CS370 Project Two

Text

Description automatically generatedI am a developer that has been tasked with creating an intelligent agent for a pirate NPC character in a treasure hunt game. This game consists of a maze that must be navigated to find hidden treasure. The player’s goal would be to find the treasure before the pirate that was created. Here is the maze that must be navigated designed as an 8x8 matrix:

Text, letter

Description automatically generatedA human would typically start from the start point and look at the available path options. The person would look at each option an see which directions had routes they would go and try one. Each decision they make at each point would likely be remembered and factored into future decisions to allow for the person to learn as they go. Eventually, they would find their way through by repeating these steps. The intelligent agent takes a similar sort of approach that is more methodical. The agent starts at a random cell and places themselves in that cell.

Text

Description automatically generatedThe agent then maps out the possible movements, then tests each one and calculates an overall score for how well the maze was traversed. The path with the highest score in the end is the optimal path.

The human and the computer approaches both start from a random point in the maze. Each looks in all directions and tries new directions to go based on previous choices and some random choice. The difference between the human and the computer approach is in some of the intuitions that humans have that a computer taking a methodical approach to solving a maze. Their selections and decisions will be different which may have an effect on how efficiently the maze is solved.

In machine learning, exploration is the concept where an agent is allowed to improve its knowledge about each action. Exploitation is the concept where an agent exploits the current value and chooses the highest value. Exploration as a concept is much more forward-looking and offers benefits over a long period. Exploitation is looks at more of the short-term benefits. Finding the right balance between making an exploration decisions and exploitation decisions will allow for optimal maze solving. (Lamba 2018) For the agent that was created, epsilon, the value that controls the balance between exploration and exploitation, was set to 0.1. It can be set anywhere between 0 and 1. This means that the probability that an explorative action would be taken next was 0.1. There was not much experimentation with other values, but this epsilon value provided an excellent balance between explorative and exploitative action.

Reinforcement learning is the process be which reward is given for correct or positive action and punishment is given for negative or incorrect action. Through this process, we can create a system that can learn to do certain things. (Yoon 2019) In this case, we want a pirate NPC agent to traverse the maze to get to the treasure. We can help determine the path to the goal with reinforcement learning by implementing a point system that rewards the agent with points as they take positive steps through the maze while punishing them by removing points when they take negative steps through the maze.

For the pirate NPC agent, a deep Q-learning neural network was created. This was done by first defining the maze and the actions that can be taken. Then a random open cell in the maze is selected and the agent is placed in the cell. A list of discrete actions that can be taken is generated. Each action has a potential future reward score calculated, then the highest scoring action is taken. This process is repeated until the agent reaches the treasure. Through this process, a Q-table map of the actions to take to traverse the maze is created. (Lamba 2018)

References:

Lamba, A. (2018, September 3). *An introduction to Q-learning: Reinforcement learning*. Medium. Retrieved August 21, 2022, from https://medium.com/free-code-camp/an-introduction-to-q-learning-reinforcement-learning-14ac0b4493cc

Yoon, C. (2019, July 17). *Understanding actor critic methods*. Medium. Retrieved August 21, 2022, from https://towardsdatascience.com/understanding-actor-critic-methods-931b97b6df3f