Jason Marcil

CS-370 Project Two Design Defense

22EW6

The ways in which a human and a machine would approach solving this maze are somewhat different. A human would observe the maze, analyze it to determine the best possible path, and execute a path to solve it. They may also start executing a path before fully determining the correct solution and analyze the result of choosing directions as they approach possible crossroads. However, a human decides to approach the maze, they are using a biological form of a neural network to make decisions and learn how to solve the problem.

An intelligent agent on the other hand takes a slightly different approach to solving the maze. The agent will make random decisions that either result in a reward or penalty for making decisions. As these rewards and penalties rack up with every epoch, the agent has a better chance of making correct decisions that lead to a solution. This requires the use of a neural network so that the agent can take in information at the input layer, process it within hidden layers, and output a solution over several cycles through an output layer.

While the two approaches to solving the maze are certainly different, there is a similarity that they share. It’s important to remember that intelligent agents are built by humans, so there is a human bias to how intelligent agents make decisions. Where humans can analyze a path before executing the solution, machines need to make attempts to learn how to do it. The only noticeable difference here is that humans don’t need to make attempts to solve the problem, but they are processing information in a similar manner. Both humans and machines make decisions based on the information they already know. For humans, they can look at the entire maze and make decisions as they work through it based on what they can see in front of them. Intelligent agents make attempts and will fail until they succeed, but they learn the solution based on the results of their previous attempts. Humans have the advantage of being able to look ahead, but other than that, they both are doing some form decision making based on what they know.

Learning algorithms can be abstracted as being a balance of exploration and exploitation. Exploitation is a strategy that improves an existing model that is being learned, while exploration attempts to learn better models. In exploitation, an intelligent agent pursues the path of learning that provides immediate rewards, whereas exploration leads the agent to decide to forgo an immediate reward and take a directionless approach with the intent of discovering greater rewards elsewhere (Perez, 2017). The ideal proportion for this problem would be that the agent begins learning by using exploration while progressively utilizing a higher degree of exploitation to progress through epochs. Exploitation will ultimately be favored since it is more likely to solve the maze in a shorter period.

Reinforcement learning can help to determine the path to the goal because this type of learning teaches the agent through a rewards system. This is an example of how a machine will use exploitation to achieve the goal of solving the maze. When the agent makes a correct decision, the system rewards the agent with a point, but if it makes the wrong decision, a point is deducted. The agent wants more points, so it will make decisions it thinks will lead to that outcome. Ultimately, this will lead the agent to the treasure if this goal is continually achieved.

To implement deep Q-learning using neural networks for this game, we first need to implement a Q-table, which is a lookup table where we calculate the maximum expected future rewards for action at each state (Cheung, 2018). We then use the Q-Learning algorithm to learn the value of the Q-table. Once the table is initialized, the agent, utilizing its neural network, will choose and perform an action by means of exploration and progressively use exploitation as it learns through reinforcement. After each action, the state changes for the agent, so it evaluates what actions led to the current state before going through the next epoch. This process will be repeated until the agent has made it through the entire maze. The speed at which the machine will be able to finish the maze will depend on how the algorithm has been configured and fine-tuned.

References:

Cheung, D. (2018, September 3). An introduction to Q-Learning: reinforcement learning. Medium. https://medium.com/free-code-camp/an-introduction-to-q-learning-reinforcement-learning-14ac0b4493cc

Perez, C. (2017, October 7). Exploration, Exploitation and Imperfect Representation in Deep Learning. Medium. https://medium.com/intuitionmachine/exploration-exploitation-and-imperfect-representation-in-deep-learning-9472b67fdecd#:~:text=In%20exploitation%2C%20an%20automation%20greedily%20pursues%20a%20path,the%20intent%20of%20discovering%20a%20greater%20reward%20elsewhere.