Calvin Kin Lok YEUNG
Github: https://github.com/calvinKLyeung/Calvin_3202_Taxi

Project overview

This project aims to implement the Q-Learning Algorithm and the Deep Q Learning Network (DQN) and compare the learning process and prediction result in the same environment. The implementation and the comparison is done with the Taxi game version 3 under the Toy Text environment provided by the OpenAI Gymnasium library. The Q-learning agent and the DQN agent are built using mainly NumPy and PyTorch libraries.

Approach

Environment

Description In the Taxi game environment, the Taxi agent is randomly spawned into a 5x5 grid world with 4 pick-up/drop-off locations on the Map in Red, Green, Yellow, and Blue colour respectively.
The goal of the game is to let the Taxi agent

1. navigate to the passenger's location
2. pick up the passenger
3. head to the passenger's desired destination (1 of the 4 locations)
4. drop off the passenger at that destination

End of Episode The Taxi game is terminated once the Taxi agent drops off the passenger.
An episode is truncated if the Taxi agent fails to terminate after 200 steps by default.

Game Id
"Taxi-v3"

Action Space
The Taxi agent has totally 6 actions to choose from the range of {0, 1, 2, 3, 4, 5}.
From 0 to 3, each of the number indicates the direction of action.
From 4 to 5, each of the number represents picking up or dropping off the passenger.

- 0 = Head South (Down)
- 1 = Head North (Up)
- 2 = Head East (Right)
- 3 = Head West (Left)

- 4 = Picking up the passenger
- 5 = Dropping off the passenger

Observation Space

In the 5x5 grid world, there are 25 valid position for the Taxi agent to be in.
The passenger can be in 5 unique locations {Red, Greed, Yellow, Blue, Taxi}.
And there are 4 destinations to drop off the passenger {Red, Greed, Yellow, Blue}.
In total, the Taxi game has $(5 \times 5) \times 5 \times 4 = 500$ discrete number of states with 6 possible actions.

Rewards

Incorrect attempts to pick up or drop off the passenger at the wrong location is punished with negative reward of -10.
Each step also comes with -1 negative reward as living penalty.
Successfully dropping off the passenger at the correct destination is encouraged with positive reward of $+20$.

Choice of model(s)

Q-learning Algorithm

Q-learning in reinforcement learning is essentially a lookup table algorithm. The Q-learning algorithm uses function approximator from a single experience to estimate the state-action and updates the result Q values in a lookup Q value table as reference for the agent to navigate. With the help of the epsilon greedy policy to switch between exploration and exploitation, Q-learning learns the optimal Q-values for every Q state and eventually converge to the optimal policy by taking suboptimal and random actions iteratively.

Deep-Q learning Network

The Deep-Q Network (DQN) is a reinforcement learning algorithm that combines the Q-learning Algorithm and multilayer networks (or neural networks). In essence, the DQN is the same idea as Q-learning in general but replaces the function approximator in the Q-learning with a deep neural network. The DQN takes the representation of the states in something like a one-hot encoded vector to approximate the Q value for output actions along with the weights and bias in the neural network. DQN learns from experience which consists of a state, an action, a reward, and the next state. To prevent learning from sequential experience which usually correlate with each other and lead to overfitting, DQN randomizes the training data before sending to the neural network. DQN does so by utilizing a Replay Memory buffer where a batch of random sample experience is passed on to the multilayer neural network for training only if the batch of experience is filled.

Methods and Purpose of Testing and Experiments

The project focuses on comparing the Q-learning and DQN algorithms in 4 major aspects:

- compare the average reward improvement during the training phase with controlled parameters
- compare prediction performance with controlled parameters
- compare the average reward improvement the training phase with fine-tuned parameters
- compare the prediction performance with fine-tuned parameters

Note:
Controlled parameters refer to using the same controlled value for parameters in common, like learning rate alpha, and discount rate gamma.
Fine-tuned parameters refer to manually fine-tuned and optimized parameters to maximize the reward prediction of the models.
Taxi agent is encouraged with higher reward if the shortest path and correct actions are being taken due to the living penalty, illegal action punishment, and correct action rewards.

Challenges and Trouble Shooting

Difficulty in setting up Stable Baselines for Taxi environment
Initially, the original plan of this project was to do a implementation and comparison between Q-learning, Deep-Q Network (DQN), Advantage Actor Critic (A2C), and Proximal Policy Optimization (PPO). However, after a long struggle, I was not able set the DQN, A2C, PRO models up to perform any decent prediction in the Taxi environment. The training phase of the models often ends abruptly after only a few episodes and the result prediction keeps telling the Taxi agent to run into a wall. I suspect the complex multilayer networks models do not really fit well for the Discrete and relatively small action and observation space of the Taxi problem and, hence, perform significantly worse than the Q-learning algorithm. Therefore, the project is eventually shifted its focus to the implementation and comparison between Q-learning and DQN using the NumPy and PyTorch libraries.

Designing and Implementing Q Learning Agent
Unlike the previous reinforcement learning homework where the series of Agents are already welled designed and implemented for students to only fill in the blanks of the Q-learning algorithm, this project needs to recreate a Q-learning agent in a completely new and different environment. Initially, I tried to mimic the Pacman agent design but soon realized that a lot of the component are just not necessary or do not fit for Taxi game environment, like the `getQvalue()`, `getPolicy()`, `getAction()` functions. The Taxi environment already has built-in solutions, like `env.step(action)`. Therefore, the project leans the essence of the Pacman agent design, that is, the object-oriented class hierarchy where there is first the Agent class, then the ValueEstimationAgent class, the ReinforcementLearningAgent class,

the QLearningAgent class, and finally the DQNAgent class. Based on this Object-Oriented design and the guidance of the huggingface deep reinforcement learning hands-on tutorial, this project adds its own constructors, fields, methods for the agent classes, including fields to instantiate and store the Taxi game environment and to represent Q value table space in NumPy array form for better performance.

Designing and Implementing DQN Agent

The DQN Agent design is heavily influenced by the PyTorch Reinforcement Learning (DQN) Tutorial. This is partly because I only have limited knowledge and understanding on how the neural networks works exactly and partly because I have no previous experience in building a neural network in PyTorch from scratch. After some struggling, I realized that the training and optimization functions from the PyTorch tutorial DQN cannot really run on the Taxi environment due to difference in PyTorch version and the Observation space. Unlike the ready-to-use DQN model from Stable Baselines which can be easily adjusted by fiddling the function parameters, the PyTorch DQN design requires manually delve into the nitty-gritty of the model which, unfortunately, is something I am not confident enough to get into at the moment. Eventually, I discovered another DQN tutorial that runs on a Discrete action and observation space of the Frozen Lake, a game environment similar yet different from the Taxi game. Following the tutorial video on YouTube Deep Q-Learning/Deep Q-Network (DQN) Explained | Python PyTorch Deep Reinforcement Learning and the demo code on GitHub, I gradually walk through and modify the DQN PyTorch model to fit for the requirement of the Taxi environment. Though, it was a massive challenge to figure out which part of the code is corresponding to which part of the DQN algorithm, like how to store the Replay Memory, when to batch the experience, how to train on the batch data in the 3 layers neural network, and when to use the optimizer properly.

Result: show the result and interpretation of your experiment. Any iterative improvements summary.

The following is the code implemented for Q-learning agent, DQN agent, and the performance comparison.

```
In [16]: !pip install -U pip
         !pip install -U torch
         !pip install -U torchvision
         !pip install gymnasium
         !pip install matplotlib
         !pip install imageio
```

```
Requirement already satisfied: pip in /opt/anaconda3/lib/python3.12/site-p
ackages (24.2)
Requirement already satisfied: torch in /opt/anaconda3/lib/python3.12/site
-packages (2.4.0)
Requirement already satisfied: filelock in /opt/anaconda3/lib/python3.12/s
ite-packages (from torch) (3.13.1)
```

Requirement already satisfied: typing-extensions>=4.8.0 in /opt/anaconda3/
lib/python3.12/site-packages (from torch) (4.11.0)
Requirement already satisfied: sympy in /opt/anaconda3/lib/python3.12/site
-packages (from torch) (1.12)
Requirement already satisfied: networkx in /opt/anaconda3/lib/python3.12/s
ite-packages (from torch) (3.2.1)
Requirement already satisfied: jinja2 in /opt/anaconda3/lib/python3.12/sit
e-packages (from torch) (3.1.4)
Requirement already satisfied: fsspec in /opt/anaconda3/lib/python3.12/sit
e-packages (from torch) (2024.3.1)
Requirement already satisfied: setuptools in /opt/anaconda3/lib/python3.1
2/site-packages (from torch) (69.5.1)
Requirement already satisfied: MarkupSafe>=2.0 in /opt/anaconda3/lib/pytho
n3.12/site-packages (from jinja2->torch) (2.1.3)
Requirement already satisfied: mpmath>=0.19 in /opt/anaconda3/lib/python3.
12/site-packages (from sympy->torch) (1.3.0)
Requirement already satisfied: torchvision in /opt/anaconda3/lib/python3.1
2/site-packages (0.19.0)
Requirement already satisfied: numpy in /opt/anaconda3/lib/python3.12/site
-packages (from torchvision) (1.26.4)
Requirement already satisfied: torch==2.4.0 in /opt/anaconda3/lib/python3.
12/site-packages (from torchvision) (2.4.0)
Requirement already satisfied: pillow!=8.3.*,>=5.3.0 in /opt/anaconda3/li
b/python3.12/site-packages (from torchvision) (10.3.0)
Requirement already satisfied: filelock in /opt/anaconda3/lib/python3.12/s
ite-packages (from torch==2.4.0->torchvision) (3.13.1)
Requirement already satisfied: typing-extensions>=4.8.0 in /opt/anaconda3/
lib/python3.12/site-packages (from torch==2.4.0->torchvision) (4.11.0)
Requirement already satisfied: sympy in /opt/anaconda3/lib/python3.12/site
-packages (from torch==2.4.0->torchvision) (1.12)
Requirement already satisfied: networkx in /opt/anaconda3/lib/python3.12/s
ite-packages (from torch==2.4.0->torchvision) (3.2.1)
Requirement already satisfied: jinja2 in /opt/anaconda3/lib/python3.12/sit
e-packages (from torch==2.4.0->torchvision) (3.1.4)
Requirement already satisfied: fsspec in /opt/anaconda3/lib/python3.12/sit
e-packages (from torch==2.4.0->torchvision) (2024.3.1)
Requirement already satisfied: setuptools in /opt/anaconda3/lib/python3.1
2/site-packages (from torch==2.4.0->torchvision) (69.5.1)
Requirement already satisfied: MarkupSafe>=2.0 in /opt/anaconda3/lib/pytho
n3.12/site-packages (from jinja2->torch==2.4.0->torchvision) (2.1.3)
Requirement already satisfied: mpmath>=0.19 in /opt/anaconda3/lib/python3.
12/site-packages (from sympy->torch==2.4.0->torchvision) (1.3.0)
Requirement already satisfied: gymnasium in /opt/anaconda3/lib/python3.12/
site-packages (1.0.0a2)
Requirement already satisfied: numpy>=1.21.0 in /opt/anaconda3/lib/python
3.12/site-packages (from gymnasium) (1.26.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /opt/anaconda3/lib/py
thon3.12/site-packages (from gymnasium) (2.2.1)
Requirement already satisfied: typing-extensions>=4.3.0 in /opt/anaconda3/
lib/python3.12/site-packages (from gymnasium) (4.11.0)
Requirement already satisfied: farama-notifications>=0.0.1 in /opt/anacond
a3/lib/python3.12/site-packages (from gymnasium) (0.0.4)
Requirement already satisfied: matplotlib in /opt/anaconda3/lib/python3.1
2/site-packages (3.8.4)
Requirement already satisfied: contourpy>=1.0.1 in /opt/anaconda3/lib/pyth
on3.12/site-packages (from matplotlib) (1.2.0)
Requirement already satisfied: cycler>=0.10 in /opt/anaconda3/lib/python3.

