**Module 7 Project Two Design Defense**

**Michael Reynolds**

**Analyze the differences**

When solving the maze, a human would most of the time start by assessing the maze visually, visually a human may behind the scenes be ‘trying attempts’ in their head but won’t move until finding what they consider the most direct path, then begin moving in that path either reaching the end or hitting a dead end and backtracking.

An intelligent agent solving the maze may instead, and in this case, start by utilizing parameters to determine sporadic movement, they’ll be fed ‘reward’ to determine correct pathing. The agent will try many different times with initially randomized input, and then ‘remember’ what worked, this will then allow the agent to have a memory regarding the most rewarding possible solutions.

There are strong similarities between these approaches, but the differences are also notable. In comparison to the agent, humans generally will “abstract” their solutions, their analysis of the problem and possible solutions can happen almost entirely through familiarity and ‘simulation’ rather than execution. Agents rely on execution and memory learned within the environment of the problem. Humans can do trial and error and can even fail to remember a solution, but Agents can achieve 100% accuracy as they can have solid memory regarding solutions. Humans being more organic can be faulty, but this allows them to, at least to date, adapt more quickly in ever-changing environments.

**Assess the purpose of the intelligent agent & Q-learning**

When it comes to pathfinding, exploitation means making choices based on existing knowledge, in this case that applies to our agent’s memory, or a human’s memory. Exploration instead involves trying new moves/solutions to discover new and hopefully better paths. The ideal proportion of exploitation and exploration depends on the problem at hand, there are different parameters involved like an agent's initial knowledge and the environment involved. Commonly the move is to start with exploration to learn about the environment and then gradually incorporate exploitation from memory. For example, using an epsilon-greedy approach with an epsilon value of 0.1 means that 10% of the time, the agent explores, and 90% of the time, it exploits. The balance can be changed through experimentation and tuning based on the complexity of the environment and the agent's goals.

Reinforcement learning can help an agent determine the path to a goal by defining states, defining actions, assigning rewards, etc. The agent explores different actions in different states to learn which actions yield the highest rewards and will use this to shift towards exploitation as it learns. Ultimately, this adoption of exploitation guides the agent to find the optimal path to the goal, such as the treasure. The key is to set up the reinforcement learning problem correctly, to enable the agent to discover the best path to the goal.