It is interesting to think about the differences of how a human and a machine solve a problem such as a maze. Humans tend to be aware of what their goal is and have an immediate understanding of how they should try to obtain that goal. However, the intelligent agent's approach is algorithmic, and uses reinforcement learning to decipher the best path. When considering human-solving steps for the maze, one would typically look at the environment as a whole and see where to avoid and which path deems most optimal. They could think about the outcomes of each direction before making their first move. The intelligent agent, however, follows a structured algorithmic process. For example, in the case of reinforcement learning, the agent explores different paths while updating its knowledge through reward signals (Sutton & Barto, 2018).

The similarities between these approaches lie in the common goal: finding an optimal path. They both are trying to reach the end as efficiently as possible. The differences are humans’ ability to change decisions without a substantial amount of previous data and the ability to come up with more creative solutions. Whereas the intelligence agent explores the environment while updating its knowledge and learns from receiving rewards or penalties.

Exploitation refers to selecting known optimal actions, while exploration involves trying new actions to discover better strategies (Sutton & Barto, 2018). The ideal proportion between these two varies with the problem's context. In pathfinding, early on, more exploration might be beneficial to gain data of the environment. Over time, however, a balance shifts toward exploitation as the agent refines its strategy. For me, using an exploration factor of ten percent seemed to show optimal results.

Reinforcement learning can help determine the path to the goal by rewarding and penalizing the agent’s actions. The agent will naturally try to minimize penalties and maximize rewards which in turn will be finding the most optimal path to the treasure. The agent learns through exploration, receiving rewards for good actions, and adjusting its strategy accordingly.

In this project, I implemented deep Q-learning using neural networks to optimize the agent's pathfinding. The agent observes its environment, takes actions, receives rewards, and updates its Q-values through neural networks.

In conclusion, the approach to pathfinding encompasses both human and algorithmic intelligence. The intelligence agent used a small percent of exploration to learn its environment and used exploitation to optimize its moves based on the rewards system. The result of this reinforcement learning led the agent to find the optimal path to the treasure and do it consistently.

**Resources:**

**Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT** Press**.**