**Project 2**

Samuel A. Meade

Southern New Hampshire University

[CS-370-T5546 Current/Emerging Trends in CS 23EW5](https://learn.snhu.edu/d2l/home/1313000)

Timothy (Tim) Alexander, M.S., CIS

6/17/2023

Design Defense

Human vs. Machine Problem Solving

The Treasure Hunt video game instantiates a 2D environment, containing paths, obstacles and treasure, simulating Markov Decision Process. As a player, the goal is to traverse the maze and find the treasure. Moves are limited from one to four directions/actions, one being in the face of an obstacle, four being a crossroads.

As a human player, recollection plays a large part in the decision-making process when navigating the maze. A human may be presented with a crossroads, take a path, find an obstacle, and know enough to not repeat the path that ended abruptly. The fault in human play comes in that same strength of recollection. It is easy for a human to forget the paths traversed as the dead-end paths become longer and options increase, while an AI, once trained, will be able to reference all attempts in a given environment.

Training an intelligent agent to solve the maze is a bit more involved. Utilizing Q-learning, a reinforcement learning policy that will find the next best action, given a current state, rewards the agent based on its predicted action (Banoula, 2023). When the action chosen is favorable, the agent is given a reward, when the prediction results in a terminal state, the reward is negative. Ultimately, this is intended to mimic a person’s ability to make informed decision.

Intelligent Agent in Pathfinding

Training an intelligent agent in pathfinding requires the consideration of exploitation and exploration principles. Exploitation refers to the approach of incentivizing the acquisition of rewards through projected value, while exploration encourages the improvement of action comprehension (Yang, 2022). While exploitation may result in rapid growth, exploration will more accurately train the agent in the long run. Depending on the application, the ratio of exploitation to exploration will shift. As the environment grows and more complex decision making is needed by the agent, exploration will take priority over exploitation. If the risk to reward warrants a quick solution within a smaller environment, exploitation principles will prove the be the better perspective. Analyzing and knowing which tactic to implement given the need of the project is imperative.

In the case of the Treasure Hunt game, the agent (pirate), is able to learn through reinforcement training by avoiding terminal states. For the pirate, a terminal state would be considered a move resulting in a dead end. Utilizing exploitation, a larger positive reward in an unexplored path may encourage the model to keep moving forward. Coupling exploitation with a basic reward of successful moves tracking and avoiding known dead ends, the pirate will likely learn at an ideal rate.

Algorithmic Problem Solving

A screenshot of a computer program

Description automatically generated with medium confidence For the Project 2 Treasure Hunt, the Pirate utilizes Q-learning to reach the treasure. In the example to the right, the pirate references past actions through the use of episodes, taking them into consideration for future action. Leveraging episodes, the model can tune and grow similar to how a human would through recollection of previous attempts. That makes this neural network rather advanced and allows it the ability to traverse more complex problems.

Citations

Banoula, M. (2023, February 22). *What is Q-learning: Everything you need to know: Simplilearn*. Simplilearn.com. <https://www.simplilearn.com/tutorials/machine-learning-tutorial/what-is-q-learning#:~:text=Q%2Dlearning%20is%20a%20model,next%20action%20to%20be%20taken>.

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>