When humans solve a maze problem there are a couple of different actions taking place. Eye movement to discover a route is important. Hand eye coordination becomes a factor to move the mouse where the subject is looking. Memory plays a factor by remembering where they have been but is shared in new areas to explore. There are a couple different strategies to use but they all have these same attributes. “available visual information, memory, confidence, the estimated cost in time for exploration, and idiosyncratic tolerance for error”(Zhao&Marquez, 2013) For a human to solve this problem it would be a very simple procedure. In a quick glance the path from the starting point to the finish is simple to extrapolate. This is because humans have trained their brains for much more complex problems.

The intelligent agent created to solve this maze was able to do so fairly quickly. This is in part of the simplicity of the maze. Without having eyes. The agent must try to solve the maze by exploration and relying on memory. When the agent runs into a barrier, there is a negative reward given so that the agent will learn that colliding with a barrier is a bad decision. When the agent does not collide with a barrier there is a positive reward given. The agent will continue to get positive rewards until they arrive at the assigned grid. Upon arrival at the assigned grid locations a larger reward will be given so that the agent stores that as the most important thing.

Humans and agents both share some common approaches. Both use exploration, input, and trial and error. The way the input is received is different from a machine to a human.

“In reinforcement learning developers devise a method of rewarding desired behaviors and punishing negative behaviors. This method assigns positing values to the desired action to encourage the agent and negative values to undesired behaviors. ”(Carew,2023) This trains the agent to look for long term success and to maximize the rewards to achieve the optimal solution. Trial and error is the simplest way reinforcement learning helps to determine the correct path. “Reinforcement learning is a trial and error process where an AI (agent) performs a number of actions in an environment. Each unique moment the agent has a state and acts from this given state to a new one. This particular action may on may not has a reward.”(Karagiannakos,2018)

To implement deep Q-learning it started by randomly selecting a cell for the agent to start in. Every new game the agent would start in a random cell. This agent would play 15000 games. The agent would do fifty games before updating its plan. This lets each training session ensure that the new route has been tested and is indeed correct.

Reference

Zhao, M., & Marquez, A. G. (2013, July 22). *Understanding Humans' Strategies in Maze Solving*. Understanding Humans' Strategies in Maze Solving -arXiv. Retrieved February 15, 2023, from https://arxiv.org/pdf/1307.5713.pdf

Carew, J. M. (2023, February 10). *What is reinforcement learning?: Definition from TechTarget*. Enterprise AI. Retrieved February 16, 2023, from https://www.techtarget.com/searchenterpriseai/definition/reinforcement-learning#:~:text=How%20does%20reinforcement%20learning%20work,negative%20values%20to%20undesired%20behaviors.

Karagiannakos, S. (2018, September 23). *The Secrets Behind Reinforcement Learning*. AI Summer. Retrieved February 16, 2023, from https://theaisummer.com/Reinforcement\_learning/#:~:text=Reinforcement%20learning%20is%20a%20trial,may%20not%20has%20a%20reward.