Michael Garramone

CS 370 Current/Emerging Trends in CS

Professor Alexander

2/21/2024

**Design Defense:**

Analyze the differences between human and machine approaches to solving problems.

I have been asked to create an intelligent agent that would act as an NPC (non-playable character) for the gaming company that I work for. The intelligent agent will be known as the pirate for the treasure hunt maze game. The pirate is rewarded for all the correct moves that they will make to end up at the spot the treasure is at and will be penalized for making any moves that are not allowed (making an invalid move and moving backwards to a space they already had visited) as well as blocking themselves from moving forward to a new block with a wall in the maze.

For a human walking through a maze compared to a machine learning how to navigate a maze to get to the treasure, there are both similarities and differences between the two. As humans we can see what is in front of us and come up with an idea of how far we can go and if the maze ends up being a dead end or not. For the machine, it has no visual conception of what is forward, backwards, to the left or the right of itself, and must decide on what direction it must go before choosing the next block it can move to. Where both human and machine are similar, is that humans have a curiosity to explore and with machine learning, exploration helps with reinforced learning. Humans at the start of a maze have no idea of what is going to be the path to the end like the machine. Both use exploration to find paths that are best suited to get to the goal. However, humans might be able to discover the path and remember the path more quickly than a machine would. It may take a few tries at the maze, but humans would not have to run the maze as much as a machine would to get the best path for the treasure. Also, to learn how the to get the best path both machine and human would have to take steps back to old positions and choose a new direction to go.

Access the purpose of the intelligent agent in pathfinding.

In this model where the agent or pirate for this matter is learning to move around the maze, I used the concepts of exploration and exploitation for reinforcement learning. “In exploration techniques, agents primarily focus on improving their knowledge about each action instead of getting more rewards so that they can get long-term benefits” (Yang, 2022). The agent is paying more attention to learning the maze rather than focusing on the reward and penalization mechanic that is being instituted. After the agent has reached a good level of reinforced learning from exploration it is transitioned over to the exploitation technique. “Exploitation is defined as a greedy approach in which agents try to get more rewards by using estimated value but not the actual value” (Yang, 2022). Based on the knowledge that the agent received during its time of using the exploration technique, it makes decisions on what direction is best suited for the optimal reward and path that needs to be taken to get to the treasure.

Using Learning rate decay, we can find the needed amount of time the agent uses exploration techniques and when it should switch over to exploitation techniques to improve the agents path taken through rewards and penalizations to the treasure. “Learning rate decay is a technique for training modern neural networks. It starts training the network with a large learning rate and then slowly reducing/decaying it until local minima is obtained” (Haswani, 2021).

Reinforcement learning helps the agent to determine the best path to the treasure through the learned actions that were taken by the agent with the rewards system. While the agent is in the beginning stages of learning each state gives a score of how much the agent is rewarded and penalized for the moves that were taken. During this time the agent is learning as it explores the maze. While the agent explores the maze, it updates its action-value based on the feedback it receives, and over time that action value is optimized by the rewards and penalties it receives. Creating more accurate data needed to optimize the best path needed to reach the treasure.

Evaluate the use of algorithms to solve complex problems.

Deep Q-Learning using neural networks were implemented for this game to function properly with the agent finding the treasure. The agent first started exploring the maze using exploration techniques. This allowed the agent to pay more attention to learning the maze rather than focusing on the reward and penalization mechanic that is being instituted. With the use of decay learning, the agent would move from exploration after a set period of time and use exploitation techniques that focused purely on what the best action for the agent to take to optimize the best path to the treasure based on the reward system. If the agent was to move in a wrong direction such moving backwards or choose to move forward and there was a wall it would be penalized. By epoch 5, the agent had made its first correct run through the maze finding the treasure, and at epoch 446 the agent started winning the treasure hunt maze 100% of the time.

References:

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

Haswani, V. (2021, May 30). *Learning rate decay and methods in deep learning*. Medium. https://medium.com/analytics-vidhya/learning-rate-decay-and-methods-in-deep-learning-2cee564f910b