Stephen Cardone

Professor Balogun

CS-370

February 21, 2021

Pathfinding Algorithm: Design Defense

In this project I have implemented a Q-learning neural network to create an AI actor that solves a maze. The purpose of this paper is to defend the design by explaining and evaluating the design and the process used to come up with the design. I will do this by analyzing differences between human and machine approaches to solving problems, explaining the fundamentals of an intelligent agent in pathfinding, and finally, I will evaluate how I implemented the deep Q-learning algorithm for pathfinding in this game.

Humans and machines solve problems very differently. Let us look at the maze solving process from the perspective of a human versus that of my intelligent agent. Firstly, lets assess how a human would solve a maze. Think of if you where put into a corn maze and asked to find the exit. The first thing you will do is look around, and you would sub-consciously form a memory for what the maze looks like in your current perspective. Then, as you walk around, you will expand your memories, and make some connections about what each perspective looks like, and you would begin to form connections between each perspective. For example, you may run into a wall (lets call it wall “A”). You then realize and memorize that, if you follow wall “A” to the left, you will run into a T in the patch where you can choose left or right. You may choose left, which turns out to be a dead end which you memorize. Then you go right, and the maze continues. As you learn more about the maze, eventually you will solve the maze. At this point, if you are put in a random location in this same maze, you will recognize what you see and be able to find a familiar landmark that allows you to get on a familiar path to solve the maze.

A computer agent on the other hand will function slightly differently. The computer does not have the ability to form the arbitrary visual memories like a human does. Instead, the computer utilizes known information and a system of rules to identify a solution to the maze. In my maze solving algorithm, we supply rules for rewards, and punishments. For example, if the computer attempts to move a direction, and ends up hitting a wall, the reward is -1. The computer doesn’t want a negative reward, so it knows, that the next time it is in this same location, it does not want to try to move that direction, because the reward will be -1. Therefore, it moves a different direction. Eventually, it will find the exit which has the highest reward of +1. We let the agent continue to explore the maze many times, and it continues to store the reward/punishment values in as many situations as possible. Then, when we want to solve the maze, we just tell the agent to “go down the path that leads to the highest possible reward”. The computer agent then identifies a path that never runs into any walls, and ends up at the maze solution. Then the agent just follows those directions and solves the maze.

Given the two explanations above, there are similarities and differences to how a human and how my computer agent are solving the maze. The similarities are that, in both cases, the actor has some way of recognizing where they are, looks back on some form of memory and identifies that they know how to go from where they are, to the solution of the maze. The differences are on how they gather and store information, how they determine good from bad, and how they picture the solution from their current location. The human relies on visual memories and keeps track of what visual memory is near another visual memory eventually leading to the end, while the computer relies on a system of rewards and punishments and it tries to maximize the rewards and minimize the punishments which leads to the solution.

Exploration and Exploitation are a fundamental concept in reinforcement learning. At a fundamental level, “Exploration is all about finding more information about an environment, whereas exploitation is exploiting already known information to maximize the rewards.” (Fernandez) Exploitation is the mechanism that allows the agent to say “I know if I go up right now, this leads down a path with high rewards” which implies that going up solves the maze. Exploration is the mechanism that says “If I go up, I know I will have a high reward (solving the maze) however what If I go down instead? It is possible that I will find an even higher reward (solving the maze faster).” By adjusting how often the system chooses to exploit vs explore, we can improve the speed at which the neural network reaches 100% success. For my algorithm, I ended up using an exploration factor of .01 which means that the algorithm would exploit an average of 99 times and exploit 1 time. This solution yielded good results for me and allowed for a 100% solve rate after 329 iterations.

Next, I will explain how I implemented deep a Q-learning neural network for this pathfinding game. Firstly, I set a number of episodes (number of times to play the game). Then, for each episode, I do the following process. Set the actor to a random location in the maze. Then, begin the “Action Loop” until either the game is won, or a failure condition is met. The “Action Loop” does the following: Perform an action (either exploration or exploitation). Receive the reward for the action taken. Store the action and result in a memory object. Train the AI model on the data. This process populates a Q-table which stores the potential future rewards from any given location in the maze. Eventually, once the successful training is completed, the Q-table has 100% coverage and we will solve the maze correctly 100% of the time.

In conclusion, we have investigated the similarities and differences between human and machine problem solving, specifically in the case of solving a maze. Also, we investigated the functions of the intelligent agent in pathfinding, including the principals of exploration and exploitation. And Finally, we analyzed the Q-Learning algorithm I used to implement the pathfinding for the treasure hunting game in this project to achieve 100% accuracy in solving the maze.

Works Cited

Fernandez, E. (2019, November 28). Ai is not similar to human intelligence. thinking so could be dangerous. Retrieved February 21, 2021, from https://www.forbes.com/sites/fernandezelizabeth/2019/11/30/ai-is-not-similar-to-human-intelligence-thinking-so-could-be-dangerous/?sh=1e33cf156c22

Lamba, A. (2018, August 27). A brief introduction to reinforcement learning. Retrieved February 21, 2021, from https://medium.com/free-code-camp/a-brief-introduction-to-reinforcement-learning-7799af5840db