**7-3 Project Two Design Defense**

Antonio Sobalvarro

CS 370 Current/Emerging Trends in CS, Southern New Hampshire University

Dr. Obafemi Balogun

October 16, 2022

A human being might try to solve this maze by simply walking through the maze, trying to remember what turns led to dead ends and then tracing their steps over until finally finding the end. They could also try an approach where they make either only right turns or left turns in order to find some sort of pattern. An intelligent agent, by contrast, might solve this solution through a method of assigning values to each step it takes along the maze and then compiling a sort of road map where it then follows the maximum values that lead to the end, thus maximizing the rewards. Both humans and an intelligent agent use a system of what is essentially trial and error while attempting to reach the end goal. A human, however, does not assign a numerical value between 0 and 1 to each step they have taken, something an intelligent agent would do, which means that an intelligent agent is much more strictly methodical and calculating in its approach. Essentially the intelligent agent is “maximi[zing] its performance measure” (Arun, 2021) by collecting data about the maze, then analyzing the paths it can take from the start to the end point, and finally taking the path with the maximum performance measure based on the output of the data it collected when compared to the possible paths.

Exploration and exploitation are very important to the field of artificial intelligence, but they are sometimes confused because they can be similar, and they can intertwine. Exploration, in simplest terms, occurs when the AI is introduced to a new environment and so it needs to explore it and find the best solutions in order to maximize its output. After this exploration has been completed, the AI then uses all the knowledge taken from the exploration phase in order to quickly and directly take the path that aligns with the maximized reward goal. In other words, the AI exploits its accumulated knowledge of the environment to obtain the maximum reward more directly. As with everything, there needs to be a balance of both exploration and exploitation if the AI is to do anything effectively. Having too much exploration can obviously lead to a delayed exploitation of the accumulated resources; it’s even safe to assume that if too much exploration is done, the AI runs the risk of never being able to even being the exploitation phase. However, if the AI overdoes the exploitation, then that leaves no room for exploration, which translates to a lack of properly learning from the environment and a cap on available knowledge is can then further exploit. One interesting thing to note is that “…the more mature a system becomes, the more the role of exploration diminishes” (Bachmann, 2018). I would venture to say that the particular pathfinding problem for this project should explore just enough before exploiting its knowledge. There is no clear-cut way to say exactly what proportion of each it should have, the program will run, and it will determine this proportion based on its findings, which I imagine would be that it does not need to explore much since the size of the map is not large. Reinforcement learning helps the agent determine the path to the goal by considering the assigned values of each move, or change of state, then it will use that information to find the best possible route to the goal, which in terms of AI is deemed maximizing the reward.

Since deep Q-learning formed an integral part of this course and project, the most important thing to do is to import the needed libraries, which allowed for the construction of the environment where the neural network does its learning via exploration. The algorithm trains the neural network by allowing the agent to use a value system to assign each stage a numeric value. Once the exploration had occurred, the agent then was able to design a plan to reach the goal by way of maximizing rewards. The last step, a test, assures proper functioning within the confines of the environment.

References

Arun. (2021, October 31). *Artificial intelligence series: Problem solving agents*. Medium. <https://medium.com/geekculture/artificial-intelligence-series-problem-solving-agents-2ee405ddf4d0>

Bachmann, A. (2018, February 12). *Exploration vs. exploitation — An innovators dilemma*. Medium. <https://alex-bachmann.medium.com/exploration-vs-exploitation-an-innovators-dilemma-4f0014f32b94>

Zhang, J. (2020, July 20). *Algorithm of life: The exploration exploitation tradeoff*. Medium. <https://medium.com/@jerrywrites/algorithm-of-life-the-exploration-exploitation-tradeoff-67943d4668c>