Carl Allen

April 17, 2025

Current/trends emerging

Design Defense for Pirate Intelligent Agent in Treasure Hunt

**Introduction**

The purpose of this project was to develop an intelligent agent—a pirate—capable of navigating a Treasure Hunt Game environment and efficiently locating treasure. The agent was designed using reinforcement learning techniques, which allow it to learn optimal behaviors through trial and error interactions with the game environment. This document outlines the core design decisions, agent architecture, learning approach, and challenges encountered throughout development.

**Design Decisions**

The design of the pirate agent was centered on maximizing the treasure discovery rate while minimizing steps taken. After evaluating multiple options, Q-learning was selected as the reinforcement learning algorithm due to its simplicity and effectiveness in grid-based environments. A discrete state-action representation aligned well with the game board layout, where each grid cell represented a unique state.

To encourage the agent to explore the environment without getting stuck in suboptimal paths, an epsilon-greedy strategy was implemented. This approach allowed the agent to balance exploration of unknown areas with exploitation of known high-reward paths.

**Agent Architecture**

The pirate agent operates in a 2D grid-based environment, where each state corresponds to the agent’s location. The available actions include moving **up**, **down**, **left**, or **right**. The environment responds to these actions by transitioning the agent to a new location and providing a reward: +1 for treasure, -1 for invalid moves (e.g., walking into a wall), and 0 otherwise.

The agent maintains a Q-table, a matrix that maps state-action pairs to expected future rewards. As the agent plays more episodes, it updates the Q-values using the Bellman Equation:

Q(s,a)←Q(s,a)+α[r+γmax⁡a′Q(s′,a′)−Q(s,a)]Q(s,a)←Q(s,a)+α[r+γa′max​Q(s′,a′)−Q(s,a)]

Where:

* ss is the current state
* aa is the action taken
* rr is the reward received
* s′s′ is the resulting state
* αα is the learning rate
* γγ is the discount factor

**Learning Strategy**

The agent was trained over multiple episodes, each consisting of a full run through the environment until the treasure was found or a maximum number of steps was reached. To promote learning, the exploration rate (epsilon) was gradually decayed over time, encouraging the agent to explore early on and exploit learned knowledge later.

Hyperparameters were tuned as follows:

* Learning Rate (α): 0.1
* Discount Factor (γ): 0.9
* Initial Epsilon: 1.0 (decaying to 0.01)
* Episodes: 500 to 1000, depending on performance

The training progress was monitored by tracking the number of steps taken to find treasure and the agent’s average reward per episode.

**Challenges and Improvements**

One of the main challenges was preventing the agent from looping or revisiting the same areas unnecessarily. This was mitigated by assigning a small negative reward to non-progressive moves and ensuring valid boundary conditions for actions.

Additionally, sparse rewards made it difficult for the agent to learn in early episodes. To address this, a shaping reward system was briefly introduced to guide the agent toward the treasure before transitioning to sparse rewards for realism.

Future improvements could involve:

* Switching to Deep Q-Networks (DQN) for continuous or large-scale environments.
* Adding obstacles or dynamic treasure locations.
* Introducing multiple agents and cooperative behavior.

**Conclusion**

The pirate intelligent agent demonstrated the effectiveness of reinforcement learning for decision-making in a simulated environment. Through iterative learning, it became proficient in navigating the grid and locating the treasure. This project provided valuable insight into agent-based modeling, learning strategies, and the practical challenges of developing intelligent systems.

**References**

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

Russell, S. J., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.