**Design Defense**

Human intelligence, a highly debated and researched topic, has no single definition or theory that fully captures its complexity. It extends beyond simple information processing, encompassing emotional intelligence, creativity, and critical thinking. “Human intelligence evolved over millions of years through the process of natural selection” (Khurana, 2023). Intelligence is a multidimensional construct, and cognitive abilities, including reasoning, learning, problem-solving, adaptation, and effective interaction with the environment, only demonstrate a small part of human intelligence’s intricacies. Therefore, writing about this complex topic is far beyond the purpose of this short paper. However, this design defense paper utilizes a few abilities of human intelligence to build a foundation for measuring Artificial Intelligence’s ability to solve a similar problem, such as a maze.

Humans combine observation, critical thinking, problem-solving, and navigation skills to solve a treasure hunt maze. “Human beings’ ability to solve mazes differs because their level of cognitive skill differs from one another” (Fitchett, 2019). First, they study the instructions and rules to find any helpful clue or hint that could lead to capturing the treasure. Observation is the next step since they try to familiarize themselves with the game environment and identify any essential visual information and the route leading to the treasure location. Then, they use their critical thinking skills to follow clues and extract useful information. At this point, they engage their problem-solving skills to make logical connections between gathered information and consider any possibilities to navigate the maze. Finally, according to their observations and gathered clues, they navigate the maze, keeping track of the route to avoid any unnecessary step retracing to solve the maze. However, it is important to remember that humans can make mistakes during any stage of solving the maze, which can lead them to a dead end or an obstacle. In these circumstances, they backtrack and explore other possible solutions to find an alternative route that puts them on the right path for capturing the treasure.

On the other hand, this intelligent agent utilizes the Deep Q-learning algorithm to solve the treasure hunt game’s pathfinding problem. “One of the core strategies of any AI movement system is pathfinding strategies” (Barnouti et al., 2016). The goal of the intelligent agent is to find an optimal path to reach the treasure. Remember that the intelligent agent needs to know its environment and the algorithm to follow. Therefore, it requires all the initial preparation before it can start solving the problem. As a result, the first step is to define the game environment, including the states (different configurations or positions within the game), actions (possible moves in each state), rewards (positive rewards for reaching desired states or the goal (treasure) and negative ones for undesired states), and final goal (capturing the treasure). Initializing the neural network (Q-network) is the next step of the process since Deep Q-learning utilizes the neural network to estimate the output Q-values for each possible action based on the input state. Also, the intelligent agent needs to know the learning rate, discount factor (gamma), epsilon-greedy strategy (exploration-exploitation trade-off), and the size of the experience replay buffer for solving this problem. At this point, the intelligent agent can start gathering experiences through exploration. By randomly selecting actions (based on a certain probability), the intelligent agent explores the environment to discover the optimal policy. Now, the intelligent agent will repeat the following steps in a sequence until either it reaches a predefined number of episodes or converges to an optimal policy. Those sequence steps are as follows:

* For each episode, set the environment to its state, select an action using the epsilon-greedy strategy, perform the selected action, observe the next state and the associated reward, and store the experience (state, action, reward, next state) in the experience replay buffer.
* Choose some experiences from experience replay as samples to train the Q-network. Use the observed rewards and predicted Q-values to update the Q-values.
* Compute the loss function and then use it to train the Q-network on the previously sampled experiences from experience replay. Update the network weights and backpropagate the policy gradients.

There are similarities in the fundamental strategies humans and intelligent agents use to solve the treasure hunt game. “A truly smart and fully autonomous agent creation to handle a complex game can be as challenging as replicating a large part of human intelligence” (Safadi et al., 2015). Humans and intelligent agents must understand the game environment and its goal to choose the appropriate strategy for effective and efficient game navigation. More importantly, both utilize problem-solving, learning (from their environment), and adaptability to capture the treasure successfully. However, they execute those strategies differently. Humans usually use their cognitive abilities, creative thinking, and intuition to solve a maze, while intelligent agents follow systematic approaches to analyze the game that involves algorithms and machine learning. Humans can rely on subjective judgment and emotional intelligence to select solutions and are highly prone to errors, while intelligent agents utilize predefined criteria to prioritize their optimal solutions.