In [17]:
```python
import gymnasium as gym
import ale_py
# if using gymnasium
import shimmy

# print(gym.envs.registry.keys())
gym.envs.registration.registry.keys()
```

Out[17]:
```
dict_keys(['CartPole-v0', 'CartPole-v1', 'MountainCar-v0', 'MountainCarC
ontinuous-v0', 'Pendulum-v1', 'Acrobot-v1', 'phys2d/CartPole-v0', 'phys2
d/CartPole-v1', 'phys2d/Pendulum-v0', 'LunarLander-v3', 'LunarLanderCont
inuous-v3', 'BipedalWalker-v3', 'BipedalWalkerHardcore-v3', 'CarRacing-v
2', 'Blackjack-v1', 'FrozenLake-v1', 'FrozenLake8x8-v1', 'CliffWalking-v
0', 'Taxi-v3', 'tabular/Blackjack-v0', 'tabular/CliffWalking-v0', 'Reach
er-v2', 'Reacher-v4', 'Reacher-v5', 'Pusher-v2', 'Pusher-v4', 'Pusher-v
5', 'InvertedPendulum-v2', 'InvertedPendulum-v4', 'InvertedPendulum-v5',
'InvertedDoublePendulum-v2', 'InvertedDoublePendulum-v4', 'InvertedDoubl
ePendulum-v5', 'HalfCheetah-v2', 'HalfCheetah-v3', 'HalfCheetah-v4', 'Ha
lfCheetah-v5', 'Hopper-v2', 'Hopper-v3', 'Hopper-v4', 'Hopper-v5', 'Swim
mer-v2', 'Swimmer-v3', 'Swimmer-v4', 'Swimmer-v5', 'Walker2d-v2', 'Walke
r2d-v3', 'Walker2d-v4', 'Walker2d-v5', 'Ant-v2', 'Ant-v3', 'Ant-v4', 'An
t-v5', 'Humanoid-v2', 'Humanoid-v3', 'Humanoid-v4', 'Humanoid-v5', 'Huma
noidStandup-v2', 'HumanoidStandup-v4', 'HumanoidStandup-v5', 'Adventure-
v0', 'AdventureDeterministic-v0', 'AdventureNoFrameskip-v0', 'Adventure-
v4', 'AdventureDeterministic-v4', 'AdventureNoFrameskip-v4', 'Adventure-
ram-v0', 'Adventure-ramDeterministic-v0', 'Adventure-ramNoFrameskip-v0',
'Adventure-ram-v4', 'Adventure-ramDeterministic-v4', 'Adventure-ramNoFra
meskip-v4', 'AirRaid-v0', 'AirRaidDeterministic-v0', 'AirRaidNoFrameskip
-v0', 'AirRaid-v4', 'AirRaidDeterministic-v4', 'AirRaidNoFrameskip-v4',
'AirRaid-ram-v0', 'AirRaid-ramDeterministic-v0', 'AirRaid-ramNoFrameskip
-v0', 'AirRaid-ram-v4', 'AirRaid-ramDeterministic-v4', 'AirRaid-ramNoFra
meskip-v4', 'Alien-v0', 'AlienDeterministic-v0', 'AlienNoFrameskip-v0',
'Alien-v4', 'AlienDeterministic-v4', 'AlienNoFrameskip-v4', 'Alien-ram-v
```

0', 'Alien-ramDeterministic-v0', 'Alien-ramNoFrameskip-v0', 'Alien-ram-v4', 'Alien-ramDeterministic-v4', 'Alien-ramNoFrameskip-v4', 'Amidar-v0', 'AmidarDeterministic-v0', 'AmidarNoFrameskip-v0', 'Amidar-v4', 'AmidarDeterministic-v4', 'AmidarNoFrameskip-v4', 'Amidar-ram-v0', 'Amidar-ramDeterministic-v0', 'Amidar-ramNoFrameskip-v0', 'Amidar-ram-v4', 'Amidar-ramDeterministic-v4', 'Amidar-ramNoFrameskip-v4', 'Assault-v0', 'AssaultDeterministic-v0', 'AssaultNoFrameskip-v0', 'Assault-v4', 'AssaultDeterministic-v4', 'AssaultNoFrameskip-v4', 'Assault-ram-v0', 'Assault-ramDeterministic-v0', 'Assault-ramNoFrameskip-v0', 'Assault-ram-v4', 'Assault-ramDeterministic-v4', 'Assault-ramNoFrameskip-v4', 'Asterix-v0', 'AsterixDeterministic-v0', 'AsterixNoFrameskip-v0', 'Asterix-v4', 'AsterixDeterministic-v4', 'AsterixNoFrameskip-v4', 'Asterix-ram-v0', 'Asterix-ramDeterministic-v0', 'Asterix-ramNoFrameskip-v0', 'Asterix-ram-v4', 'Asterix-ramDeterministic-v4', 'Asterix-ramNoFrameskip-v4', 'Asteroids-v0', 'AsteroidsDeterministic-v0', 'AsteroidsNoFrameskip-v0', 'Asteroids-v4', 'AsteroidsDeterministic-v4', 'AsteroidsNoFrameskip-v4', 'Asteroids-ram-v0', 'Asteroids-ramDeterministic-v0', 'Asteroids-ramNoFrameskip-v0', 'Asteroids-ram-v4', 'Asteroids-ramDeterministic-v4', 'Asteroids-ramNoFrameskip-v4', 'Atlantis-v0', 'AtlantisDeterministic-v0', 'AtlantisNoFrameskip-v0', 'Atlantis-v4', 'AtlantisDeterministic-v4', 'AtlantisNoFrameskip-v4', 'Atlantis-ram-v0', 'Atlantis-ramDeterministic-v0', 'Atlantis-ramNoFrameskip-v0', 'Atlantis-ram-v4', 'Atlantis-ramDeterministic-v4', 'Atlantis-ramNoFrameskip-v4', 'BankHeist-v0', 'BankHeistDeterministic-v0', 'BankHeistNoFrameskip-v0', 'BankHeist-v4', 'BankHeistDeterministic-v4', 'BankHeistNoFrameskip-v4', 'BankHeist-ram-v0', 'BankHeist-ramDeterministic-v0', 'BankHeist-ramNoFrameskip-v0', 'BankHeist-ram-v4', 'BankHeist-ramDeterministic-v4', 'BankHeist-ramNoFrameskip-v4', 'BattleZone-v0', 'BattleZoneDeterministic-v0', 'BattleZoneNoFrameskip-v0', 'BattleZone-v4', 'BattleZoneDeterministic-v4', 'BattleZoneNoFrameskip-v4', 'BattleZone-ram-v0', 'BattleZone-ramDeterministic-v0', 'BattleZone-ramNoFrameskip-v0', 'BattleZone-ram-v4', 'BattleZone-ramDeterministic-v4', 'BattleZone-ramNoFrameskip-v4', 'BeamRider-v0', 'BeamRiderDeterministic-v0', 'BeamRiderNoFrameskip-v0', 'BeamRider-v4', 'BeamRiderDeterministic-v4', 'BeamRiderNoFrameskip-v4', 'BeamRider-ram-v0', 'BeamRider-ramDeterministic-v0', 'BeamRider-ramNoFrameskip-v0', 'BeamRider-ram-v4', 'BeamRider-ramDeterministic-v4', 'BeamRider-ramNoFrameskip-v4', 'Berzerk-v0', 'BerzerkDeterministic-v0', 'BerzerkNoFrameskip-v0', 'Berzerk-v4', 'BerzerkDeterministic-v4', 'BerzerkNoFrameskip-v4', 'Berzerk-ram-v0', 'Berzerk-ramDeterministic-v0', 'Berzerk-ramNoFrameskip-v0', 'Berzerk-ram-v4', 'Berzerk-ramDeterministic-v4', 'Berzerk-ramNoFrameskip-v4', 'Bowling-v0', 'BowlingDeterministic-v0', 'BowlingNoFrameskip-v0', 'Bowling-v4', 'BowlingDeterministic-v4', 'BowlingNoFrameskip-v4', 'Bowling-ram-v0', 'Bowling-ramDeterministic-v0', 'Bowling-ramNoFrameskip-v0', 'Bowling-ram-v4', 'Bowling-ramDeterministic-v4', 'Bowling-ramNoFrameskip-v4', 'Boxing-v0', 'BoxingDeterministic-v0', 'BoxingNoFrameskip-v0', 'Boxing-v4', 'BoxingDeterministic-v4', 'BoxingNoFrameskip-v4', 'Boxing-ram-v0', 'Boxing-ramDeterministic-v0', 'Boxing-ramNoFrameskip-v0', 'Boxing-ram-v4', 'Boxing-ramDeterministic-v4', 'Boxing-ramNoFrameskip-v4', 'Breakout-v0', 'BreakoutDeterministic-v0', 'BreakoutNoFrameskip-v0', 'Breakout-v4', 'BreakoutDeterministic-v4', 'BreakoutNoFrameskip-v4', 'Breakout-ram-v0', 'Breakout-ramDeterministic-v0', 'Breakout-ramNoFrameskip-v0', 'Breakout-ram-v4', 'Breakout-ramDeterministic-v4', 'Breakout-ramNoFrameskip-v4', 'Carnival-v0', 'CarnivalDeterministic-v0', 'CarnivalNoFrameskip-v0', 'Carnival-v4', 'CarnivalDeterministic-v4', 'CarnivalNoFrameskip-v4', 'Carnival-ram-v0', 'Carnival-ramDeterministic-v0', 'Carnival-ramNoFrameskip-v0', 'Carnival-ram-v4', 'Carnival-ramDeterministic-v4', 'Carnival-ramNoFrameskip-v4', 'Centipede-v0', 'CentipedeDeterministic-v0', 'CentipedeNoFrameskip-v0', 'Centipede-v4', 'CentipedeDeterminis

tic-v4', 'CentipedeNoFrameskip-v4', 'Centipede-ram-v0', 'Centipede-ramDeterministic-v0', 'Centipede-ramNoFrameskip-v0', 'Centipede-ram-v4', 'Centipede-ramDeterministic-v4', 'Centipede-ramNoFrameskip-v4', 'ChopperCommand-v0', 'ChopperCommandDeterministic-v0', 'ChopperCommandNoFrameskip-v0', 'ChopperCommand-v4', 'ChopperCommandDeterministic-v4', 'ChopperCommandNoFrameskip-v4', 'ChopperCommand-ram-v0', 'ChopperCommand-ramDeterministic-v0', 'ChopperCommand-ramNoFrameskip-v0', 'ChopperCommand-ram-v4', 'ChopperCommand-ramDeterministic-v4', 'ChopperCommand-ramNoFrameskip-v4', 'CrazyClimber-v0', 'CrazyClimberDeterministic-v0', 'CrazyClimberNoFrameskip-v0', 'CrazyClimber-v4', 'CrazyClimberDeterministic-v4', 'CrazyClimberNoFrameskip-v4', 'CrazyClimber-ram-v0', 'CrazyClimber-ramDeterministic-v0', 'CrazyClimber-ramNoFrameskip-v0', 'CrazyClimber-ram-v4', 'CrazyClimber-ramDeterministic-v4', 'CrazyClimber-ramNoFrameskip-v4', 'Defender-v0', 'DefenderDeterministic-v0', 'DefenderNoFrameskip-v0', 'Defender-v4', 'DefenderDeterministic-v4', 'DefenderNoFrameskip-v4', 'Defender-ram-v0', 'Defender-ramDeterministic-v0', 'Defender-ramNoFrameskip-v0', 'Defender-ram-v4', 'Defender-ramDeterministic-v4', 'Defender-ramNoFrameskip-v4', 'DemonAttack-v0', 'DemonAttackDeterministic-v0', 'DemonAttackNoFrameskip-v0', 'DemonAttack-v4', 'DemonAttackDeterministic-v4', 'DemonAttackNoFrameskip-v4', 'DemonAttack-ram-v0', 'DemonAttack-ramDeterministic-v0', 'DemonAttack-ramNoFrameskip-v0', 'DemonAttack-ram-v4', 'DemonAttack-ramDeterministic-v4', 'DemonAttack-ramNoFrameskip-v4', 'DoubleDunk-v0', 'DoubleDunkDeterministic-v0', 'DoubleDunkNoFrameskip-v0', 'DoubleDunk-v4', 'DoubleDunkDeterministic-v4', 'DoubleDunkNoFrameskip-v4', 'DoubleDunk-ram-v0', 'DoubleDunk-ramDeterministic-v0', 'DoubleDunk-ramNoFrameskip-v0', 'DoubleDunk-ram-v4', 'DoubleDunk-ramDeterministic-v4', 'DoubleDunk-ramNoFrameskip-v4', 'ElevatorAction-v0', 'ElevatorActionDeterministic-v0', 'ElevatorActionNoFrameskip-v0', 'ElevatorAction-v4', 'ElevatorActionDeterministic-v4', 'ElevatorActionNoFrameskip-v4', 'ElevatorAction-ram-v0', 'ElevatorAction-ramDeterministic-v0', 'ElevatorAction-ramNoFrameskip-v0', 'ElevatorAction-ram-v4', 'ElevatorAction-ramDeterministic-v4', 'ElevatorAction-ramNoFrameskip-v4', 'Enduro-v0', 'EnduroDeterministic-v0', 'EnduroNoFrameskip-v0', 'Enduro-v4', 'EnduroDeterministic-v4', 'EnduroNoFrameskip-v4', 'Enduro-ram-v0', 'Enduro-ramDeterministic-v0', 'Enduro-ramNoFrameskip-v0', 'Enduro-ram-v4', 'Enduro-ramDeterministic-v4', 'Enduro-ramNoFrameskip-v4', 'FishingDerby-v0', 'FishingDerbyDeterministic-v0', 'FishingDerbyNoFrameskip-v0', 'FishingDerby-v4', 'FishingDerbyDeterministic-v4', 'FishingDerbyNoFrameskip-v4', 'FishingDerby-ram-v0', 'FishingDerby-ramDeterministic-v0', 'FishingDerby-ramNoFrameskip-v0', 'FishingDerby-ram-v4', 'FishingDerby-ramDeterministic-v4', 'FishingDerby-ramNoFrameskip-v4', 'Freeway-v0', 'FreewayDeterministic-v0', 'FreewayNoFrameskip-v0', 'Freeway-v4', 'FreewayDeterministic-v4', 'FreewayNoFrameskip-v4', 'Freeway-ram-v0', 'Freeway-ramDeterministic-v0', 'Freeway-ramNoFrameskip-v0', 'Freeway-ram-v4', 'Freeway-ramDeterministic-v4', 'Freeway-ramNoFrameskip-v4', 'Frostbite-v0', 'FrostbiteDeterministic-v0', 'FrostbiteNoFrameskip-v0', 'Frostbite-v4', 'FrostbiteDeterministic-v4', 'FrostbiteNoFrameskip-v4', 'Frostbite-ram-v0', 'Frostbite-ramDeterministic-v0', 'Frostbite-ramNoFrameskip-v0', 'Frostbite-ram-v4', 'Frostbite-ramDeterministic-v4', 'Frostbite-ramNoFrameskip-v4', 'Gopher-v0', 'GopherDeterministic-v0', 'GopherNoFrameskip-v0', 'Gopher-v4', 'GopherDeterministic-v4', 'GopherNoFrameskip-v4', 'Gopher-ram-v0', 'Gopher-ramDeterministic-v0', 'Gopher-ramNoFrameskip-v0', 'Gopher-ram-v4', 'Gopher-ramDeterministic-v4', 'Gopher-ramNoFrameskip-v4', 'Gravitar-v0', 'GravitarDeterministic-v0', 'GravitarNoFrameskip-v0', 'Gravitar-v4', 'GravitarDeterministic-v4', 'GravitarNoFrameskip-v4', 'Gravitar-ram-v0', 'Gravitar-ramDeterministic-v0', 'Gravitar-ramNoFrameskip-v0', 'Gravitar-ram-v4', 'Gravitar-ramDeterministic-v4', 'Gravitar-ramNoFrameskip-v4', 'Hero-v0', 'HeroDeterministic-v0', 'HeroNoFr

