**CS 370 Project Two: Design Defense**

Harshilkumar Jayswal

**Differences between human and machine approaches to solving problems**

Humans can solve the maze with visual observation, where individuals rely on their senses to understand the maze's layout. This initial step involves not only perceiving the physical barriers and pathways but also mentally simulating potential routes towards the goal. They draw upon their cognitive abilities to plan and strategize, considering factors such as dead ends and obstacles. Decision-making in this context is influenced by a blend of prior knowledge, intuition, and problem-solving skills. Humans possess the unique ability to foresee potential dead ends based on their mental map of the maze, allowing them to make decisions that minimize the risk of getting trapped and maximize the likelihood of reaching the goal.

While intelligent agents approach maze-solving through a computational perspective, leveraging algorithms and machine learning techniques. These agents receive input representing the maze's structure and employ a balance between exploration and exploitation. Through Q-learning, the agent assigns values to different actions based on their expected outcomes in specific environmental states. This process allows the agent to probabilistically select actions that maximize the expected cumulative reward over time. They benefit from experience replay, a technique where past experiences are stored and replayed during training to enhance learning efficiency. They follow systematic algorithms and computational techniques to make decisions, leveraging data-driven insights to optimize their pathfinding strategies.

**Assessing the Purpose of the Intelligent Agent**

For an intelligent agent’s problem-solving efficiency there needs to be a balance between Exploration and Exploitation. Exploration allows the agent to discover new paths and gather information about the maze's layout, while exploitation involves leveraging known information to maximize rewards. The ideal proportion of exploitation and exploration depends on the agent's stage of learning and the complexity of the maze. Initially, a higher emphasis on exploration facilitates comprehensive understanding of the environment, while gradually shifting towards exploitation as the agent accumulates knowledge and identifies optimal pathways.

Reinforcement learning serves as a foundational framework for the intelligent agent's path to the goal. By associating actions with expected rewards, reinforcement learning enables the agent to iteratively refine its strategy towards achieving the goal while navigating through the maze. Through trial and error, the agent learns from its interactions with the environment, adjusting its behavior based on the received rewards and environmental feedback. This adaptive learning process empowers the agent to effectively explore the maze, identify optimal paths, and ultimately reach the desired goal.

**Implementing Deep Q-Learning**

First, we set up a neural network, which acts as the brain of our agent, to predict the value of different actions in various situations. Next, we represent the game environment using a grid, where each cell has distinct properties like walls or paths. Then, we create a system to remember the agent's experiences during training, akin to keeping a journal of past events to learn from.

As the agent plays through the game multiple times, each playthrough representing an "epoch," it makes decisions, moves around, and collects rewards based on its actions. Just as we weigh different options before making a decision, the agent decides what action to take based on its learning, sometimes trying new strategies and other times relying on what it knows works well. After each action, the agent learns from its experiences, storing this information in its memory to inform future decisions. Periodically, the agent reviews its experiences and updates its understanding of which actions are best in different situations. Throughout training, we monitor the agent's progress, observing metrics like win rate, learning speed, and overall improvement. Training continues until the agent consistently achieves its goal, whether it's finding the treasure at the end of the maze. Once training is complete, we evaluate the agent's performance on new challenges, assessing how well it adapts to different situations and how efficiently it learns. Through this iterative process, deep Q-learning enables our agent to learn to navigate the maze effectively, resembling how humans learn through experience and practice.

**References**

Beysolow, T. (2019). Applied Reinforcement Learning with python: With Openai Gym, tensorflow, and keras. Apress.

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](https://doi.org/10.1016/j.tics.2019.02.006%20)

Dugmeci, D. (2021, November 10). Finding shortest path using Q-learning algorithm. Medium. [https://towardsdatascience.com/finding-shortest-path-using-q-learning-algorithm-1c1f39e89505](https://towardsdatascience.com/finding-shortest-path-using-q-learning-algorithm-1c1f39e89505%20)

Lamba, A. (2018, September 3). An introduction to Q-learning: Reinforcement learning. Medium. [https://medium.com/free-code-camp/an-introduction-to-q-learning-reinforcement-learning-14ac0b4493cc](https://medium.com/free-code-camp/an-introduction-to-q-learning-reinforcement-learning-14ac0b4493cc%20)

Stern, R., Sturtevant, N., Felner, A., Koenig, S., Ma, H., Walker, T., Li, J., Atzmon, D., Cohen, L., Kumar, T. K. S., Boyarski, E., & Bartak, R. (2019, June 19). Multi-agent pathfinding: Definitions, variants, and benchmarks. arXiv.org. <https://doi.org/10.48550/arXiv.1906.08291>