“Exploration is trying new possibilities to find better rewards. Exploitation is to keep opting for the same actions that have given some significant rewards” (Makone, 2021). Exploitation is a strategy for maximizing immediate rewards based on the existing information. Exploitation enables the intelligent agent to select actions known to be effective or have resulted in success in the past by exploiting its current knowledge. In contrast, exploration is a strategy for discovering more optimal or potentially better solutions that can be immediately unclear. Exploration enables the intelligent agent to gather more information about the environment by trying new routes or actions. In the game’s early stages, the intelligent agent must have more emphasis on exploration because it is essential for gathering information about the game environment and potential routes to the treasure. As the intelligent agent continues to collect and accumulate knowledge, it should put more emphasis on exploitation to prioritize routes with higher success possibilities based on learned experiences to maximize efficiency. However, it is essential to maintain a balance between exploitation and exploration. Relying too much on exploitation can cause the intelligent agent to get stuck in local optima and miss out on potentially better solutions. Excessive exploration, on the other hand, may lead to inefficiency and wasted resources. Consequently, achieving a balance between exploitation and exploration enables the intelligent agent to combine the advantages of exploiting known routes with the necessity of exploring new ones to find the optimal route for capturing the treasure.

Reinforcement learning provides an adaptive and dynamic framework that helps the intelligent agent make decisions based on each interaction with the game environment. “Reinforcement Learning (RL) is an approach to learning from interaction” (Singhal, 2018). The intelligent agent interacts with the game environment by taking actions (selecting routes), receiving feedback (rewards or penalties), and adjusting its strategy over time. Reinforcement learning helps the intelligent agent to learn optimal policies for selecting actions that lead to the highest cumulative rewards over time. In the treasure hunt game, the intelligent agent starts exploring the game environment to understand its dynamics and discover available routes. After learning more about the game environment, the intelligent agent exploits its knowledge to identify known efficient routes. The reward helps the intelligent agent to reinforce behaviors that contribute to finding the treasure. At this point, the intelligent agent refines its pathfinding strategy using iterative exploitation and exploration processes to lead to the most effective route based on its learned experiences that maximize efficiently reaching the treasure.

When handling complex state spaces, Q-learning struggles to produce desired results. Deep Q-learning leverages the power of neural networks to overcome the Q-learning problem-solving deficiency in complex and high-dimensional environments. Deep Q-learning utilizes experience replay to efficiently learn from past experiences and addresses instability concerns associated with Q-learning. Also, Deep Q-learning’s ability to generalize across states makes it suitable for more advanced and challenging tasks. In the treasure hunt game, Deep Q-learning predicts Q-values of actions based on the current state of the environment. The intelligent agent performs a randomly selected action, observes the reward and the next state, and then updates the environment with this new result. Initially, the intelligent agent relies more on exploration to gather information about the environment. As the intelligent agent accumulates some knowledge, it focuses more on exploiting its accumulated knowledge to maximize the immediate rewards. The intelligent agent repeats the process for its predefined number of episodes and saves its experiences in the experience replay. Then, it utilizes its experiences to train the model. And finally, it finds an optimal route to capture the treasure.

**References**

Barnouti, N. H., Al-Dabbagh, S. S. M., & Naser, M. a. S. (2016). Pathfinding in Strategy Games and Maze Solving Using A\* Search Algorithm. *Journal of Computer and Communications*, *04*(11), 15–25. https://doi.org/10.4236/jcc.2016.411002

Engström, M. (2019, March 8). Building an AI to navigate a maze. *Medium*. https://magnus-engstrom.medium.com/building-an-ai-to-navigate-a-maze-899bf03f224d

Fitchett, A. (2019, March 26). *The Porteus Maze & the Brain*. Maze Engineers. https://conductscience.com/maze/the-porteus-maze-the-brain/

Khurana, R. (2023, June 1). The Differences Between Artificial Intelligence and Human Intelligence. *Scaler Topics*. https://www.scaler.com/topics/differences-between-artificial-intelligence-and-human-intelligence/

Makone, A. (2021, August 22). Reinforcement Learning 6: Exploration vs Exploitation. *Medium*. https://ashutoshmakone.medium.com/reinforcement-learning-5-exploration-vs-exploitation-c1bae5a2ea42

Safadi, F., Fonteneau, R., & Ernst, D. (2015). Artificial Intelligence in Video Games: Towards a Unified Framework. *International Journal of Computer Games Technology*, *2015*, 1–30. https://doi.org/10.1155/2015/271296

Singhal, R. (2018, December 13). Reinforcement Learning: A Beginner’s Guide. *Medium*. https://medium.com/@rsinghal757/reinforcement-learning-a-beginners-guide-56de9c7fe1d9