ameskip-v0', 'Hero-v4', 'HeroDeterministic-v4', 'HeroNoFrameskip-v4', 'Hero-ram-v0', 'Hero-ramDeterministic-v0', 'Hero-ramNoFrameskip-v0', 'Hero-ram-v4', 'Hero-ramDeterministic-v4', 'Hero-ramNoFrameskip-v4', 'IceHockey-v0', 'IceHockeyDeterministic-v0', 'IceHockeyNoFrameskip-v0', 'IceHockey-v4', 'IceHockeyDeterministic-v4', 'IceHockeyNoFrameskip-v4', 'IceHockey-ram-v0', 'IceHockey-ramDeterministic-v0', 'IceHockey-ramNoFrameskip-v0', 'IceHockey-ram-v4', 'IceHockey-ramDeterministic-v4', 'IceHockey-ramNoFrameskip-v4', 'Jamesbond-v0', 'JamesbondDeterministic-v0', 'JamesbondNoFrameskip-v0', 'Jamesbond-v4', 'JamesbondDeterministic-v4', 'JamesbondNoFrameskip-v4', 'Jamesbond-ram-v0', 'Jamesbond-ramDeterministic-v0', 'Jamesbond-ramNoFrameskip-v0', 'Jamesbond-ram-v4', 'Jamesbond-ramDeterministic-v4', 'Jamesbond-ramNoFrameskip-v4', 'JourneyEscape-v0', 'JourneyEscapeDeterministic-v0', 'JourneyEscapeNoFrameskip-v0', 'JourneyEscape-v4', 'JourneyEscapeDeterministic-v4', 'JourneyEscapeNoFrameskip-v4', 'JourneyEscape-ram-v0', 'JourneyEscape-ramDeterministic-v0', 'JourneyEscape-ramNoFrameskip-v0', 'JourneyEscape-ram-v4', 'JourneyEscape-ramDeterministic-v4', 'JourneyEscape-ramNoFrameskip-v4', 'Kangaroo-v0', 'KangarooDeterministic-v0', 'KangarooNoFrameskip-v0', 'Kangaroo-v4', 'KangarooDeterministic-v4', 'KangarooNoFrameskip-v4', 'Kangaroo-ram-v0', 'Kangaroo-ramDeterministic-v0', 'Kangaroo-ramNoFrameskip-v0', 'Kangaroo-ram-v4', 'Kangaroo-ramDeterministic-v4', 'Kangaroo-ramNoFrameskip-v4', 'Krull-v0', 'KrullDeterministic-v0', 'KrullNoFrameskip-v0', 'Krull-v4', 'KrullDeterministic-v4', 'KrullNoFrameskip-v4', 'Krull-ram-v0', 'Krull-ramDeterministic-v0', 'Krull-ramNoFrameskip-v0', 'Krull-ram-v4', 'Krull-ramDeterministic-v4', 'Krull-ramNoFrameskip-v4', 'KungFuMaster-v0', 'KungFuMasterDeterministic-v0', 'KungFuMasterNoFrameskip-v0', 'KungFuMaster-v4', 'KungFuMasterDeterministic-v4', 'KungFuMasterNoFrameskip-v4', 'KungFuMaster-ram-v0', 'KungFuMaster-ramDeterministic-v0', 'KungFuMaster-ramNoFrameskip-v0', 'KungFuMaster-ram-v4', 'KungFuMaster-ramDeterministic-v4', 'KungFuMaster-ramNoFrameskip-v4', 'MontezumaRevenge-v0', 'MontezumaRevengeDeterministic-v0', 'MontezumaRevengeNoFrameskip-v0', 'MontezumaRevenge-v4', 'MontezumaRevengeDeterministic-v4', 'MontezumaRevengeNoFrameskip-v4', 'MontezumaRevenge-ram-v0', 'MontezumaRevenge-ramDeterministic-v0', 'MontezumaRevenge-ramNoFrameskip-v0', 'MontezumaRevenge-ram-v4', 'MontezumaRevenge-ramDeterministic-v4', 'MontezumaRevenge-ramNoFrameskip-v4', 'MsPacman-v0', 'MsPacmanDeterministic-v0', 'MsPacmanNoFrameskip-v0', 'MsPacman-v4', 'MsPacmanDeterministic-v4', 'MsPacmanNoFrameskip-v4', 'MsPacman-ram-v0', 'MsPacman-ramDeterministic-v0', 'MsPacman-ramNoFrameskip-v0', 'MsPacman-ram-v4', 'MsPacman-ramDeterministic-v4', 'MsPacman-ramNoFrameskip-v4', 'NameThisGame-v0', 'NameThisGameDeterministic-v0', 'NameThisGameNoFrameskip-v0', 'NameThisGame-v4', 'NameThisGameDeterministic-v4', 'NameThisGameNoFrameskip-v4', 'NameThisGame-ram-v0', 'NameThisGame-ramDeterministic-v0', 'NameThisGame-ramNoFrameskip-v0', 'NameThisGame-ram-v4', 'NameThisGame-ramDeterministic-v4', 'NameThisGame-ramNoFrameskip-v4', 'Phoenix-v0', 'PhoenixDeterministic-v0', 'PhoenixNoFrameskip-v0', 'Phoenix-v4', 'PhoenixDeterministic-v4', 'PhoenixNoFrameskip-v4', 'Phoenix-ram-v0', 'Phoenix-ramDeterministic-v0', 'Phoenix-ramNoFrameskip-v0', 'Phoenix-ram-v4', 'Phoenix-ramDeterministic-v4', 'Phoenix-ramNoFrameskip-v4', 'Pitfall-v0', 'PitfallDeterministic-v0', 'PitfallNoFrameskip-v0', 'Pitfall-v4', 'PitfallDeterministic-v4', 'PitfallNoFrameskip-v4', 'Pitfall-ram-v0', 'Pitfall-ramDeterministic-v0', 'Pitfall-ramNoFrameskip-v0', 'Pitfall-ram-v4', 'Pitfall-ramDeterministic-v4', 'Pitfall-ramNoFrameskip-v4', 'Pong-v0', 'PongDeterministic-v0', 'PongNoFrameskip-v0', 'Pong-v4', 'PongDeterministic-v4', 'PongNoFrameskip-v4', 'Pong-ram-v0', 'Pong-ramDeterministic-v0', 'Pong-ramNoFrameskip-v0', 'Pong-ram-v4', 'Pong-ramDeterministic-v4', 'Pong-ramNoFrameskip-v4', 'Pooyan-v0', 'PooyanDeterministic-v0', 'PooyanNoFrameskip-v0', 'Pooyan-v4', 'PooyanDeterministic-v4', 'PooyanNoFrameskip-v

4', 'Pooyan-ram-v0', 'Pooyan-ramDeterministic-v0', 'Pooyan-ramNoFrameski
p-v0', 'Pooyan-ram-v4', 'Pooyan-ramDeterministic-v4', 'Pooyan-ramNoFrame
skip-v4', 'PrivateEye-v0', 'PrivateEyeDeterministic-v0', 'PrivateEyeNoFr
ameskip-v0', 'PrivateEye-v4', 'PrivateEyeDeterministic-v4', 'PrivateEyeN
oFrameskip-v4', 'PrivateEye-ram-v0', 'PrivateEye-ramDeterministic-v0', '
PrivateEye-ramNoFrameskip-v0', 'PrivateEye-ram-v4', 'PrivateEye-ramDeter
ministic-v4', 'PrivateEye-ramNoFrameskip-v4', 'Qbert-v0', 'QbertDetermin
istic-v0', 'QbertNoFrameskip-v0', 'Qbert-v4', 'QbertDeterministic-v4', '
QbertNoFrameskip-v4', 'Qbert-ram-v0', 'Qbert-ramDeterministic-v0', 'Qber
t-ramNoFrameskip-v0', 'Qbert-ram-v4', 'Qbert-ramDeterministic-v4', 'Qber
t-ramNoFrameskip-v4', 'Riverraid-v0', 'RiverraidDeterministic-v0', 'Rive
rraidNoFrameskip-v0', 'Riverraid-v4', 'RiverraidDeterministic-v4', 'Rive
rraidNoFrameskip-v4', 'Riverraid-ram-v0', 'Riverraid-ramDeterministic-v
0', 'Riverraid-ramNoFrameskip-v0', 'Riverraid-ram-v4', 'Riverraid-ramDet
erministic-v4', 'Riverraid-ramNoFrameskip-v4', 'RoadRunner-v0', 'RoadRun
nerDeterministic-v0', 'RoadRunnerNoFrameskip-v0', 'RoadRunner-v4', 'Road
RunnerDeterministic-v4', 'RoadRunnerNoFrameskip-v4', 'RoadRunner-ram-v
0', 'RoadRunner-ramDeterministic-v0', 'RoadRunner-ramNoFrameskip-v0', 'R
oadRunner-ram-v4', 'RoadRunner-ramDeterministic-v4', 'RoadRunner-ramNoFr
ameskip-v4', 'Robotank-v0', 'RobotankDeterministic-v0', 'RobotankNoFrame
skip-v0', 'Robotank-v4', 'RobotankDeterministic-v4', 'RobotankNoFrameski
p-v4', 'Robotank-ram-v0', 'Robotank-ramDeterministic-v0', 'Robotank-ramN
oFrameskip-v0', 'Robotank-ram-v4', 'Robotank-ramDeterministic-v4', 'Robo
tank-ramNoFrameskip-v4', 'Seaquest-v0', 'SeaquestDeterministic-v0', 'Sea
questNoFrameskip-v0', 'Seaquest-v4', 'SeaquestDeterministic-v4', 'Seaque
stNoFrameskip-v4', 'Seaquest-ram-v0', 'Seaquest-ramDeterministic-v0', 'S
eaquest-ramNoFrameskip-v0', 'Seaquest-ram-v4', 'Seaquest-ramDeterministi
c-v4', 'Seaquest-ramNoFrameskip-v4', 'Skiing-v0', 'SkiingDeterministic-v
0', 'SkiingNoFrameskip-v0', 'Skiing-v4', 'SkiingDeterministic-v4', 'Skii
ngNoFrameskip-v4', 'Skiing-ram-v0', 'Skiing-ramDeterministic-v0', 'Skiin
g-ramNoFrameskip-v0', 'Skiing-ram-v4', 'Skiing-ramDeterministic-v4', 'Sk
iing-ramNoFrameskip-v4', 'Solaris-v0', 'SolarisDeterministic-v0', 'Solar
isNoFrameskip-v0', 'Solaris-v4', 'SolarisDeterministic-v4', 'SolarisNoFr
ameskip-v4', 'Solaris-ram-v0', 'Solaris-ramDeterministic-v0', 'Solaris-r
amNoFrameskip-v0', 'Solaris-ram-v4', 'Solaris-ramDeterministic-v4', 'Sol
aris-ramNoFrameskip-v4', 'SpaceInvaders-v0', 'SpaceInvadersDeterministic
-v0', 'SpaceInvadersNoFrameskip-v0', 'SpaceInvaders-v4', 'SpaceInvadersD
eterministic-v4', 'SpaceInvadersNoFrameskip-v4', 'SpaceInvaders-ram-v0',
'SpaceInvaders-ramDeterministic-v0', 'SpaceInvaders-ramNoFrameskip-v0',
'SpaceInvaders-ram-v4', 'SpaceInvaders-ramDeterministic-v4', 'SpaceInvad
ers-ramNoFrameskip-v4', 'StarGunner-v0', 'StarGunnerDeterministic-v0', '
StarGunnerNoFrameskip-v0', 'StarGunner-v4', 'StarGunnerDeterministic-v
4', 'StarGunnerNoFrameskip-v4', 'StarGunner-ram-v0', 'StarGunner-ramDete
rministic-v0', 'StarGunner-ramNoFrameskip-v0', 'StarGunner-ram-v4', 'Sta
rGunner-ramDeterministic-v4', 'StarGunner-ramNoFrameskip-v4', 'Tennis-v
0', 'TennisDeterministic-v0', 'TennisNoFrameskip-v0', 'Tennis-v4', 'Tenn
isDeterministic-v4', 'TennisNoFrameskip-v4', 'Tennis-ram-v0', 'Tennis-ra
mDeterministic-v0', 'Tennis-ramNoFrameskip-v0', 'Tennis-ram-v4', 'Tennis
-ramDeterministic-v4', 'Tennis-ramNoFrameskip-v4', 'TimePilot-v0', 'Time
PilotDeterministic-v0', 'TimePilotNoFrameskip-v0', 'TimePilot-v4', 'Time
PilotDeterministic-v4', 'TimePilotNoFrameskip-v4', 'TimePilot-ram-v0', '
TimePilot-ramDeterministic-v0', 'TimePilot-ramNoFrameskip-v0', 'TimePilo
t-ram-v4', 'TimePilot-ramDeterministic-v4', 'TimePilot-ramNoFrameskip-v
4', 'Tutankham-v0', 'TutankhamDeterministic-v0', 'TutankhamNoFrameskip-v
0', 'Tutankham-v4', 'TutankhamDeterministic-v4', 'TutankhamNoFrameskip-v
4', 'Tutankham-ram-v0', 'Tutankham-ramDeterministic-v0', 'Tutankham-ramN
oFrameskip-v0', 'Tutankham-ram-v4', 'Tutankham-ramDeterministic-v4', 'Tu

