Reinforcement Learning: Training an Agent to Balance a Pole

Ian McConnell

Introduction

- The agent's goal in the environment is to balance a pole by moving a cart left and right on a track
- Reward is gained for every tick forward in a given episode
- The episode ends if the pole leans > 12° in either direction or 500 reward is received

1. How does the performance of an agent using a reinforcement learning algorithm compare to agents using simple solutions?
2. Is reinforcement learning a practical solution for this type of problem?

Methods

$$[Position, Velocity, Pole\ Angle, Pole\ Velocity] \longrightarrow f \longrightarrow left\ or\ right?$$

- Create an agent that selects completely random actions to use as a control
- Create an agent that implements a simple solution of moving left of the pole is leaning to the left and moving right if leaning to the right
- Create an agent that utilizes a reinforcement learning algorithm to improve its performance in the environment
- The algorithm is initialized with randomized weights which are slightly randomized in each episode. If the newly randomized weights perform better than the base weights, they take their place as the new base weights
- Record the total reward gained in each episode by every agent for interpretation
- Reinforcement learning was implemented because the environment was designed to be solved with this type of algorithm

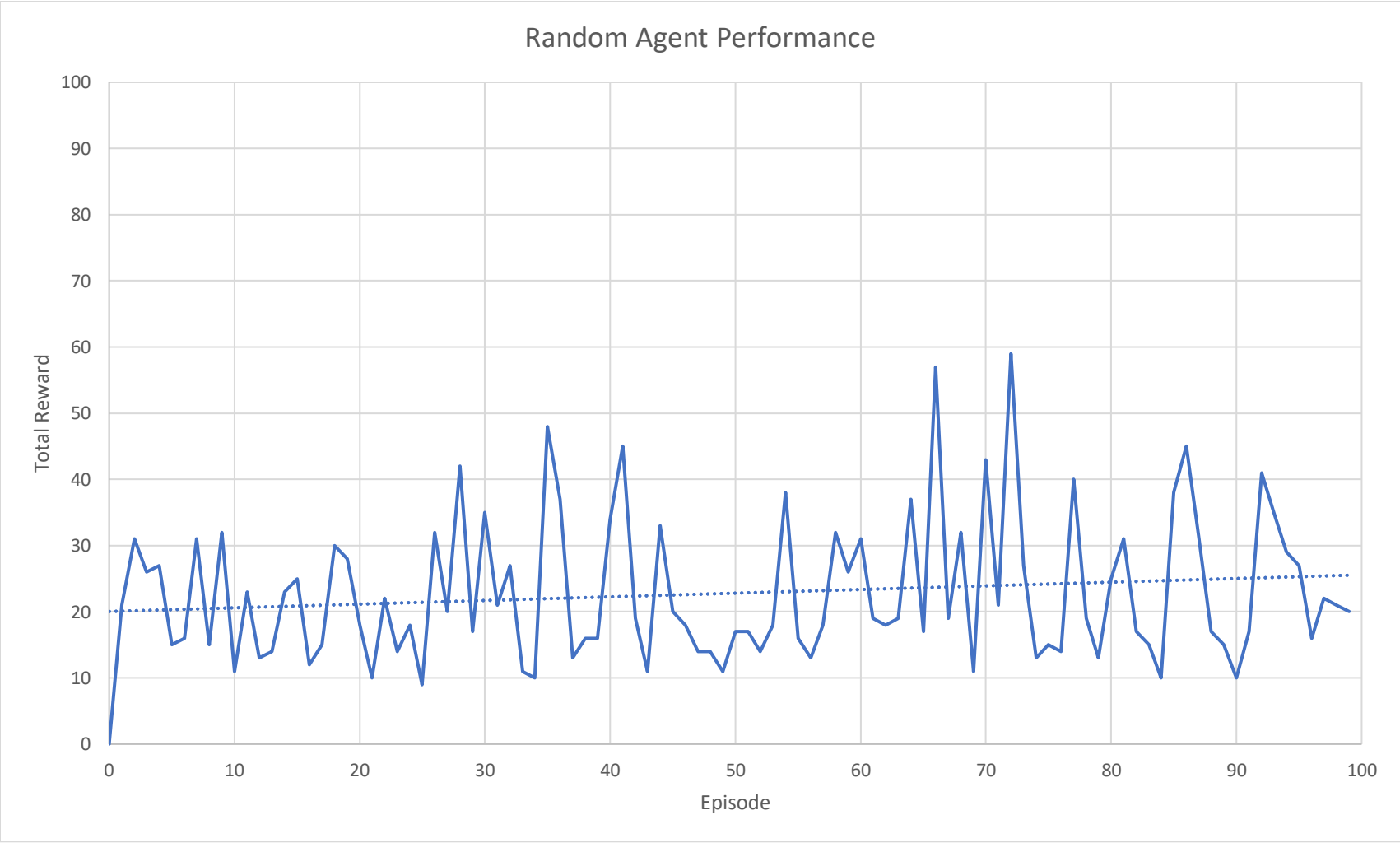


Fig 1. Graphical representation of random agent's performance generated from 100 episodes

