**Design Defense**

**7.3 Assignment**

**Created by Issaih Perez**

**06/19/2025**

As a part of your project, you will also submit a design defense. This design defense will demonstrate the approach you took in solving this problem, explain how the intelligent agent works, and evaluate the algorithm you chose to use. In order to adequately defend your designs, you will need to support your ideas with research from your readings. You must include citations for sources that you used.

1. **Analyze the differences between human and machine approaches to solving problems.**
   1. Describe the steps a human being would take to solve this maze.
      * When looking at the overall design of the 'Maze', there are many ways a human could solve this 'Maze'. Assuming the player character was able to see all the spaces on the grid, the Human User would immediately look for which specific point they start out in and look at the location of the end point and head towards that goal. Once the user has found a path that works, they would then start looking for ways to optimize or reduce the overall steps it would take to get to the end point. If we were to assume that the user initially did not know the contents of the grid but only knew the possible locations they could go towards, then it might take a bit longer, but the user could make moves and chart each movement on a paper or take mental notes on what steps worked until they reach the goal. Once the goal has been reached, the player would then utilize their notes or memory from the previous actions to see how they could refine their approach when starting from the end point, and could either try moving into different squares on specific paths to go into places they haven't moved before to see how that would change the outcome or maybe try going an entirely different route to see if that route yields the same results or not.
   2. Describe the steps your intelligent agent is taking to solve this path-finding problem.
      * According to the documentation and code, the Intelligence Agent makes moves in the four cardinal directions; Up, Down, Left, Right, and primarily learns from experience, meaning the agent would reuse their previous path consistently. One thing to keep in mind is the Epsilon Value, which would translate to the Agents 'learning rate'. For example, if this value is set as 0.1, this would mean that for every ten attempts, nine will be made primarily based on experiences while the one remaining attempt will be an 'exploration' attempt where the agent tries to learn a new path. These attempts are then stored with 'memories' which records the attempt and the results. Once the algorithm has found a 'memory' that produced better results, it would use that path as the primary route until a new, better route has been found. This program would continue until the most optimal route has been found within the given grid, as no 'better' route could be found if all paths are explored and assessed completely, or the maximum number of 'memories' has been reached.
   3. What are the similarities and differences between these two approaches?
      * Some similarities with the two approaches would be that both user(s) are trying to find the most optimal path, so they would benefit a lot from constant testing and exploration to discover the more 'ideal' routes. They both also learn from experience, as both approaches are experience based, which would mean that the user(s) would increase their overall performance the more experience they have with the game. Some differences between the two approaches, aside from the user, would be how they learn or figure out the ideal pathing. Humans, for example, do not need to attempt the game a minimum number of times to gain a learning curve. Specifically, the Intelligence Agent attempts a new path 1/10 times, while humans do not have that restriction. Humans can learn a new path immediately after attempting a path, which can reduce the amount of attempts the user might take when playing the game. Another difference is that humans are limited by their desire to learn or improve, as if someone gets bored or frustrated with the assignment, there generally is not something preventing the user(s) from walking away and not attempting the game anymore. The Agent, on the other hand, is limited by the amount of attempts they can make via memories. Depending on the size of the grid they are 'exploring', humans are only limited by their will to complete the task while the algorithm is limited by their maximum number of attempts.
2. **Assess the purpose of the intelligent agent in pathfinding.**
   1. What is the difference between exploitation and exploration? What is the ideal proportion of exploitation and exploration for this pathfinding problem? Explain your reasoning.
      * Exploration would be the process where we learn what overall options and/or opportunities we have available to us while Exploitation is when we would act upon those discovered options. (Bauman, 2023) When comparing the two, one action would come before the other, as if we do not explore our options, we cannot effectively use all our resources effectively nor efficiently because we are not giving ourselves all the available options for us to potentially exploit. Without the exploration of our options or potential avenues that we can take, we cannot make informed decisions that could potentially improve our overall efficiency or show us how to exploit the current scenario into something that benefits us more than the current set up.
   2. How can reinforcement learning help to determine the path to the goal (the treasure) by the agent (the pirate)?
      * Reinforcement learning involves an agent, environment, and goal, where the agent operates or moves around the environment and figures out a way to accomplish the predetermined goal. (Murel & Kavlakoglu, 2024) There are at least four components that are used within the Reinforcement Learning method: Policy, Reward Signal, Value Function, and Model. **Policy**defines the agent's behavior by mapping the perceived environment and environmental states, which would then lead the agent to take specific actions depending on the state. (Murel & Kavlakoglu, 2024) **Reward Signal** is the reward the agent gets when taking a specific action, with the overall goal being to maximize the number of 'rewards' the agent receives when it takes an action. **Value Function** is different from Reward Signal as this is the overall value the agent gets when looking into the longer term. Lastly a **Model** describes a way for the agent to predict their environment, allowing it to guide the algorithm into assessing different solutions/outcomes based on various scenarios.
3. **Evaluate the use of algorithms to solve complex problems.**
   1. How did you implement deep Q-learning using neural networks for this game?
      * Within the program there is a Reward/Punishment portion that awards/reprimands the Agent depending on the specific action they took. A snippet of the code will be attached but as it will show the number of points/scores the agent gains/loses depending on if the action is valid or invalid. This decision was based off the decaying learning function, which shows the learning curve the agent would take when trying to determine the ‘best’ route to reach its target. This helps incentivize the agent to win quickly and find the most ‘ideal’ path based off the experiences and history the agent had with the game/project.

A screen shot of a computer code

AI-generated content may be incorrect.

**Resources:**

(Indent Here) Bauman, D. (2023, October 21). Exploitation and Exploration. *Standford University*. <https://danielbauman.sites.stanford.edu/exploitation-and-exploration#:~:text=Exploration%20is%20the%20process%20by,I%20bring%20it%20up%20today.>

(Indent Here) Murel, J., & Kavlakoglu, E. (2024, March 25). What is reinforcement learning?. *IBM*. <https://www.ibm.com/think/topics/reinforcement-learning#:~:text=The%20reinforcement%20learning%20agent%20learns,actions(s)%20to%20take.>