tankham-ramNoFrameskip-v4', 'UpNDown-v0', 'UpNDownDeterministic-v0', 'UpNDownNoFrameskip-v0', 'UpNDown-v4', 'UpNDownDeterministic-v4', 'UpNDownNoFrameskip-v4', 'UpNDown-ram-v0', 'UpNDown-ramDeterministic-v0', 'UpNDown-ramNoFrameskip-v0', 'UpNDown-ram-v4', 'UpNDown-ramDeterministic-v4', 'UpNDown-ramNoFrameskip-v4', 'Venture-v0', 'VentureDeterministic-v0', 'VentureNoFrameskip-v0', 'Venture-v4', 'VentureDeterministic-v4', 'VentureNoFrameskip-v4', 'Venture-ram-v0', 'Venture-ramDeterministic-v0', 'Venture-ramNoFrameskip-v0', 'Venture-ram-v4', 'Venture-ramDeterministic-v4', 'Venture-ramNoFrameskip-v4', 'VideoPinball-v0', 'VideoPinballDeterministic-v0', 'VideoPinballNoFrameskip-v0', 'VideoPinball-v4', 'VideoPinballDeterministic-v4', 'VideoPinballNoFrameskip-v4', 'VideoPinball-ram-v0', 'VideoPinball-ramDeterministic-v0', 'VideoPinball-ramNoFrameskip-v0', 'VideoPinball-ram-v4', 'VideoPinball-ramDeterministic-v4', 'VideoPinball-ramNoFrameskip-v4', 'WizardOfWor-v0', 'WizardOfWorDeterministic-v0', 'WizardOfWorNoFrameskip-v0', 'WizardOfWor-v4', 'WizardOfWorDeterministic-v4', 'WizardOfWorNoFrameskip-v4', 'WizardOfWor-ram-v0', 'WizardOfWor-ramDeterministic-v0', 'WizardOfWor-ramNoFrameskip-v0', 'WizardOfWor-ram-v4', 'WizardOfWor-ramDeterministic-v4', 'WizardOfWor-ramNoFrameskip-v4', 'YarsRevenge-v0', 'YarsRevengeDeterministic-v0', 'YarsRevengeNoFrameskip-v0', 'YarsRevenge-v4', 'YarsRevengeDeterministic-v4', 'YarsRevengeNoFrameskip-v4', 'YarsRevenge-ram-v0', 'YarsRevenge-ramDeterministic-v0', 'YarsRevenge-ramNoFrameskip-v0', 'YarsRevenge-ram-v4', 'YarsRevenge-ramDeterministic-v4', 'YarsRevenge-ramNoFrameskip-v4', 'Zaxxon-v0', 'ZaxxonDeterministic-v0', 'ZaxxonNoFrameskip-v0', 'Zaxxon-v4', 'ZaxxonDeterministic-v4', 'ZaxxonNoFrameskip-v4', 'Zaxxon-ram-v0', 'Zaxxon-ramDeterministic-v0', 'Zaxxon-ramNoFrameskip-v0', 'Zaxxon-ram-v4', 'Zaxxon-ramDeterministic-v4', 'Zaxxon-ramNoFrameskip-v4', 'ALE/Adventure-v5', 'ALE/Adventure-ram-v5', 'ALE/AirRaid-v5', 'ALE/AirRaid-ram-v5', 'ALE/Alien-v5', 'ALE/Alien-ram-v5', 'ALE/Amidar-v5', 'ALE/Amidar-ram-v5', 'ALE/Assault-v5', 'ALE/Assault-ram-v5', 'ALE/Asterix-v5', 'ALE/Asterix-ram-v5', 'ALE/Asteroids-v5', 'ALE/Asteroids-ram-v5', 'ALE/Atlantis2-v5', 'ALE/Atlantis2-ram-v5', 'ALE/Atlantis-v5', 'ALE/Atlantis-ram-v5', 'ALE/Backgammon-v5', 'ALE/Backgammon-ram-v5', 'ALE/BankHeist-v5', 'ALE/BankHeist-ram-v5', 'ALE/BasicMath-v5', 'ALE/BasicMath-ram-v5', 'ALE/BattleZone-v5', 'ALE/BattleZone-ram-v5', 'ALE/BeamRider-v5', 'ALE/BeamRider-ram-v5', 'ALE/Berzerk-v5', 'ALE/Berzerk-ram-v5', 'ALE/Blackjack-v5', 'ALE/Blackjack-ram-v5', 'ALE/Bowling-v5', 'ALE/Bowling-ram-v5', 'ALE/Boxing-v5', 'ALE/Boxing-ram-v5', 'ALE/Breakout-v5', 'ALE/Breakout-ram-v5', 'ALE/Carnival-v5', 'ALE/Carnival-ram-v5', 'ALE/Casino-v5', 'ALE/Casino-ram-v5', 'ALE/Centipede-v5', 'ALE/Centipede-ram-v5', 'ALE/ChopperCommand-v5', 'ALE/ChopperCommand-ram-v5', 'ALE/Combat-v5', 'ALE/Combat-ram-v5', 'ALE/CrazyClimber-v5', 'ALE/CrazyClimber-ram-v5', 'ALE/Crossbow-v5', 'ALE/Crossbow-ram-v5', 'ALE/Darkchambers-v5', 'ALE/Darkchambers-ram-v5', 'ALE/Defender-v5', 'ALE/Defender-ram-v5', 'ALE/DemonAttack-v5', 'ALE/DemonAttack-ram-v5', 'ALE/DonkeyKong-v5', 'ALE/DonkeyKong-ram-v5', 'ALE/DoubleDunk-v5', 'ALE/DoubleDunk-ram-v5', 'ALE/Earthworld-v5', 'ALE/Earthworld-ram-v5', 'ALE/ElevatorAction-v5', 'ALE/ElevatorAction-ram-v5', 'ALE/Enduro-v5', 'ALE/Enduro-ram-v5', 'ALE/Entombed-v5', 'ALE/Entombed-ram-v5', 'ALE/Et-v5', 'ALE/Et-ram-v5', 'ALE/FishingDerby-v5', 'ALE/FishingDerby-ram-v5', 'ALE/FlagCapture-v5', 'ALE/FlagCapture-ram-v5', 'ALE/Freeway-v5', 'ALE/Freeway-ram-v5', 'ALE/Frogger-v5', 'ALE/Frogger-ram-v5', 'ALE/Frostbite-v5', 'ALE/Frostbite-ram-v5', 'ALE/Galaxian-v5', 'ALE/Galaxian-ram-v5', 'ALE/Gopher-v5', 'ALE/Gopher-ram-v5', 'ALE/Gravitar-v5', 'ALE/Gravitar-ram-v5', 'ALE/Hangman-v5', 'ALE/Hangman-ram-v5', 'ALE/HauntedHouse-v5', 'ALE/HauntedHouse-ram-v5', 'ALE/Hero-v5', 'ALE/Hero-ram-v5', 'ALE/HumanCannonball-v5', 'ALE/HumanCannonball-ram-v5', 'ALE/IceHockey-v5', 'ALE/IceHockey-ram-v5', 'ALE/Jamesbond-v5', 'ALE/Jamesbond-ram-v5', 'ALE/JourneyEscape-v5', 'ALE/JourneyEscape-

```
ram-v5', 'ALE/Joust-v5', 'ALE/Joust-ram-v5', 'ALE/Kaboom-v5', 'ALE/Kaboo
m-ram-v5', 'ALE/Kangaroo-v5', 'ALE/Kangaroo-ram-v5', 'ALE/KeystoneKapers
-v5', 'ALE/KeystoneKapers-ram-v5', 'ALE/KingKong-v5', 'ALE/KingKong-ram-
v5', 'ALE/Klax-v5', 'ALE/Klax-ram-v5', 'ALE/Koolaid-v5', 'ALE/Koolaid-ra
m-v5', 'ALE/Krull-v5', 'ALE/Krull-ram-v5', 'ALE/KungFuMaster-v5', 'ALE/K
ungFuMaster-ram-v5', 'ALE/LaserGates-v5', 'ALE/LaserGates-ram-v5', 'ALE/
LostLuggage-v5', 'ALE/LostLuggage-ram-v5', 'ALE/MarioBros-v5', 'ALE/Mari
oBros-ram-v5', 'ALE/MazeCraze-v5', 'ALE/MazeCraze-ram-v5', 'ALE/Miniatur
eGolf-v5', 'ALE/MiniatureGolf-ram-v5', 'ALE/MontezumaRevenge-v5', 'ALE/M
ontezumaRevenge-ram-v5', 'ALE/MrDo-v5', 'ALE/MrDo-ram-v5', 'ALE/MsPacman
-v5', 'ALE/MsPacman-ram-v5', 'ALE/NameThisGame-v5', 'ALE/NameThisGame-ra
m-v5', 'ALE/Othello-v5', 'ALE/Othello-ram-v5', 'ALE/Pacman-v5', 'ALE/Pac
man-ram-v5', 'ALE/Phoenix-v5', 'ALE/Phoenix-ram-v5', 'ALE/Pitfall2-v5',
'ALE/Pitfall2-ram-v5', 'ALE/Pitfall-v5', 'ALE/Pitfall-ram-v5', 'ALE/Pong
-v5', 'ALE/Pong-ram-v5', 'ALE/Pooyan-v5', 'ALE/Pooyan-ram-v5', 'ALE/Priv
ateEye-v5', 'ALE/PrivateEye-ram-v5', 'ALE/Qbert-v5', 'ALE/Qbert-ram-v5',
'ALE/Riverraid-v5', 'ALE/Riverraid-ram-v5', 'ALE/RoadRunner-v5', 'ALE/Ro
adRunner-ram-v5', 'ALE/Robotank-v5', 'ALE/Robotank-ram-v5', 'ALE/Seaques
t-v5', 'ALE/Seaquest-ram-v5', 'ALE/SirLancelot-v5', 'ALE/SirLancelot-ram
-v5', 'ALE/Skiing-v5', 'ALE/Skiing-ram-v5', 'ALE/Solaris-v5', 'ALE/Solar
is-ram-v5', 'ALE/SpaceInvaders-v5', 'ALE/SpaceInvaders-ram-v5', 'ALE/Spa
ceWar-v5', 'ALE/SpaceWar-ram-v5', 'ALE/StarGunner-v5', 'ALE/StarGunner-r
am-v5', 'ALE/Superman-v5', 'ALE/Superman-ram-v5', 'ALE/Surround-v5', 'AL
E/Surround-ram-v5', 'ALE/Tennis-v5', 'ALE/Tennis-ram-v5', 'ALE/Tetris-v
5', 'ALE/Tetris-ram-v5', 'ALE/TicTacToe3D-v5', 'ALE/TicTacToe3D-ram-v5',
'ALE/TimePilot-v5', 'ALE/TimePilot-ram-v5', 'ALE/Trondead-v5', 'ALE/Tron
dead-ram-v5', 'ALE/Turmoil-v5', 'ALE/Turmoil-ram-v5', 'ALE/Tutankham-v
5', 'ALE/Tutankham-ram-v5', 'ALE/UpNDown-v5', 'ALE/UpNDown-ram-v5', 'AL
E/Venture-v5', 'ALE/Venture-ram-v5', 'ALE/VideoCheckers-v5', 'ALE/VideoC
heckers-ram-v5', 'ALE/VideoChess-v5', 'ALE/VideoChess-ram-v5', 'ALE/Vide
oCube-v5', 'ALE/VideoCube-ram-v5', 'ALE/VideoPinball-v5', 'ALE/VideoPinb
all-ram-v5', 'ALE/Warlords-v5', 'ALE/Warlords-ram-v5', 'ALE/WizardOfWor-
v5', 'ALE/WizardOfWor-ram-v5', 'ALE/WordZapper-v5', 'ALE/WordZapper-ram-
v5', 'ALE/YarsRevenge-v5', 'ALE/YarsRevenge-ram-v5', 'ALE/Zaxxon-v5', 'A
LE/Zaxxon-ram-v5', 'GymV26Environment-v0', 'GymV21Environment-v0'])
```

In [18]:
```python
import gymnasium as gym
# import math
import random
import matplotlib
import matplotlib.pyplot as plt
import numpy as np

import torch
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F

# For visualization
import os
import imageio
from PIL import Image, ImageDraw, ImageFont
import matplotlib.pyplot as plt
%matplotlib inline
from IPython import display
# from IPython.display import Image
```

```python
import plotly.graph_objects as go
from plotly.subplots import make_subplots

from collections import deque

import json

# env = gym.make("CartPole-v1")
# env = gym.make("Taxi-v3")#, render_mode="rgb_array")
# env = gym.make("ALE/DonkeyKong-v5")

# set up matplotlib
is_ipython = 'inline' in matplotlib.get_backend()
if is_ipython:
    from IPython import display

plt.ion()

# if GPU exist, use GPU
device = torch.device(
    "cuda" if torch.cuda.is_available() else
    "mps" if torch.backends.mps.is_available() else
    "cpu"
)
```

In [19]: `gym.__version__`

Out[19]: `'1.0.0a2'`

Making Agent

In [21]:
```python
class Agent:
    def __init__(self, game_id, render_mode):
        self.game_id = game_id # game id
        self.render_mode = render_mode
        self.env = gym.make(game_id, render_mode=render_mode)

    def select_action(self, state, action):
        return action

    # ========== Display on screen ==========
    # Reference: https://stackoverflow.com/questions/40195740/how-to-run-
    def display_state_on_screen(self, episode, step, info=""):
        plt.figure(3)
        plt.clf()
        plt.imshow(self.env.render())
        plt.title("%s | Episode: %d | Step: %d %s" % (self.env.spec.id, e
        plt.axis('off')

        display.clear_output(wait=True)
        display.display(plt.gcf())
    # ========== Display on screen ==========

    # ========== to gif ==========
    # Reference: https://github.com/yuchen071/DQN-for-LunarLander-v2/blob
```

```python
    def add_text_on_img(self, img, score, episode, step):
        img = Image.fromarray(img)
        draw = ImageDraw.Draw(img)
        draw.text((20, 20), f"Score={score: .2f}, Episode={episode}, Step

        return np.array(img)

    def save_frames_as_gif(self, frames, filename, path="gifs/"):
        if not os.path.exists(path):
            os.makedirs(path)

        print("Saving gif...", end="")
        imageio.mimsave(path + filename + ".gif", frames, fps=60)

        print("Done!")
    # ========== to gif ==========
```

