CS-370-12784-M01 Current/Emerging Trends in CS

Adam Vosburg

Southern New Hampshire University

Project Two – Design Defense Plam

02/17/2025

**Project Two – Design Defense Plan:**

**Human vs. Machine: Problem Solving Approaches**

When solving a maze, humans typically employ a combination of visual processing and intuitive strategies. The human approach begins with an initial visual survey to understand the overall maze layout. During this process, humans create mental maps of potential routes from start to finish. Additionally, we will often employ techniques such as wall-following or looking for obvious pathways. When encountering dead ends, humans naturally backtrack and learn from their previous attempts through spatial memory. Humans excel at pattern recognition and can quickly eliminate obviously poor paths, but are limited by working memory capacity and may struggle with systematic exploration of all possibilities (Russell & Norvig, 2021).

Comparatively, the designed and implemented intelligent agent uses a more systematic approach through Q-learning. The agent processes the maze environment as a series of states and possible actions, without the benefit of visual pattern recognition that humans possess. The agent learns through state-value associations and repeated interactions with the environment, storing experiences in memory and using them to improve future decisions (Sutton & Barto, 2018). This systematic approach allows the agent to explore possibilities that humans might overlook due to cognitive biases.

The similarities between human and machine approaches include the fundamental concept of learning from experience and using feedback from previous attempts to improve future performance. However, the differences are significant and measurable. While humans rely on intuitive pattern recognition, agents use systematic value estimation through mathematical calculations. Humans have limited memory but excel at generalization, whereas agents can store vast amounts of experience but require explicit training to generalize effectively. Furthermore, humans can apply previous maze-solving experiences immediately to new mazes, while agents typically start fresh with each new environment.

**Assessing Purpose in Intelligent Agent Pathfinding**

In reinforcement learning, the concepts of exploitation and exploration represent a fundamental trade-off in the learning process. Exploitation refers to the agent's tendency to choose actions it knows will yield good results, while exploration involves trying new actions to potentially find better solutions (Mnih et al., 2015). In this implementation, I manage this trade-off through an epsilon-greedy strategy. This is where the exploration rate starts high and gradually decreases as the agent becomes more confident in its learned values. The ideal proportion shifts during training, with higher exploration early to discover potential paths. It will then eventually reduce to favor exploitation once the environment is well-understood.

Reinforcement learning helps determine the optimal path through a carefully designed reward system. The agent receives a positive reward for reaching the treasure, penalties for invalid moves and boundary violations, and a small movement cost to encourage efficient paths. Through experience replay, the agent stores and reuses past experiences, making efficient use of training data and enabling stable learning through randomized batch sampling (Lin, 1992). This approach allows the agent to learn from both successful and unsuccessful attempts, gradually building a comprehensive understanding of the environment and optimal pathfinding strategies.

**Evaluating Algorithms in Problem Solving**

This implementation combines deep learning with Q-learning through a sophisticated neural network architecture. I designed a system that effectively integrates experience replay and neural network function approximation. The neural network takes the current state as input and predicts Q-values for each possible action (Hasselt et al, 2015). This architecture allows the agent to generalize across similar states and make informed decisions in previously unseen situations.

The training process involves randomly initializing starting positions to ensure good generalization, with batch updates for efficient learning. The agent uses temporal difference learning to update its value estimates, combining immediate rewards with estimated future returns. This approach allows the agent to learn optimal pathfinding strategies through trial and error while gradually improving its performance until it achieves consistent success in reaching the goal.

**References**

Hasselt, H. V., Guez, A., & Silver, D. (2015). Deep reinforcement learning with double Q-learning. Proceedings of the AAAI Conference on Artificial Intelligence, 30(1), 2094-2100. [[1509.06461] Deep Reinforcement Learning with Double Q-learning](https://arxiv.org/abs/1509.06461)

Lin, L. J. (1992). Self-improving reactive agents based on reinforcement learning, planning and teaching. Machine Learning, 8(3), 293-321. <https://doi.org/10.1007/BF00992699>

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., ... & Hassabis, D. (2015). Human-level control through deep reinforcement learning. Nature, 518(7540), 529-533. [Human-level control through deep reinforcement learning | Nature](https://www.nature.com/articles/nature14236)

Russell, S., & Norvig, P. (2021). Artificial Intelligence: A Modern Approach (4th ed.). Pearson. [Artificial Intelligence: A Modern Approach, 4th US ed.](https://aima.cs.berkeley.edu/index.html)

Sutton, R. S., & Barto, A. G. (2018). Reinforcement Learning: An Introduction (2nd ed.). MIT Press. [SuttonBartoIPRLBook2ndEd.pdf](https://web.stanford.edu/class/psych209/Readings/SuttonBartoIPRLBook2ndEd.pdf)