Omar Al Yousuf

Southern New Hampshire University

CS 370: Current/Emerging Trends in CS

Prof. Kalysa Wilson

02/15/2022

**Design Defense**

* **Analyze the differences between human and machine approaches to solving problems.**
  + Describe the steps a human being would take to solve this maze.

For a human, this maze problem is considered a very easy task. A human might look at this maze and have it possibly solved in minutes, probably at a young age. A human can look at the maze and determine the shortest and the best way to solve it by processing a few possible outputs without the need to actually try the different paths that might lead or not lead to the treasure.

* + Describe the steps your intelligent agent is taking to solve this pathfinding problem.

For an artificial agent, the maze is not viewed as it was with human eyes. The agent will have to try multiple trials to find the best path to the treasure. This could be hundreds, thousands, and more depending on the size of the maze. The agent starts by getting some information about the game such as where to start, where the treasure is at, and the possible moves it’s able to make. Utilizing algorithms, the agent tries to find ways to get to the treasure by trials and errors. Finally, the agent finds the best way to get to the treasure and produce that output. The agent reached a %100 winning rate with 130 epochs after 14.34 minutes.

* + What are the similarities and differences between these two approaches?  
    Similarities:

One similarity to mention is that both the human and the agent know where the start point, and endpoint (treasure) are at.

Differences:

* A human is able to solve the maze in a matter of a minute or two, while the agent would need more time, possibly 10 minutes or more to solve it.
* Humans get the problem input using their view (eyes), while the agent receives its input as an array of numbers.
* A human doesn’t necessarily require trials of different moves to find its optimal path, as an agent will. Instead, a human can look at the problem and solve it by only looking at it without the need pf trials.
* **Assess the purpose of the intelligent agent in pathfinding.**
  + What is the difference between exploitation and exploration? What is the ideal proportion of exploitation and exploration for this pathfinding problem? Explain your reasoning.

Exploitation refers to using existing solutions and refinement to improve them. “Exploitation refers to the refinement of existing products, resources, knowledge, and competencies, and is associated with incremental changes and learning through local search” (Benner & Tushman, 2003).

Exploration, on the other hand, refers to exploring, examining, and looking for new solutions. “Exploration refers to the discovery of new products, resources, knowledge, and opportunities, and it is associated with radical changes and learning through experimentation” (Sinah, 2015).

An ideal proportion of the two types for this pathfinding problem are %90 exploitation, and %10 explorations, which could be represented by an “Exploration Factor” of 0.1. This means the agent will take one random action every 10 moves.

* + How can reinforcement learning help to determine the path to the goal (the treasure) by the agent (the pirate)?

Reinforcement learning could potentially help the agent solve the maze problem and find its way to the treasure by introducing a rewards and penalties system. The pirate will go through a series of experiments and exploitation to find its optimal way to the treasure through trials and errors. The agent will always try to accumulate the maximum number of rewards which eventually results in finding its way to the goal.

* **Evaluate the use of algorithms to solve complex problems.**
  + How did you implement deep Q-learning using neural networks for this game?

To train the game neural networks, we followed the steps below:

1. First, we generate thousands of games by exploiting the derived policy making 90% of our game move, and the rest, %10 would be random moves (exploration) as explained earlier in this paper.
2. Inject the most recent training samples in our neural network and ignore any older ones.
3. Finally, after each game move, we will generate an episode that consists of five elements:
   1. Envstate - environment state: which represents a full picture of the maze cells.
   2. Action: the moves the pirate can takes which is up, down, left, and right.
   3. Reward: is the reward received from the action the agent takes.
   4. Envstate\_next: which is the new maze environment state which resulted from the last action.
   5. Game\_over: which indicates if the game is over or not.

“The weights of network N are initialized with random values, so in the beginning N will produce awful results, but if our model parameters are chosen properly, it should converge to a solution of the Bellman Equation, and therefore later experiments are expected to be more truthful” (Zak).

**Reference**

Benner, M.J., Tushman, M.L. (2003). Exploitation, exploration, and process management: The productivity dilemma revisited. Academy of Management Review, 28(2), 238–256.

Sinha, S. (2015). The Exploration–Exploitation Dilemma: A Review in the Context of Managing Growth of New Ventures. Vikalpa, 40(3), 313323. <https://doi.org/10.1177/0256090915599709>

Zac, S. (n.d.). *Deep reinforcement learning for maze solving¶*. qmaze. Retrieved February 20, 2022, from https://www.samyzaf.com/ML/rl/qmaze.html