In [22]:
```python
class ValueEstimationAgent(Agent):
    def __init__(self, game_id, render_mode, gamma):
        Agent.__init__(self, game_id, render_mode)
        self.gamma = gamma  # discount rate gamma


    def select_action(self, state, action):
        return action
```

In [23]:
```python
class ReinforcementLearningAgent(ValueEstimationAgent):
    def __init__(self, game_id, render_mode, gamma, alpha):
        ValueEstimationAgent.__init__(self, game_id, render_mode, gamma)
        self.alpha = alpha   # alpha learning rate

    def select_action(self, state, action):
            return action
```

Making Q Learning Agent

In [25]:
```python
# Reference: https://huggingface.co/learn/deep-rl-course/unit2/hands-on

class QLearningAgent(ReinforcementLearningAgent):
    def __init__(self, game_id, render_mode, gamma, alpha, epsilon, epsil
        ReinforcementLearningAgent.__init__(self, game_id, render_mode, g
        #self.game_id
        #self.render_mode
        #self.env=env
        #self.alpha=alpha
        #self.gamma=gamma

        self.epsilon = epsilon                # epsilon greedy policy
        self.epsilon_start = epsilon_start    # starting epsilon, usually
        self.epsilon_end = epsilon_end        # ending epsilon, usually 0
        self.epsilon_decay = epsilon_decay    # fixed epsilon decay value
```

```python
        self.max_steps = max_steps            # set max steps for each ep
        # replace max_episode_steps of the env with the given max_steps
        self.env._max_episode_steps = self.max_steps  # TAXI default max

        # store the numer of states and actions of the env for easier acc
        self.num_states = self.env.observation_space.n
        self.num_actions = self.env.action_space.n

        # Q value table in numpy array of array form initialize to 0
        self.Q_space = np.zeros((self.num_states, self.num_actions))  # s

        self.result_model_name = result_model_name


    def select_action(self, state, action):
        return action

    def train(self, episodes):

        # rewards from each episode !
        rewards_per_episode = np.zeros(episodes)

        # record epsilon for display
        epsilon_history = []


        for episode in range(episodes):
            # Reset the environment
            observation = self.env.reset()
            state = observation[0]    # state is the 1st elem in the self.

            # acc total reward per episode
            total_reward = 0



            for step in range(self.max_steps):

                # =============== Epsilon Greedy Algorithm Section ======
                # if random number between [0 : 1) > epsilon --> exploita
                if(random.random() > self.epsilon):
                    # read argument action with max q value among the 6 a
                    action = np.argmax(self.Q_space[state][:])
                # else --> exploration
                else:
                    action = self.env.action_space.sample()  # pick rando
                # =============== Epsilon Greedy Algorithm Section ======

                # Epsilon decay as we need fewer exploration as time pass
                # Exponential Decay of epsilon
                # refernce: https://medium.com/@CalebMBowyer/strategies-f
                decay_factor = (self.epsilon_end/self.epsilon_start)**(1/
                # update self.epsilon
                self.epsilon = self.epsilon * decay_factor
```

```python
            # append updated epsilon to the record
            epsilon_history.append(self.epsilon)

            # https://colab.research.google.com/github/erodola/ML-s2-
            # decay_rate = 0.0005 # 1000
            # self.epsilon = self.epsilon_end + (self.epsilon_start -


            # Take the action and observe the outcome new state and r
            new_state, reward, terminated, truncated, info = self.env

            # acc reward to toal_reward to keep track of the total_re
            total_reward += reward

            # Q-learning actual part
            sample = reward + self.gamma * np.max(self.Q_space[new_st
            self.Q_space[state][action] = (1 - self.alpha) * self.Q_s

            # if the number of steps elapsed >= max episode steps (20
            # If terminated --> finish current episode
            if terminated == True or truncated == True:
                break

            # update to state to new_state
            state = new_state


        # store the total reward of each episode for record
        rewards_per_episode[episode] = total_reward

        # print the progress
        i = episode
        if (i + 1) % 100 == 0:
            print(f"Episode {i+1}/{episodes}, Avg Reward: {np.mean(re


    self.env.close()

    # Create new graph
    plt.figure(1)

    # Plot average rewards vs episodes
    sum_rewards = np.zeros(episodes)
    for x in range(episodes):
        sum_rewards[x] = np.sum(rewards_per_episode[max(0, x-100):(x+
    plt.subplot(121) # plot on a 1 row x 2 col grid, at cell 1
    plt.plot(sum_rewards)

    # Plot epsilon decay vs episodes
    plt.subplot(122) # plot on a 1 row x 2 col grid, at cell 2
    plt.plot(epsilon_history)

    # Save plots
    plt.savefig(self.result_model_name + '.png')

    return rewards_per_episode
```

```python
def test(self, test_episodes=10, display=False):

    episode_reward_lst = []

    frames = []


    for episode in range(test_episodes): # total_test_episodes == 10
        observation = self.env.reset()
        state = observation[0]

        total_reward = 0
        print("========================================================
        print("EPISODE ", episode)


        # added
        steps_count = 0
        for step in range(self.max_steps):

            # Take the action (index) that have the maximum expected
            action = np.argmax(self.Q_space[state][:])

            # apply selected action
            new_state, reward, terminated, truncated, info = self.env

            # add up reward for later calculation
            total_reward += reward
            # increment step for displaying state on screen
            steps_count += 1

            if (display == True):
                # Display on Screen
                self.display_state_on_screen(episode, steps_count)  #


            if terminated == True or truncated == True :
                episode_reward_lst.append(total_reward)
                print ("Score of total rewards:", total_reward)
                break
            state = new_state  # Update state to new_state

    self.env.close()
    # print("Total Steps taken:", steps_count)
    print ("Average score over time: " +  str(sum(episode_reward_lst)
    return episode_reward_lst



# turn test scenario to gif
def test_to_gif(self, filename="q_learning_result", episodes=5):
```

```python
        frames = []
        for episode in range(episodes):
            observation = self.env.reset()
            state = observation[0]

            score = 0
            for step in range(self.max_steps):
                # store rgb ender to frame
                frame = self.env.render()
                # store the rendered frame to list
                frames.append(self.add_text_on_img(frame, score, episode,

                # selected from the Q value table
                action = np.argmax(self.Q_space[state][:])


                # print current episode, action, state
                print(f"episodes: {episode}, acton: {action}, state: {sta


                # apply selected action
                state, reward, terminated, truncated, info = self.env.ste

                score += reward

                if terminated == True or truncated == True:
                    break
        self.env.close()
        # turn frames into gif
        self.save_frames_as_gif(frames, filename=filename)
```

Deep Q Learning (DQN) agent

```python
In [27]:  # Defining models
          # Reference:
          # https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.htm
          # https://colab.research.google.com/github/pytorch/tutorials/blob/gh-page
          # https://www.youtube.com/watch?si=fXl8PEnb0TrV9Vwt&v=EUrWGTCGzlA&feature

          # DQN model
          class DQN(nn.Module):
              # 3 layers multi layers learning network
              def __init__(self, in_states, h1_nodes, out_actions):
                  super().__init__()
                  self.fc1 = nn.Linear(in_states, h1_nodes)       # Input layer = 500
                  self.fc2 = nn.Linear(h1_nodes, h1_nodes)        # hidden layer
                  self.out = nn.Linear(h1_nodes, out_actions)     # output layer = 6

              def forward(self, x):
                  # Rectified Linear Unit (ReLU)
```

```python
        x = F.relu(self.fc1(x))        # run Input layer
        x = F.relu(self.fc2(x))        # run hidden layer
        x = self.out(x)                # run output layer
        return x

# Define memory for Experience Replay
class ReplayMemory():
    def __init__(self, maxlen):
        self.memory = deque([], maxlen=maxlen)

    def append(self, transition):
        self.memory.append(transition)

    def sample(self, sample_size):
        return random.sample(self.memory, sample_size)

    def __len__(self):
        return len(self.memory)
```

In [28]:
```python
class DQNAgent(QLearningAgent):
    def __init__(self, game_id, render_mode, gamma, alpha, epsilon, epsil
        QLearningAgent.__init__(self, game_id, render_mode, gamma, alpha,

        #self.game_id
        #self.render_mode
        #self.env=env
        #self.alpha=alpha
        #self.gamma=gamma

        #self.epsilon = epsilon                     # epsilon greedy policy
        #self.epsilon_start = epsilon_start     # starting epsilon, usuall
        #self.epsilon_end = epsilon_end         # ending epsilon, usually
        #self.epsilon_decay = epsilon_decay     # fixed epsilon decay valu
        #self.max_steps = max_steps             # set max steps for each e
        ## replace max_episode_steps of the env with the given max_steps
        #self.env._max_episode_steps = self.max_steps  # TAXI default max

        ## store the numer of states and actions of the env for easier ac
        #self.num_states = self.env.observation_space.n
        #self.num_actions = self.env.action_space.n

        ## Q value table in numpy array of array form initialize to 0
        #self.Q_space = np.zeros((self.num_states, self.num_actions))  #

        #self.result_model_name = result_model_name

    # ---------------------- Special Parameters for DQN only -----------
    network_sync_rate_tau = 10   # after 10 steps, sync policy network an
    replay_memory_size = 10000
    mini_batch_size = 64

    loss_fn = nn.MSELoss()   # Mean Squared Error Loss function
    optimizer = None         # initialize later on

    ACTIONS = ['S', 'N', 'E', 'W', 'P', 'D']  # South, North, East, West,
    # ---------------------- Special Parameters for DQN only -----------
```

```python
    def train(self, episodes):

        # epsilon = 1


        # set up ReplayMemory obj to store the experience
        memory = ReplayMemory(self.replay_memory_size)

        # need 2 neural networks
        # 1) policy network for training
        # 2) target network makes use of the deep q learning equation
        policy_dqn = DQN(in_states=self.num_states, h1_nodes=128, out_act
        target_dqn = DQN(in_states=self.num_states, h1_nodes=128, out_act


        # copy policy network to target network
        target_dqn.load_state_dict(policy_dqn.state_dict())

        # neural network initially comes with a random set of weights and
        print('Non-trained Policy Network Model with randomly initialized
        self.print_dqn(policy_dqn)

        # initialize and store Adam optimizer
        self.optimizer = torch.optim.Adam(policy_dqn.parameters(), lr=sel

        # rewards from each episode !
        rewards_per_episode = np.zeros(episodes)

        # record epsilon for display
        epsilon_history = []

        # track how many steps has been taken, sync policy network and ta
        step_count = 0

        for episode in range(episodes):
            state, _ = self.env.reset()
            total_reward = 0
            terminated = False    # True when the taxi drops off the pass
            truncated = False     # True when the length of steps in a si

            while not terminated and not truncated:


                # =============== Epsilon Greedy Algorithm Section ======
                if random.random() < self.epsilon:
                    # Random action
                    action = self.env.action_space.sample()
                else:
                    # torch.no_grad() tells pytorch not to calculate new
                    with torch.no_grad():
                        # Use policy network to calculate a set of q valu
                        # self.state_to_dqn_input() encodes state to one-
                        action = policy_dqn(self.state_to_dqn_input(state
                # =============== Epsilon Greedy Algorithm Section ======

                # run the action in the env
                new_state, reward, terminated, truncated, _ = self.env.st
```

```python
            # Memoize the experience, append to the ReplayMemory
            memory.append((state, action, new_state, reward, terminat

            # update state to the new state
            state = new_state

            # increment
            total_reward += reward
            step_count += 1

        # store the total reward of each episode for record
        rewards_per_episode[episode] = total_reward


        # Check if enough experience has been collected for a batch
        if len(memory) > self.mini_batch_size:
            mini_batch = memory.sample(self.mini_batch_size)


            # optimize
            self.optimize(mini_batch, policy_dqn, target_dqn)


            # Exponential epsilon decay factor !!!!
            decay_factor = (self.epsilon_end/self.epsilon_start)**(1/
            self.epsilon = self.epsilon * decay_factor

            # https://colab.research.google.com/github/erodola/ML-s2-
            # decay_rate = 0.0005 # 1000
            # episode = i  # rename i to episode for the calculation
            # print(self.epsilon, episode)
            # self.epsilon = self.epsilon_end + (self.epsilon_start -

            # record the epsilon value for later
            epsilon_history.append(self.epsilon)

            # if overstep, load instead of train
            if step_count > self.network_sync_rate_tau:
                target_dqn.load_state_dict(policy_dqn.state_dict())
                step_count = 0


        # print the progress
        if (episode + 1) % 100 == 0:
            print(f"Episode {episode+1}/{episodes}, Avg Reward: {np.m

    self.env.close()

    torch.save(policy_dqn.state_dict(), self.result_model_name + ".pt


    print("!!!!!!!!!!!!!!!!!Trained and Stored the model successfully!
```

```python
        # Display the training progress into png

        # Create new graph
        plt.figure(1)

        # Plot average rewardsvs episodes
        sum_rewards = np.zeros(episodes)
        for x in range(episodes):
            sum_rewards[x] = np.sum(rewards_per_episode[max(0, x-100):(x+
        plt.subplot(121) # plot on a 1 row x 2 col grid, at cell 1
        plt.plot(sum_rewards)

        # Plot epsilon decay vs episodes
        plt.subplot(122) # plot on a 1 row x 2 col grid, at cell 2
        plt.plot(epsilon_history)

        # Save plots
        plt.savefig(self.result_model_name + '.png')

        return rewards_per_episode

    def optimize(self, mini_batch, policy_dqn, target_dqn):
        current_q_list = []
        target_q_list = []

        for state, action, new_state, reward, terminated in mini_batch:
            if terminated:
                target = torch.FloatTensor([reward])
            else:
                with torch.no_grad():
                    # torch.no_grad() --> don't train but predict
                    target = torch.FloatTensor(
                        reward + self.gamma * target_dqn(self.state_to_dq
                    )
            # Q value update
            current_q = policy_dqn(self.state_to_dqn_input(state))
            target_q = current_q.clone().detach()
            target_q[action] = target

            current_q_list.append(current_q)
            target_q_list.append(target_q)

        # set loss function
        loss = self.loss_fn(torch.stack(current_q_list), torch.stack(targ


        self.optimizer.zero_grad()
        loss.backward()
        self.optimizer.step()

    # turn state into one-hot encoding tensor vector
    # tensor([1., 1., 0., 1., 0., 1., 0., 0., 0., 1., 1., 1., 1., 0., 1.,
    def state_to_dqn_input(self, state)-> torch.Tensor:
        input_tensor = torch.zeros(self.num_states)
        input_tensor[state] = 1
        return input_tensor
```

```python
    def test(self, test_episodes=10, display=False):

        # load and set policy network
        policy_dqn = DQN(in_states=self.num_states, h1_nodes=128, out_act
        policy_dqn.load_state_dict(torch.load(self.result_model_name + ".
        policy_dqn.eval()

        print('Trained Policy Network Model from the .pt file:')
        self.print_dqn(policy_dqn)

        episode_reward_lst = list()


        for episode in range(test_episodes):
            state, _ = self.env.reset()
            total_reward = 0
            terminated = False
            truncated = False


            # steps_count for dispaly
            steps_count = 0

            # while not terminated and not truncated:
            for step in range(self.max_steps):

                with torch.no_grad():
                    action = policy_dqn(self.state_to_dqn_input(state)).a

                state, reward, terminated, truncated, _ = self.env.step(a
                total_reward += reward

                steps_count += 1

                # only display the Game if being asked to do so
                if (display == True):
                    self.display_state_on_screen(episode, steps_count)


                if (terminated == True or truncated == True):
                    episode_reward_lst.append(total_reward)
                    print(f"Episode {episode+1}: Total Reward = {total_re
                    break

            # print(f"Episode {episode+1}: Total Reward = {total_reward}"
            # episode_reward_lst.append(total_reward)

        self.env.close()
        print(f"Average Reward across all Episodes: = {sum(episode_reward

        return episode_reward_lst

    # print the initial states and values before the training
    def print_dqn(self, dqn):
        for s in range(self.num_states):
```

```python
            q_values = dqn(self.state_to_dqn_input(s)).tolist()
            q_values_str = ' '.join([f"{q:+.2f}" for q in q_values])
            best_action = self.ACTIONS[dqn(self.state_to_dqn_input(s)).ar
            print(f'State {s:03}, Action {best_action}, Q-values: [{q_val

    def test_to_gif(self, filename="dqn_learning_result", episodes=3):
        # capture each frame
        frames = []

        # make DQN for loading policy network
        policy_dqn = DQN(in_states=self.num_states, h1_nodes=128, out_act
        # load and eval the trained policy network
        policy_dqn.load_state_dict(torch.load(self.result_model_name + ".
        policy_dqn.eval()

        print('===== Trained Policy ===== :')
        self.print_dqn(policy_dqn)

        episode_reward_lst = list()

        for episode in range(episodes):
            # reset env for each episodes == episode
            observation = self.env.reset()
            state = observation[0]


            total_reward = 0
            steps_count = 0

            # for display on the image
            score = 0
            for step in range(self.max_steps):
                frame = self.env.render()
                frames.append(self.add_text_on_img(frame, score, episode,

                # tell torch to not train but to predict
                with torch.no_grad():
                    # read max action from the trained policy network
                    action = policy_dqn(self.state_to_dqn_input(state)).a

                print(f"episodes: {episode}, acton: {action}, state: {sta

                state, reward, terminated, truncated, _ = self.env.step(a

                score += reward
                total_reward += reward
                steps_count += 1

                if terminated == True or truncated == True:

                    episode_reward_lst.append(total_reward)
                    print(f"Episode {episode+1}: Total Reward = {total_re
                    break
        self.env.close()
        self.save_frames_as_gif(frames, filename=filename)
        print(f"Average Reward across all Episodes: = {sum(episode_reward
```

