The Intelligent agent uses deep Q-learning algorithm, and it uses this to navigate the maze. The overall goal is for the agent is to find the hidden treasure and navigate the maze by finding the best path. It does this through the combination of exploration and exploitation. The algorithm overview is starts with the environment setup. This environment setup is in a class that represents the maze, the class being ‘treasure Maze. This uses the ‘qmaze’ instance for the agent to interact with the maze. The ‘GameExperience’ is a class that gets used by the agent to store the data experience that has been created by the agent through the process of playing the game. With out this the neural system would not be able to be created. The experiences would involve state, action, reward, next state and game status. A neural network is used for the agent, this allows the agent to gage the value function, this is what gets the expected reward for each of the actions in a state then estimates the reward it expected to get. During each epoch for the loop the agent gets a random position to start at and then begins to navigate through the maze, as it goes it will then store the experiences that is being made and update the Q value estimates while this training loop persists.

In detail the code initializes the maze and the experience replay and then the neural model will set-up the predict Q-values. The exploration and exploitation begin with probability that the agent chooses randomly and then chooses the action based on the highest predicted Q-value. From here the agent will begin to play the game, through the game it will gain new states and get rewards. The experiences get stored into the replay memory. Then, at random, the replay memory will sample the experiences and the neural network gets trained based on these. This is done to reduce the predicted and actual Q-value similarities. Through this process the agent gets better and begins to choose exploiting over exploring. This is important to learn and adapt to win the game. If the win rate is good and consistent the training will then be considered complete.

Humans rely on making choices that are often decided by the subjective thing sthat come up that cannot be quantified, and they are not based only on figures and data. In other words, a human would observe the maze and take in the potential paths, likely picking the path that appears the easiest or most logical based on the information they have. They would plan the route that they think they’ll go in based on the information that they have and then they would use the (“always keep the wall to your right”) logic. That is when action comes in and the human starts to navigate the maze one step at a time. Through this process the human is likely to change things based on the progress and information they are receiving as they go adjusting their path accordingly. Whenever the human runs into issues like dead end or things that are inefficient, they will remember and attempt to avoid that area or path going forward.

The AI relies on data that has been accumulated to make its choices. This makes the decision making very objective. In this case the agent will be placed with no knowledge of the maze. With random actions the agent will get data as it goes about the maze. With reinforced learning the agent will update the knowledge that it has, with that updated knowledge it will make actions based on the rewards. Through time the agent uses what it has gained in data to make better choices and reach the ultimate goes. Going from exploration to exploitation, the agent will exploit the maze.

They have differences but at the beginning they are both set to explore the environment and learn through experience leading to exploitation. Some differences would be the human cognitive strategies because we use heuristic. Machines use data and systematic exploration. The machine relies heavily on data and information taken. Humans make choices on emotion, cultural factors and personal biases. While again, machines rely on the data. Humans use reasoning and intuition while a machine will use rewards and penalties. There are big differences in how machine and humans approach tasks and how they complete them.

Exploration and Exploitation are two very different things that are both very much necessary to learning and development in humans and in machine learning. Exploration is the process of exploring and gaining information or data to get more information on the environment, in this case a maze. The exploring factor for machine learning is to start at random and begin exploring through a reward process to create data and store that data for best results in the future. Exploitation on the other hand is taking all that information gained through the exploring state and using that to get the best and highest reward by exploiting the numbers. Based on reading the ideal balance between the two would be epsilon close to 1 and a decaying it to a 0.05 as a performance improves.

With deep Q-learning it combines the Q-learning to deep neural networks to handle the spaces of the large state. For navigating the maze this is the ideal way because the maze is a complex problem, and this is good for complex problems. A neural network is used and approximate the Q-value function. taking the experience replay it stores its experience and, samples It and does this to breaks the correlation between consecutive experience. Then a target network is implemented to maintain the training. The training is done through batches of experiences, this keeps the loss minimal between predicted Q-values and target Q-values. In this code the deep Q-learning us effectively used to solve the maze pathfinding problem. Through balance of exploration and exploitation the agent can learn what path is the best to the treasure due to reinforced learning. This Displays the combined power of neural networks with a good original Q-learning technique. I found this code to be veery effective and efficient during testing even when using apporto the few times it would work.

Citations

* 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/#
* Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction* (2nd ed.). MIT Press.
* Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. *Nature, 518*(7540), 529-533. <https://doi.org/10.1038/nature14236>
* Zhao, M. (n.d.). *Understanding humans’ strategies in maze solving from eye-hand coordination*. Welcome to RUcore - Rutgers University Community Repository. https://rucore.libraries.rutgers.edu/rutgers-lib/45599/