## vi - agent - report.pdf :

***Question 2 (1 point):*** Test your Value Iteration Agent against each of the provided agents 50 times and report on the results – how many games they won, lost & drew against each of the other rule based agents. The rule based agents are: *random*, *aggressive*, *defensive*.

**Against Defensive Agent:**

Wins: 46 Draws: 4 Losses: 0

**Against Aggressive Agent:**

Wins: 50 Draws: 0 Losses: 0

**Against Random Agent:**

Wins: 50 Draws: 0 Losses: 0

### iterate :

- Implements the Value Iteration algorithm, a foundational approach in reinforcement learning.
- For each state in the `valueFunction`, it calculates the maximum expected reward achievable by taking the best possible action (move).
- Repeats this process for `k` iterations, refining the value estimates for all states.
- Terminal states are assigned a value of 0 because no further actions can occur from them.

### extractPolicy :

- Extracts an optimal policy based on the current state values (`valueFunction`).
- For each non-terminal state, it calculates the expected reward for all possible actions and selects the action that maximises this reward.
- The resulting policy maps each state to its best action, guiding decision-making to maximise rewards.

### train :

- Solves the Markov Decision Process (MDP) by combining the `iterate` and `extractPolicy` methods.
- First runs the Value Iteration algorithm (`iterate`) to compute the optimal value function.
- Then extracts the optimal policy (`extractPolicy`) from the value function and assigns it to the agent.
- Ensures the agent can act optimally in its environment after training.