Testing and Comparison

Controlled Parameter Training

Q Learning Agent with Controlled Parameters

```
In [32]: my_q_agent_controlled = QLearningAgent(game_id="Taxi-v3", render_mode="rg
                                    gamma=0.95, alpha=0.5, epsilon=1.0
                                    epsilon_decay=1000, epsilon_end=0.
                                    result_model_name="taxi_q_controll
```

```
In [33]: %%time
         my_q_agent_controlled_training_rewards = my_q_agent_controlled.train(1000
```

```
Episode 100/100000, Avg Reward: -481.41, Epsilon: 0.30
Episode 200/100000, Avg Reward: -136.33, Epsilon: 0.16
Episode 300/100000, Avg Reward: -28.87, Epsilon: 0.12
Episode 400/100000, Avg Reward: -5.96, Epsilon: 0.11
Episode 500/100000, Avg Reward: -2.52, Epsilon: 0.10
Episode 600/100000, Avg Reward: 2.18, Epsilon: 0.09
Episode 700/100000, Avg Reward: 2.26, Epsilon: 0.08
Episode 800/100000, Avg Reward: 2.91, Epsilon: 0.07
Episode 900/100000, Avg Reward: 4.50, Epsilon: 0.06
Episode 1000/100000, Avg Reward: 4.72, Epsilon: 0.06
Episode 1100/100000, Avg Reward: 4.49, Epsilon: 0.05
Episode 1200/100000, Avg Reward: 5.35, Epsilon: 0.05
Episode 1300/100000, Avg Reward: 4.72, Epsilon: 0.04
Episode 1400/100000, Avg Reward: 5.94, Epsilon: 0.04
Episode 1500/100000, Avg Reward: 6.99, Epsilon: 0.04
Episode 1600/100000, Avg Reward: 6.18, Epsilon: 0.03
Episode 1700/100000, Avg Reward: 6.21, Epsilon: 0.03
Episode 1800/100000, Avg Reward: 6.12, Epsilon: 0.03
Episode 1900/100000, Avg Reward: 7.98, Epsilon: 0.03
Episode 2000/100000, Avg Reward: 7.19, Epsilon: 0.02
Episode 2100/100000, Avg Reward: 7.10, Epsilon: 0.02
Episode 2200/100000, Avg Reward: 6.95, Epsilon: 0.02
Episode 2300/100000, Avg Reward: 7.22, Epsilon: 0.02
Episode 2400/100000, Avg Reward: 6.90, Epsilon: 0.02
Episode 2500/100000, Avg Reward: 7.21, Epsilon: 0.01
Episode 2600/100000, Avg Reward: 7.09, Epsilon: 0.01
Episode 2700/100000, Avg Reward: 6.84, Epsilon: 0.01
Episode 2800/100000, Avg Reward: 7.85, Epsilon: 0.01
Episode 2900/100000, Avg Reward: 7.20, Epsilon: 0.01
Episode 3000/100000, Avg Reward: 7.77, Epsilon: 0.01
Episode 3100/100000, Avg Reward: 7.00, Epsilon: 0.01
Episode 3200/100000, Avg Reward: 7.51, Epsilon: 0.01
Episode 3300/100000, Avg Reward: 7.49, Epsilon: 0.01
Episode 3400/100000, Avg Reward: 7.20, Epsilon: 0.01
Episode 3500/100000, Avg Reward: 7.58, Epsilon: 0.01
Episode 3600/100000, Avg Reward: 7.77, Epsilon: 0.01
Episode 3700/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 3800/100000, Avg Reward: 7.27, Epsilon: 0.00
Episode 3900/100000, Avg Reward: 7.67, Epsilon: 0.00
```

```
Episode 4000/100000, Avg Reward: 7.22, Epsilon: 0.00
Episode 4100/100000, Avg Reward: 7.21, Epsilon: 0.00
Episode 4200/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 4300/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 4400/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 4500/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 4600/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 4700/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 4800/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 4900/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 5000/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 5100/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 5200/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 5300/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 5400/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 5500/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 5600/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 5700/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 5800/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 5900/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 6000/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 6100/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 6200/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 6300/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 6400/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 6500/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 6600/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 6700/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 6800/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 6900/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 7000/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 7100/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 7200/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 7300/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 7400/100000, Avg Reward: 8.34, Epsilon: 0.00
Episode 7500/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 7600/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 7700/100000, Avg Reward: 7.26, Epsilon: 0.00
Episode 7800/100000, Avg Reward: 7.28, Epsilon: 0.00
Episode 7900/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 8000/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 8100/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 8200/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 8300/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 8400/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 8500/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 8600/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 8700/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 8800/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 8900/100000, Avg Reward: 7.44, Epsilon: 0.00
Episode 9000/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 9100/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 9200/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 9300/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 9400/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 9500/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 9600/100000, Avg Reward: 7.95, Epsilon: 0.00
```

```
Episode 9700/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 9800/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 9900/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 10000/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 10100/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 10200/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 10300/100000, Avg Reward: 8.54, Epsilon: 0.00
Episode 10400/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 10500/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 10600/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 10700/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 10800/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 10900/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 11000/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 11100/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 11200/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 11300/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 11400/100000, Avg Reward: 7.45, Epsilon: 0.00
Episode 11500/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 11600/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 11700/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 11800/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 11900/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 12000/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 12100/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 12200/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 12300/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 12400/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 12500/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 12600/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 12700/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 12800/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 12900/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 13000/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 13100/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 13200/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 13300/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 13400/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 13500/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 13600/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 13700/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 13800/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 13900/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 14000/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 14100/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 14200/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 14300/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 14400/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 14500/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 14600/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 14700/100000, Avg Reward: 8.37, Epsilon: 0.00
Episode 14800/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 14900/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 15000/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 15100/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 15200/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 15300/100000, Avg Reward: 7.68, Epsilon: 0.00
```

```
Episode 15400/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 15500/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 15600/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 15700/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 15800/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 15900/100000, Avg Reward: 7.42, Epsilon: 0.00
Episode 16000/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 16100/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 16200/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 16300/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 16400/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 16500/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 16600/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 16700/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 16800/100000, Avg Reward: 8.34, Epsilon: 0.00
Episode 16900/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 17000/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 17100/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 17200/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 17300/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 17400/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 17500/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 17600/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 17700/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 17800/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 17900/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 18000/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 18100/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 18200/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 18300/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 18400/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 18500/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 18600/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 18700/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 18800/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 18900/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 19000/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 19100/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 19200/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 19300/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 19400/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 19500/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 19600/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 19700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 19800/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 19900/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 20000/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 20100/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 20200/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 20300/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 20400/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 20500/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 20600/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 20700/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 20800/100000, Avg Reward: 7.45, Epsilon: 0.00
Episode 20900/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 21000/100000, Avg Reward: 7.85, Epsilon: 0.00
```

```
Episode 21100/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 21200/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 21300/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 21400/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 21500/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 21600/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 21700/100000, Avg Reward: 7.50, Epsilon: 0.00
Episode 21800/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 21900/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 22000/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 22100/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 22200/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 22300/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 22400/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 22500/100000, Avg Reward: 7.45, Epsilon: 0.00
Episode 22600/100000, Avg Reward: 7.25, Epsilon: 0.00
Episode 22700/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 22800/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 22900/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 23000/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 23100/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 23200/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 23300/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 23400/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 23500/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 23600/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 23700/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 23800/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 23900/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 24000/100000, Avg Reward: 7.24, Epsilon: 0.00
Episode 24100/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 24200/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 24300/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 24400/100000, Avg Reward: 8.50, Epsilon: 0.00
Episode 24500/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 24600/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 24700/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 24800/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 24900/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 25000/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 25100/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 25200/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 25300/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 25400/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 25500/100000, Avg Reward: 7.35, Epsilon: 0.00
Episode 25600/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 25700/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 25800/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 25900/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 26000/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 26100/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 26200/100000, Avg Reward: 7.23, Epsilon: 0.00
Episode 26300/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 26400/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 26500/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 26600/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 26700/100000, Avg Reward: 7.81, Epsilon: 0.00
```

```
Episode 26800/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 26900/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 27000/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 27100/100000, Avg Reward: 8.43, Epsilon: 0.00
Episode 27200/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 27300/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 27400/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 27500/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 27600/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 27700/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 27800/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 27900/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 28000/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 28100/100000, Avg Reward: 8.56, Epsilon: 0.00
Episode 28200/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 28300/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 28400/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 28500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 28600/100000, Avg Reward: 8.53, Epsilon: 0.00
Episode 28700/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 28800/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 28900/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 29000/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 29100/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 29200/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 29300/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 29400/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 29500/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 29600/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 29700/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 29800/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 29900/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 30000/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 30100/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 30200/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 30300/100000, Avg Reward: 7.45, Epsilon: 0.00
Episode 30400/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 30500/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 30600/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 30700/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 30800/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 30900/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 31000/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 31100/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 31200/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 31300/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 31400/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 31500/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 31600/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 31700/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 31800/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 31900/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 32000/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 32100/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 32200/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 32300/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 32400/100000, Avg Reward: 7.59, Epsilon: 0.00
```

```
Episode 32500/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 32600/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 32700/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 32800/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 32900/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 33000/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 33100/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 33200/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 33300/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 33400/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 33500/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 33600/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 33700/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 33800/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 33900/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 34000/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 34100/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 34200/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 34300/100000, Avg Reward: 7.42, Epsilon: 0.00
Episode 34400/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 34500/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 34600/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 34700/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 34800/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 34900/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 35000/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 35100/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 35200/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 35300/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 35400/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 35500/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 35600/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 35700/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 35800/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 35900/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 36000/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 36100/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 36200/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 36300/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 36400/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 36500/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 36600/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 36700/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 36800/100000, Avg Reward: 8.43, Epsilon: 0.00
Episode 36900/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 37000/100000, Avg Reward: 7.35, Epsilon: 0.00
Episode 37100/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 37200/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 37300/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 37400/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 37500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 37600/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 37700/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 37800/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 37900/100000, Avg Reward: 7.47, Epsilon: 0.00
Episode 38000/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 38100/100000, Avg Reward: 8.16, Epsilon: 0.00
```

```
Episode 38200/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 38300/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 38400/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 38500/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 38600/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 38700/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 38800/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 38900/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 39000/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 39100/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 39200/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 39300/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 39400/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 39500/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 39600/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 39700/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 39800/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 39900/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 40000/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 40100/100000, Avg Reward: 8.40, Epsilon: 0.00
Episode 40200/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 40300/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 40400/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 40500/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 40600/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 40700/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 40800/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 40900/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 41000/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 41100/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 41200/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 41300/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 41400/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 41500/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 41600/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 41700/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 41800/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 41900/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 42000/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 42100/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 42200/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 42300/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 42400/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 42500/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 42600/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 42700/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 42800/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 42900/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 43000/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 43100/100000, Avg Reward: 8.34, Epsilon: 0.00
Episode 43200/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 43300/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 43400/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 43500/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 43600/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 43700/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 43800/100000, Avg Reward: 7.35, Epsilon: 0.00
```

```
Episode 43900/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 44000/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 44100/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 44200/100000, Avg Reward: 7.21, Epsilon: 0.00
Episode 44300/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 44400/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 44500/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 44600/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 44700/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 44800/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 44900/100000, Avg Reward: 8.43, Epsilon: 0.00
Episode 45000/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 45100/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 45200/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 45300/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 45400/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 45500/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 45600/100000, Avg Reward: 8.43, Epsilon: 0.00
Episode 45700/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 45800/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 45900/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 46000/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 46100/100000, Avg Reward: 7.28, Epsilon: 0.00
Episode 46200/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 46300/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 46400/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 46500/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 46600/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 46700/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 46800/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 46900/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 47000/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 47100/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 47200/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 47300/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 47400/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 47500/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 47600/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 47700/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 47800/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 47900/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 48000/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 48100/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 48200/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 48300/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 48400/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 48500/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 48600/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 48700/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 48800/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 48900/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 49000/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 49100/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 49200/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 49300/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 49400/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 49500/100000, Avg Reward: 7.88, Epsilon: 0.00
```

```
Episode 49600/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 49700/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 49800/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 49900/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 50000/100000, Avg Reward: 8.34, Epsilon: 0.00