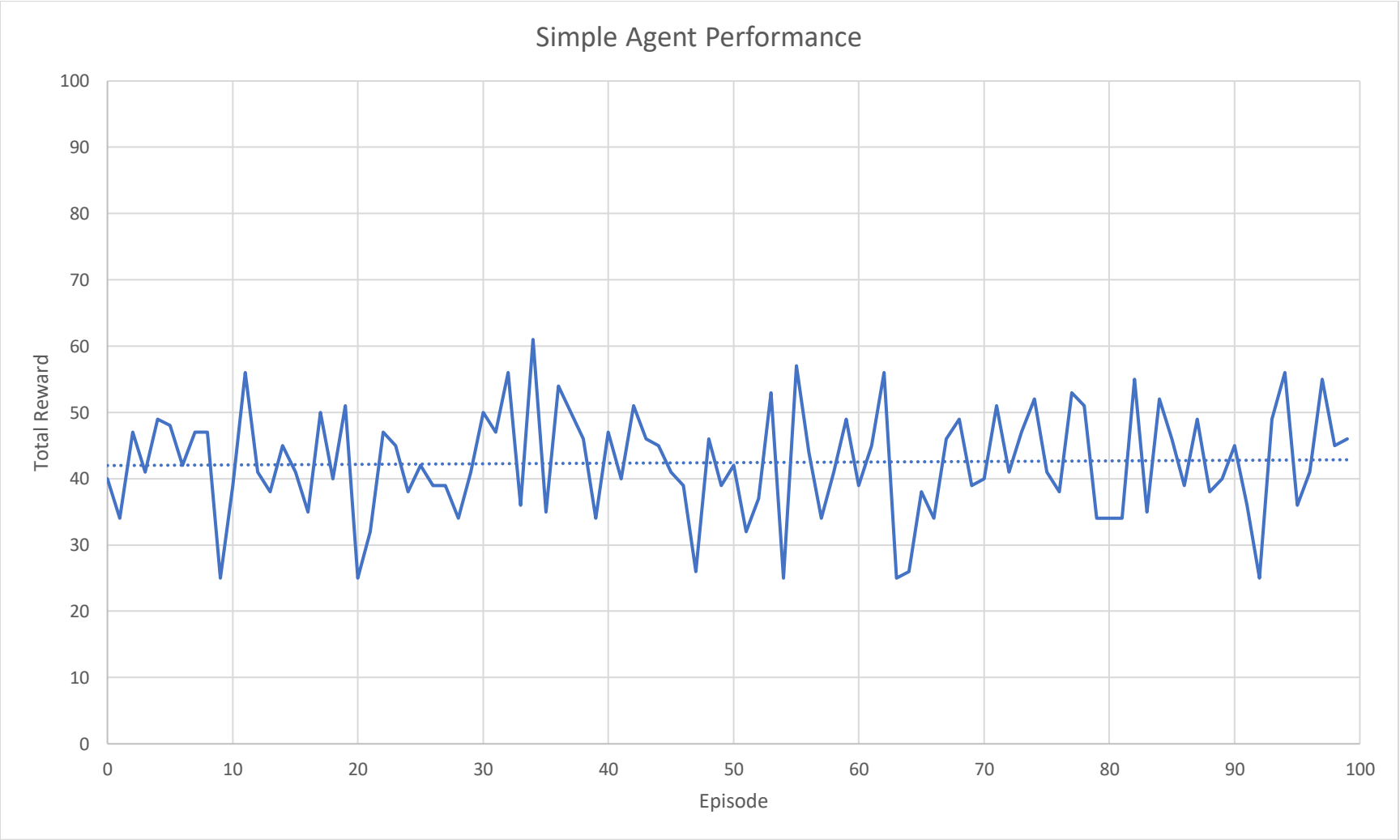


Fig 2. Graphical representation of simple agent's performance generated from 100 episodes

