorial/python-maze-solver/>

Pykes, K (2024, November 6). *Understanding the Bellman Equation in Reinforcement Learning.* Datacamp.com

https://www.datacamp.com/tutorial/bellman-equation-reinforcement-learning