Episode 50100/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 50200/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 50300/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 50400/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 50500/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 50600/100000, Avg Reward: 7.12, Epsilon: 0.00
Episode 50700/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 50800/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 50900/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 51000/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 51100/100000, Avg Reward: 7.45, Epsilon: 0.00
Episode 51200/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 51300/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 51400/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 51500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 51600/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 51700/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 51800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 51900/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 52000/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 52100/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 52200/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 52300/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 52400/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 52500/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 52600/100000, Avg Reward: 7.50, Epsilon: 0.00
Episode 52700/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 52800/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 52900/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 53000/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 53100/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 53200/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 53300/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 53400/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 53500/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 53600/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 53700/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 53800/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 53900/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 54000/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 54100/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 54200/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 54300/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 54400/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 54500/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 54600/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 54700/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 54800/100000, Avg Reward: 8.34, Epsilon: 0.00
Episode 54900/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 55000/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 55100/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 55200/100000, Avg Reward: 7.88, Epsilon: 0.00
```

```
Episode 55300/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 55400/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 55500/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 55600/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 55700/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 55800/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 55900/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 56000/100000, Avg Reward: 8.33, Epsilon: 0.00
Episode 56100/100000, Avg Reward: 8.45, Epsilon: 0.00
Episode 56200/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 56300/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 56400/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 56500/100000, Avg Reward: 7.29, Epsilon: 0.00
Episode 56600/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 56700/100000, Avg Reward: 7.44, Epsilon: 0.00
Episode 56800/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 56900/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 57000/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 57100/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 57200/100000, Avg Reward: 7.17, Epsilon: 0.00
Episode 57300/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 57400/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 57500/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 57600/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 57700/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 57800/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 57900/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 58000/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 58100/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 58200/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 58300/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 58400/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 58500/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 58600/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 58700/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 58800/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 58900/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 59000/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 59100/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 59200/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 59300/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 59400/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 59500/100000, Avg Reward: 7.47, Epsilon: 0.00
Episode 59600/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 59700/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 59800/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 59900/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 60000/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 60100/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 60200/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 60300/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 60400/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 60500/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 60600/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 60700/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 60800/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 60900/100000, Avg Reward: 7.99, Epsilon: 0.00
```

```
Episode 61000/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 61100/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 61200/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 61300/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 61400/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 61500/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 61600/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 61700/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 61800/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 61900/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 62000/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 62100/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 62200/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 62300/100000, Avg Reward: 7.25, Epsilon: 0.00
Episode 62400/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 62500/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 62600/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 62700/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 62800/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 62900/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 63000/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 63100/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 63200/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 63300/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 63400/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 63500/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 63600/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 63700/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 63800/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 63900/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 64000/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 64100/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 64200/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 64300/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 64400/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 64500/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 64600/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 64700/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 64800/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 64900/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 65000/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 65100/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 65200/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 65300/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 65400/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 65500/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 65600/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 65700/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 65800/100000, Avg Reward: 7.07, Epsilon: 0.00
Episode 65900/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 66000/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 66100/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 66200/100000, Avg Reward: 8.36, Epsilon: 0.00
Episode 66300/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 66400/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 66500/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 66600/100000, Avg Reward: 7.73, Epsilon: 0.00
```

```
Episode 66700/100000, Avg Reward: 8.34, Epsilon: 0.00
Episode 66800/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 66900/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 67000/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 67100/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 67200/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 67300/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 67400/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 67500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 67600/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 67700/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 67800/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 67900/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 68000/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 68100/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 68200/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 68300/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 68400/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 68500/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 68600/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 68700/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 68800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 68900/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 69000/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 69100/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 69200/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 69300/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 69400/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 69500/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 69600/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 69700/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 69800/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 69900/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 70000/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 70100/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 70200/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 70300/100000, Avg Reward: 8.36, Epsilon: 0.00
Episode 70400/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 70500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 70600/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 70700/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 70800/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 70900/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 71000/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 71100/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 71200/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 71300/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 71400/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 71500/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 71600/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 71700/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 71800/100000, Avg Reward: 8.71, Epsilon: 0.00
Episode 71900/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 72000/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 72100/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 72200/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 72300/100000, Avg Reward: 7.89, Epsilon: 0.00
```

```
Episode 72400/100000, Avg Reward: 8.66, Epsilon: 0.00
Episode 72500/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 72600/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 72700/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 72800/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 72900/100000, Avg Reward: 8.54, Epsilon: 0.00
Episode 73000/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 73100/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 73200/100000, Avg Reward: 7.33, Epsilon: 0.00
Episode 73300/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 73400/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 73500/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 73600/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 73700/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 73800/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 73900/100000, Avg Reward: 8.41, Epsilon: 0.00
Episode 74000/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 74100/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 74200/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 74300/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 74400/100000, Avg Reward: 8.36, Epsilon: 0.00
Episode 74500/100000, Avg Reward: 8.44, Epsilon: 0.00
Episode 74600/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 74700/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 74800/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 74900/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 75000/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 75100/100000, Avg Reward: 7.44, Epsilon: 0.00
Episode 75200/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 75300/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 75400/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 75500/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 75600/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 75700/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 75800/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 75900/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 76000/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 76100/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 76200/100000, Avg Reward: 8.36, Epsilon: 0.00
Episode 76300/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 76400/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 76500/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 76600/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 76700/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 76800/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 76900/100000, Avg Reward: 7.50, Epsilon: 0.00
Episode 77000/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 77100/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 77200/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 77300/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 77400/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 77500/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 77600/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 77700/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 77800/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 77900/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 78000/100000, Avg Reward: 7.67, Epsilon: 0.00
```

```
Episode 78100/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 78200/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 78300/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 78400/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 78500/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 78600/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 78700/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 78800/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 78900/100000, Avg Reward: 8.33, Epsilon: 0.00
Episode 79000/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 79100/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 79200/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 79300/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 79400/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 79500/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 79600/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 79700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 79800/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 79900/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 80000/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 80100/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 80200/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 80300/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 80400/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 80500/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 80600/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 80700/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 80800/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 80900/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 81000/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 81100/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 81200/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 81300/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 81400/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 81500/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 81600/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 81700/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 81800/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 81900/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 82000/100000, Avg Reward: 7.39, Epsilon: 0.00
Episode 82100/100000, Avg Reward: 7.42, Epsilon: 0.00
Episode 82200/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 82300/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 82400/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 82500/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 82600/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 82700/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 82800/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 82900/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 83000/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 83100/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 83200/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 83300/100000, Avg Reward: 7.36, Epsilon: 0.00
Episode 83400/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 83500/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 83600/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 83700/100000, Avg Reward: 8.06, Epsilon: 0.00
```

```
Episode 83800/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 83900/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 84000/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 84100/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 84200/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 84300/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 84400/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 84500/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 84600/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 84700/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 84800/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 84900/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 85000/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 85100/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 85200/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 85300/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 85400/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 85500/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 85600/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 85700/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 85800/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 85900/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 86000/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 86100/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 86200/100000, Avg Reward: 7.50, Epsilon: 0.00
Episode 86300/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 86400/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 86500/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 86600/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 86700/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 86800/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 86900/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 87000/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 87100/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 87200/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 87300/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 87400/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 87500/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 87600/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 87700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 87800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 87900/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 88000/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 88100/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 88200/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 88300/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 88400/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 88500/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 88600/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 88700/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 88800/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 88900/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 89000/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 89100/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 89200/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 89300/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 89400/100000, Avg Reward: 7.54, Epsilon: 0.00
```

```
Episode 89500/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 89600/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 89700/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 89800/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 89900/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 90000/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 90100/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 90200/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 90300/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 90400/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 90500/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 90600/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 90700/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 90800/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 90900/100000, Avg Reward: 7.39, Epsilon: 0.00
Episode 91000/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 91100/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 91200/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 91300/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 91400/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 91500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 91600/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 91700/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 91800/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 91900/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 92000/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 92100/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 92200/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 92300/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 92400/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 92500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 92600/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 92700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 92800/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 92900/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 93000/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 93100/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 93200/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 93300/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 93400/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 93500/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 93600/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 93700/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 93800/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 93900/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 94000/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 94100/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 94200/100000, Avg Reward: 8.61, Epsilon: 0.00
Episode 94300/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 94400/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 94500/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 94600/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 94700/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 94800/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 94900/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 95000/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 95100/100000, Avg Reward: 7.75, Epsilon: 0.00
```

```
Episode 95200/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 95300/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 95400/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 95500/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 95600/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 95700/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 95800/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 95900/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 96000/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 96100/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 96200/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 96300/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 96400/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 96500/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 96600/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 96700/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 96800/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 96900/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 97000/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 97100/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 97200/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 97300/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 97400/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 97500/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 97600/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 97700/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 97800/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 97900/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 98000/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 98100/100000, Avg Reward: 7.42, Epsilon: 0.00
Episode 98200/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 98300/100000, Avg Reward: 8.59, Epsilon: 0.00
Episode 98400/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 98500/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 98600/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 98700/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 98800/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 98900/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 99000/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 99100/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 99200/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 99300/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 99400/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 99500/100000, Avg Reward: 7.25, Epsilon: 0.00
Episode 99600/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 99700/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 99800/100000, Avg Reward: 7.50, Epsilon: 0.00
Episode 99900/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 100000/100000, Avg Reward: 7.99, Epsilon: 0.00
CPU times: user 11.8 s, sys: 1.67 s, total: 13.4 s
Wall time: 11.7 s
```
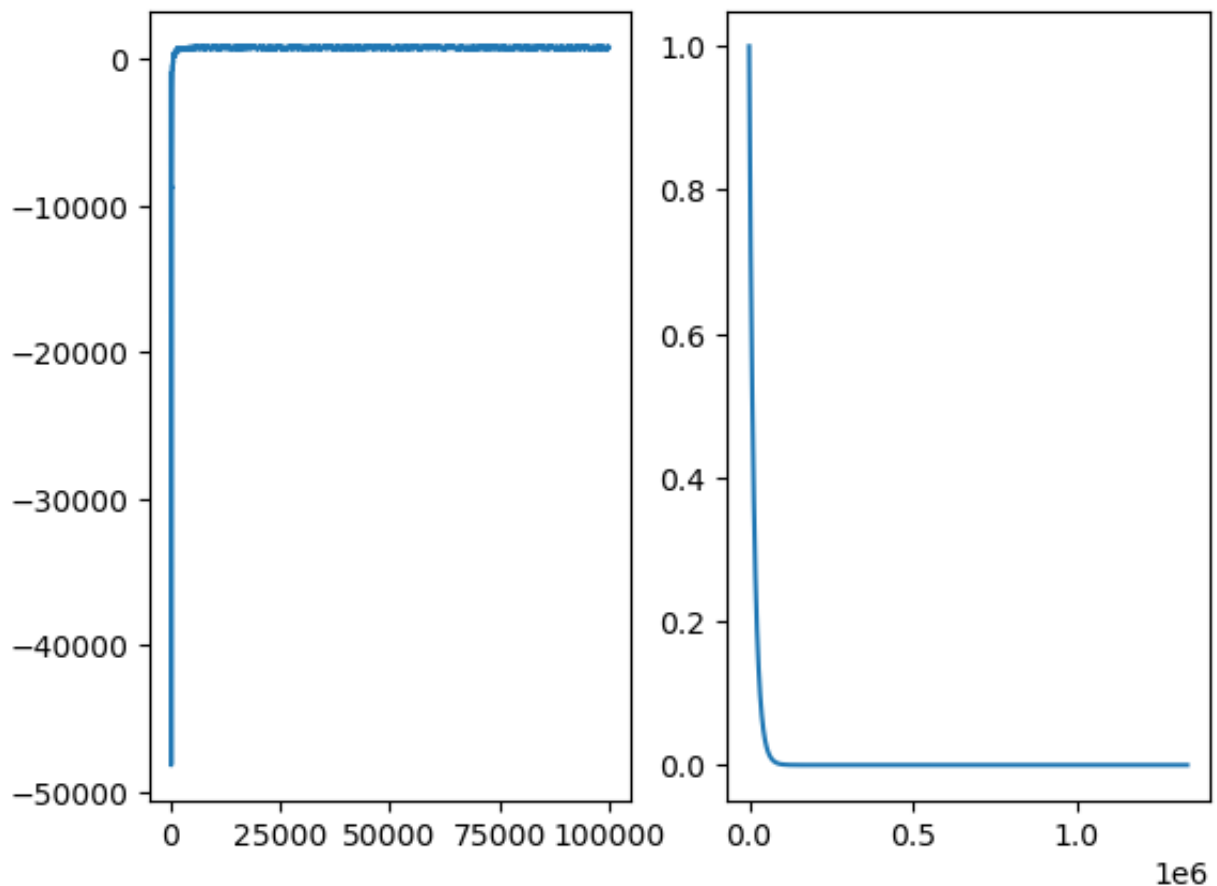
```
In [34]:  print(my_q_agent_controlled_training_rewards)

          [-920. -794. -830. ...    9.    7.   10.]
```