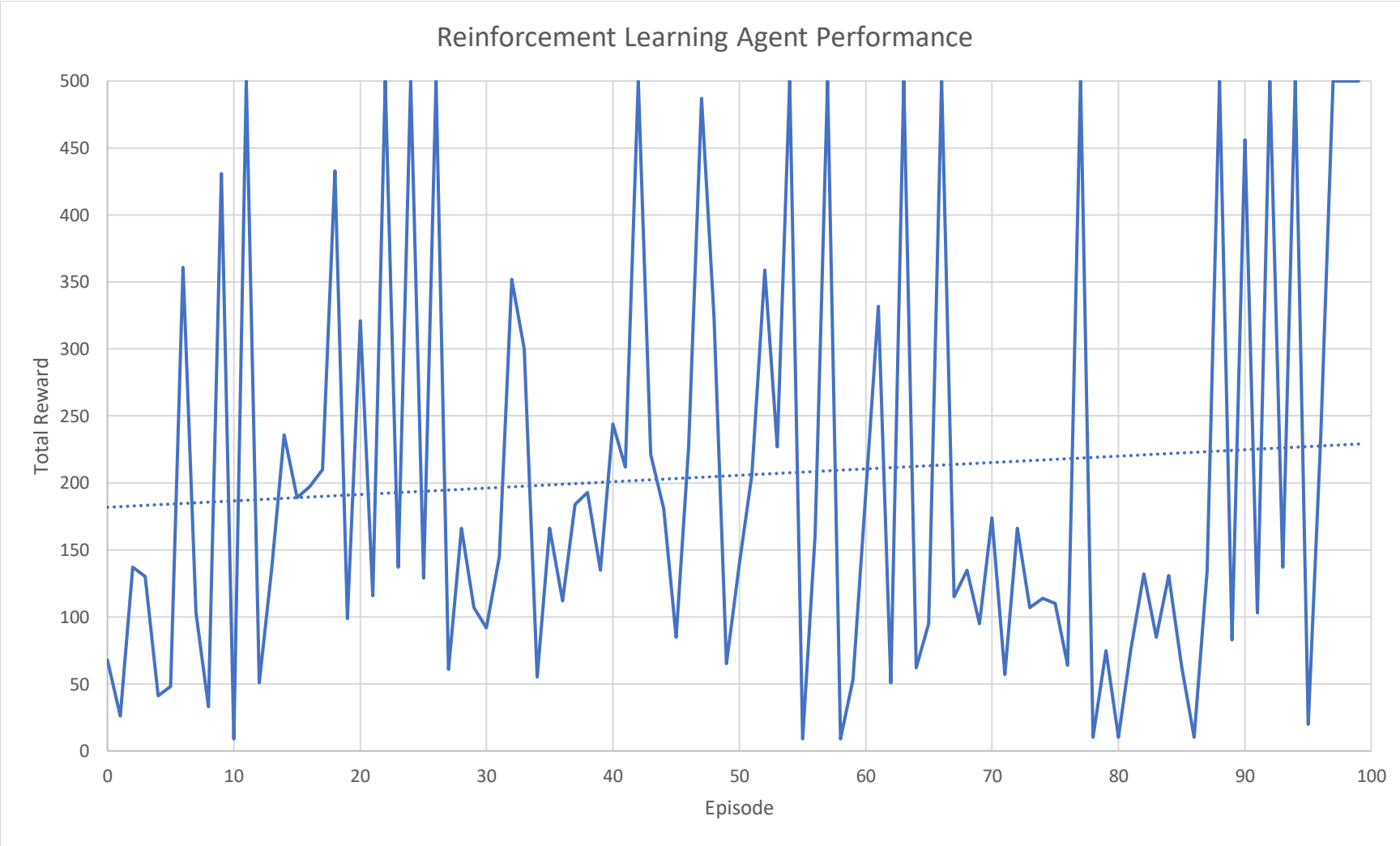
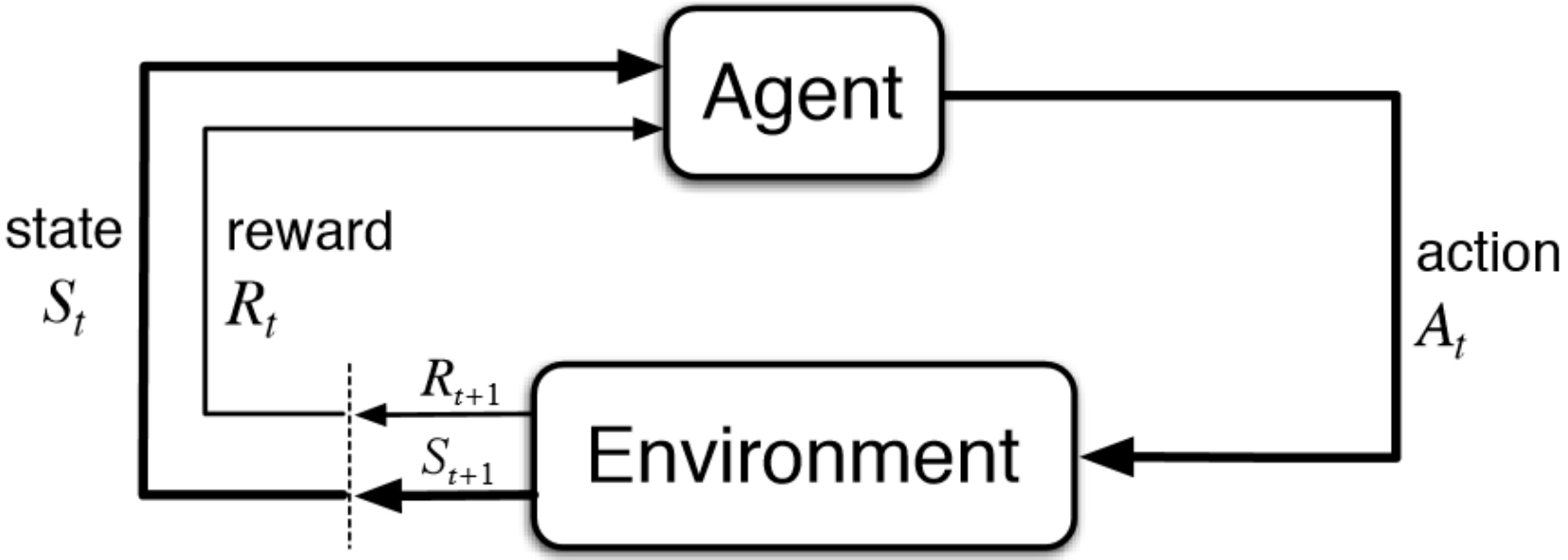
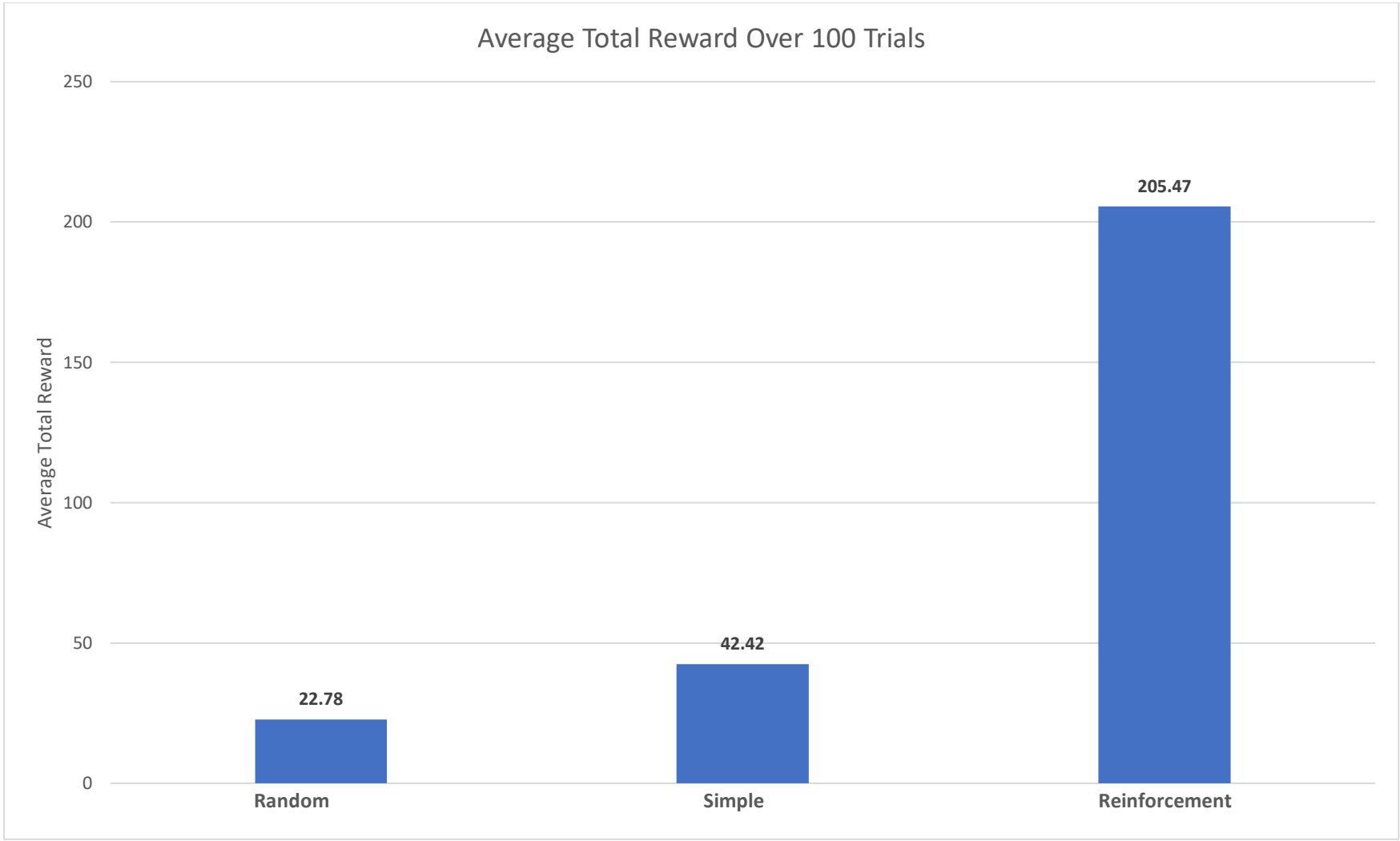


Fig 3. Graphical representation of reinforcement learning agent's performance generated during training



Reinforcement Learning Illustration (https://i.stack.imgur.com/eoeSq.png)


Average Total Reward Over 100 Trials
Average Total Reward
250
200
150
100
50
0
22.78
42.42
205.47
Random
Simple
Reinforcement

Results

- Methods were evaluated based on the average total reward received over the course of 100 episodes
- The reinforcement learning agent's average total reward grew over the course of the experiment
- The reinforcement agent greatly outperformed the random and simple agents
- Based the results, it can be concluded that reinforcement learning is a highly effective, practical approach to solving this environment