The following are the steps for a human’s approach to problem solving a maze.

1. Perception – A human would first take a good look and make astute observations about the maze like making it a point to identify the beginning and end points of said maze.
2. Planning – Then a mental map would be made of possible routes to go about utilizing prior knowledge and intuition regarding previous attempts at solving mazes or anything else to help get to a solution.
3. Trial and Error – Humans often utilize trail and error when dealing with a problem like this with minor backtracking when a dead end is hit during a trail route to the end.
4. Memory and Learning – They usually remember the paths that weren’t successful and learn from these mistakes to improve their efficiency in future maze problems but also when it comes to solving the exact same maze again.

The following are the steps for a machine’s approach to problem solving a maze.

1. Perception – Machines utilize input data or sensors to understand the maze that needs to be solved.
2. Algorithmic Planning – They employ efficient and effective algorithms like A\* or Dijkstra’s to calculate the shortest path possible to the end point.
3. Systematic Exploration – They utilize a system that systematically explores all paths which is often exhaustively more than humans.
4. Learning – By utilizing reinforcement learning they learn optimal paths to take over time through a reward system that helps determine better choices vs mediocre choices.

Both have similarities and differences it just takes some knowledge to figure them out. Both involve perception, planning, and learning. While machines are more brute force through more systematic and exhaustive approaches since humans rely on intuition and experience that machines don’t have. Also, since machines are capable of processing larger amounts of data quickly humans excel with abstract reasoning.

The purpose of the intelligent agent is to efficiently find the most optimal and shorted path from beginning to the goal within a maze usually known as the end. The process is automated reducing both human effort and error while allowing them to adapt to dynamic environments if they are initially designed with learning capabilities. Exploitation has to do with using known information to make the best decision because in pathfinding this means knowing the best path. Exploration has to do with trying new paths to discover potentially better routes that were previously unknown before. The balance that is most ideal depends heavily upon an environment’s complexity along with their dynamics. Within a static maze more exploitation is probably most beneficial once a good path is found. However, dynamic environments require a balanced approach to be necessary to adapt to changes in the environment. The way in which reinforcement helps agents is by learning optimal paths by receiving rewards for reaching specific goals and likewise receiving penalties for dead ends or bad choices. Over time, the agent improves from previous choices by maximizing cumulative rewards effectively allowing them to learn the best path to the treasure.

The following are the steps taken to implement Q-learning utilizing neural networks for this problem.

1. State Representation – A state space is defined which represents the agent’s position within the maze.
2. Action Space – All the actions that are possible for an agent. (e.g., up, down, left, or right)
3. Neural Networks – Utilizing neural networks to approximate the Q-value functions which predicts what the expected reward for each action with be in each state.
4. Training – The network is trained using experiences such as state, action, rewards, and next states and then updating the Q-values utilizing the Ballman equation.
5. Exploration Strategy – Implementation of an exploration strategy like epsilon-greedy to be used to balance exploration and exploitation.
6. Optimization – Utilizing techniques like experience replay and target networks to stabilize the agents training.

Works Cited

1. Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement learning, fast and slow. *Trends in Cognitive Sciences*, *23*(5), 408–422. https://doi.org/10.1016/j.tics.2019.02.006