**CS-370-R3356 Current/Emerging Trends in CS**

**24EW3 | 7-3 Project Two Design Defense | DEONNE LUDWIG 02/25/2024**

In this assignment the player needs to find the treasure before the pirate does. The intelligent agent (the pirate) needs to find the optimal path to the treasure using deep Q learning.

We can look at the different approaches of human vs. machine (our intelligent agent). Humans use cognitive understanding and reasoning to solve problems. They assess the problem, break it down into smaller parts, and use past experiences or knowledge to find the solution. They can be creative in their approaches and adapt to changes quickly but also can be influenced by motivation, frustration, or satisfaction. They would analyze the maze, plan a route, navigate the maze avoiding obstacles and iterate and adjust as needed to achieve success.

A machine would follow predefined algorithms or instructions to solve problems. They can use the data to improve performance and accuracy without fatigue or bias. The intelligent agent takes the following steps to solve the problem: Initialization, Training, Exploration and Exploitation, Action Selection, Feedback and Learning and Convergence.

Some similarities are: both aim to solve the maze problem by navigating from a starting point to a goal (treasure) while avoiding obstacles and both involve decision-making based on current perceptions (maze layout for humans, environment state for agents) and past experiences.

Some differences are humans use cognitive understanding and creativity, can adapt, but might include bias, while agents use predefined algorithms, learn from data, and should exclude biases.

The purpose of the intelligent agent in pathfinding is to autonomously navigate a maze or environment to find an optimal path from a starting point to a goal. The agent uses algorithms, such as deep Q-learning, to learn from experience, explore possible paths, and make decisions that lead to the highest expected rewards. The agent's goal is to solve the pathfinding problem efficiently and effectively without human intervention.

Exploitation involves leveraging known information or strategies to exploit the current best option. In the context of pathfinding, exploitation would be following paths that the agent believes are most likely to lead to the goal based on its learned experiences.

Exploration involves trying out new options or paths to gather more information and potentially discover better strategies. In pathfinding, exploration would be taking actions that the agent hasn't tried extensively before, even if they don't seem optimal based on current knowledge.

The ideal proportion of exploitation and exploration in pathfinding depends on the specific characteristics of the maze and the learning process. Early in the learning process, more exploration is favored to discover a wide range of paths and avoid premature convergence to less optimal solutions. As the agent learns and accumulates knowledge, the balance shifts towards exploitation to exploit the learned information and focus on the most promising paths. A common approach is to start with a high exploration rate and gradually decay it over time as the agent gains experience. Reinforcement learning is used in determining the path to the goal (treasure) by the agent (pirate) in the maze environment. It allows the agent to learn optimal decision-making policies through trial and error in the environment. Deep Q-learning enables the agent to learn the value of taking different actions in different states by estimating the expected rewards (Q-values). By iteratively updating its Q-values based on observed rewards, the agent gradually learns to navigate the maze more effectively, ultimately finding the optimal path to the goal.

Using algorithms to solve complex problems offers several advantages like efficiency, scalability, consistency, and optimization. In the Treasure Hunt Game, Deep Q learning (using neural networks) was implemented by initialization of the function to take a neural network model, a maze environment, and optional parameters for training and construct the environment from the provided maze array. It initializes an experience replay object to store and sample episodes for training. The training loop runs until the win rate surpasses a threshold and the agent consistently wins and completes the maze.

Das, A. (2017, March 17). *Introduction to Q-learning*. Medium. https://towardsdatascience.com/introduction-to-q-learning-88d1c4f2b49c

Singhal, R. (2018, December 16). *Reinforcement learning: A beginner’s guide*. Medium. https://medium.com/@rsinghal757/reinforcement-learning-a-beginners-guide-56de9c7fe1d9

Choudhary, A. (2023, August 21). *A hands-on introduction to deep Q-learning using openai gym in python*. Analytics Vidhya. https://www.analyticsvidhya.com/blog/2019/04/introduction-deep-q-learning-python/

Yang, A. (2022, July 25). *What is exploration vs. exploitation in reinforcement learning?*. Medium. https://angelina-yang.medium.com/what-is-exploration-vs-exploitation-in-reinforcement-learning-a3b96dcc9503

First Run at original value

Exploration factor: epsilon = 0.1

A screenshot of a computer

Description automatically generated

A screenshot of a computer code

Description automatically generatedA computer code with many words

Description automatically generated with medium confidence

A screenshot of a computer code

Description automatically generated

A screenshot of a computer program

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

A screenshot of a computer

Description automatically generated

Second Run value

Exploration factor: epsilon = 0.5

A screenshot of a computer program

Description automatically generated