DQN Agent with Controlled Parameters

```
In [36]:  my_dqn_agent_controlled = DQNAgent(game_id="Taxi-v3", gamma=0.95, alpha=0
                                    epsilon_start=1.0, epsilon_decay=1000,
                                    max_steps=200, render_mode="rgb_array"
```

```
In [37]:  %%time
          my_dqn_agent_controlled_training_rewards = my_dqn_agent_controlled.train(

          Non-trained Policy Network Model with randomly initialized values:
          State 000, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.04 +0.07]
          State 001, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
          State 002, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
          State 003, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
          State 004, Action D, Q-values: [+0.02 +0.05 -0.07 +0.06 -0.04 +0.07]
          State 005, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
          State 006, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.08]
          State 007, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
          State 008, Action N, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.06]
          State 009, Action D, Q-values: [+0.02 +0.07 -0.05 +0.07 -0.03 +0.08]
          State 010, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.02 +0.08]
          State 011, Action N, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.06]
          State 012, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
          State 013, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.04 +0.07]
          State 014, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
          State 015, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
          State 016, Action D, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.03 +0.08]
```

```
State 017, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 018, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 019, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 020, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 021, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.02 +0.08]
State 022, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 023, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 024, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 025, Action D, Q-values: [+0.02 +0.05 -0.05 +0.06 -0.03 +0.07]
State 026, Action D, Q-values: [+0.03 +0.07 -0.05 +0.06 -0.03 +0.08]
State 027, Action D, Q-values: [+0.02 +0.07 -0.05 +0.05 -0.03 +0.07]
State 028, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 029, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.07]
State 030, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 031, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.02 +0.07]
State 032, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.03 +0.08]
State 033, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 034, Action D, Q-values: [+0.02 +0.05 -0.07 +0.06 -0.03 +0.07]
State 035, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 036, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 037, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 038, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 039, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.04 +0.07]
State 040, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 041, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.04 +0.07]
State 042, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.04 +0.08]
State 043, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 044, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 045, Action D, Q-values: [+0.03 +0.06 -0.07 +0.06 -0.03 +0.07]
State 046, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 047, Action N, Q-values: [+0.03 +0.07 -0.07 +0.06 -0.03 +0.07]
State 048, Action D, Q-values: [+0.03 +0.07 -0.05 +0.07 -0.03 +0.08]
State 049, Action W, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 050, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 051, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 052, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 053, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 054, Action W, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 055, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 056, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 057, Action D, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.03 +0.07]
State 058, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.08]
State 059, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.03 +0.07]
State 060, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 061, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.08]
State 062, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 063, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 064, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.02 +0.07]
State 065, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 066, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 067, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 068, Action D, Q-values: [+0.02 +0.07 -0.05 +0.07 -0.03 +0.07]
State 069, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.04 +0.07]
State 070, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 071, Action W, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.06]
State 072, Action D, Q-values: [+0.03 +0.07 -0.07 +0.06 -0.04 +0.07]
State 073, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
```

```
State 074, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.02 +0.06]
State 075, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 076, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 077, Action W, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.07]
State 078, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.07]
State 079, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 080, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 081, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.08]
State 082, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.02 +0.07]
State 083, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 084, Action N, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 085, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.07]
State 086, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.02 +0.07]
State 087, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 088, Action N, Q-values: [+0.04 +0.07 -0.06 +0.07 -0.03 +0.06]
State 089, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 090, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 091, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.08]
State 092, Action D, Q-values: [+0.02 +0.05 -0.06 +0.06 -0.04 +0.08]
State 093, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 094, Action N, Q-values: [+0.03 +0.06 -0.07 +0.06 -0.03 +0.06]
State 095, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 096, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 097, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.04 +0.07]
State 098, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.08]
State 099, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 100, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 101, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 102, Action D, Q-values: [+0.02 +0.07 -0.05 +0.07 -0.02 +0.07]
State 103, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 104, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.02 +0.08]
State 105, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.08]
State 106, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 107, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.08]
State 108, Action D, Q-values: [+0.02 +0.06 -0.06 +0.05 -0.03 +0.07]
State 109, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 110, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.08]
State 111, Action N, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.04 +0.06]
State 112, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 113, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.07]
State 114, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.04 +0.08]
State 115, Action D, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.03 +0.07]
State 116, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 117, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.02 +0.07]
State 118, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.03 +0.08]
State 119, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.02 +0.08]
State 120, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.03 +0.07]
State 121, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.03 +0.08]
State 122, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 123, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.04 +0.07]
State 124, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 125, Action D, Q-values: [+0.02 +0.06 -0.05 +0.07 -0.02 +0.08]
State 126, Action W, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 127, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 128, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 129, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 130, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
```

```
State 131, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 132, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 133, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 134, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.07]
State 135, Action D, Q-values: [+0.02 +0.05 -0.06 +0.06 -0.04 +0.07]
State 136, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.04 +0.07]
State 137, Action N, Q-values: [+0.03 +0.07 -0.05 +0.07 -0.02 +0.07]
State 138, Action N, Q-values: [+0.04 +0.08 -0.06 +0.07 -0.03 +0.07]
State 139, Action D, Q-values: [+0.02 +0.05 -0.06 +0.06 -0.03 +0.07]
State 140, Action D, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.03 +0.08]
State 141, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.02 +0.07]
State 142, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 143, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 144, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 145, Action D, Q-values: [+0.03 +0.06 -0.07 +0.06 -0.04 +0.07]
State 146, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.04 +0.07]
State 147, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 148, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 149, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 150, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 151, Action D, Q-values: [+0.02 +0.06 -0.05 +0.07 -0.03 +0.07]
State 152, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.08]
State 153, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.08]
State 154, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 155, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 156, Action D, Q-values: [+0.03 +0.06 -0.07 +0.05 -0.04 +0.07]
State 157, Action D, Q-values: [+0.03 +0.06 -0.07 +0.06 -0.03 +0.07]
State 158, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 159, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.02 +0.08]
State 160, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 161, Action D, Q-values: [+0.03 +0.07 -0.07 +0.06 -0.03 +0.07]
State 162, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 163, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 164, Action D, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.03 +0.07]
State 165, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 166, Action W, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.04 +0.06]
State 167, Action W, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.04 +0.06]
State 168, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.04 +0.07]
State 169, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.08]
State 170, Action D, Q-values: [+0.02 +0.06 -0.07 +0.05 -0.03 +0.06]
State 171, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 172, Action D, Q-values: [+0.02 +0.06 -0.06 +0.05 -0.03 +0.07]
State 173, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.02 +0.07]
State 174, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 175, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 176, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 177, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 178, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 179, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.03 +0.07]
State 180, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.08]
State 181, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.08]
State 182, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 183, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.08]
State 184, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.08]
State 185, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.07]
State 186, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.06]
State 187, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.07]
```

```
State 188, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 189, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.04 +0.07]
State 190, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 191, Action W, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.07]
State 192, Action D, Q-values: [+0.01 +0.07 -0.06 +0.06 -0.03 +0.08]
State 193, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 194, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 195, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.08]
State 196, Action N, Q-values: [+0.04 +0.07 -0.06 +0.07 -0.03 +0.07]
State 197, Action W, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.07]
State 198, Action W, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.04 +0.07]
State 199, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.04 +0.07]
State 200, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.07]
State 201, Action D, Q-values: [+0.01 +0.07 -0.07 +0.05 -0.03 +0.07]
State 202, Action D, Q-values: [+0.03 +0.07 -0.07 +0.06 -0.03 +0.07]
State 203, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.02 +0.07]
State 204, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 205, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.08]
State 206, Action D, Q-values: [+0.03 +0.05 -0.07 +0.07 -0.03 +0.07]
State 207, Action D, Q-values: [+0.03 +0.07 -0.07 +0.06 -0.04 +0.07]
State 208, Action N, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.04 +0.07]
State 209, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.08]
State 210, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.07]
State 211, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.04 +0.07]
State 212, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 213, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 214, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.08]
State 215, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 216, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 217, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 218, Action D, Q-values: [+0.02 +0.06 -0.06 +0.05 -0.03 +0.07]
State 219, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.07]
State 220, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.02 +0.07]
State 221, Action W, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.04 +0.06]
State 222, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.02 +0.07]
State 223, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 224, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 225, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 226, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 227, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 228, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 229, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.03 +0.08]
State 230, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 231, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 232, Action D, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.04 +0.07]
State 233, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 234, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 235, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 236, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 237, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 238, Action D, Q-values: [+0.03 +0.07 -0.05 +0.07 -0.02 +0.08]
State 239, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.08]
State 240, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.03 +0.08]
State 241, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 242, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 243, Action W, Q-values: [+0.03 +0.05 -0.06 +0.07 -0.03 +0.07]
State 244, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.08]
```

```
State 245, Action D, Q-values: [+0.01 +0.05 −0.06 +0.06 −0.03 +0.07]
State 246, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.08]
State 247, Action D, Q-values: [+0.02 +0.07 −0.07 +0.06 −0.04 +0.07]
State 248, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.07]
State 249, Action W, Q-values: [+0.03 +0.07 −0.06 +0.07 −0.04 +0.06]
State 250, Action W, Q-values: [+0.03 +0.07 −0.06 +0.07 −0.02 +0.07]
State 251, Action D, Q-values: [+0.02 +0.07 −0.06 +0.07 −0.02 +0.07]
State 252, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.08]
State 253, Action D, Q-values: [+0.03 +0.06 −0.05 +0.06 −0.03 +0.08]
State 254, Action D, Q-values: [+0.02 +0.07 −0.06 +0.07 −0.03 +0.07]
State 255, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.02 +0.08]
State 256, Action D, Q-values: [+0.02 +0.07 −0.05 +0.06 −0.02 +0.07]
State 257, Action D, Q-values: [+0.02 +0.07 −0.06 +0.07 −0.04 +0.07]
State 258, Action D, Q-values: [+0.02 +0.06 −0.05 +0.06 −0.03 +0.07]
State 259, Action D, Q-values: [+0.03 +0.06 −0.06 +0.06 −0.02 +0.07]
State 260, Action D, Q-values: [+0.03 +0.06 −0.05 +0.07 −0.03 +0.07]
State 261, Action D, Q-values: [+0.04 +0.06 −0.06 +0.06 −0.03 +0.07]
State 262, Action D, Q-values: [+0.02 +0.06 −0.05 +0.07 −0.03 +0.08]
State 263, Action D, Q-values: [+0.02 +0.07 −0.06 +0.06 −0.02 +0.08]
State 264, Action D, Q-values: [+0.03 +0.07 −0.06 +0.06 −0.04 +0.07]
State 265, Action D, Q-values: [+0.03 +0.06 −0.05 +0.06 −0.03 +0.07]
State 266, Action D, Q-values: [+0.03 +0.06 −0.06 +0.07 −0.04 +0.07]
State 267, Action W, Q-values: [+0.02 +0.06 −0.07 +0.07 −0.03 +0.07]
State 268, Action D, Q-values: [+0.03 +0.05 −0.07 +0.07 −0.04 +0.07]
State 269, Action D, Q-values: [+0.02 +0.07 −0.05 +0.06 −0.04 +0.07]
State 270, Action D, Q-values: [+0.02 +0.06 −0.05 +0.06 −0.03 +0.07]
State 271, Action D, Q-values: [+0.03 +0.06 −0.06 +0.06 −0.03 +0.07]
State 272, Action D, Q-values: [+0.01 +0.07 −0.05 +0.05 −0.02 +0.07]
State 273, Action D, Q-values: [+0.03 +0.07 −0.05 +0.06 −0.03 +0.07]
State 274, Action D, Q-values: [+0.03 +0.07 −0.06 +0.06 −0.03 +0.07]
State 275, Action D, Q-values: [+0.03 +0.06 −0.07 +0.06 −0.03 +0.08]
State 276, Action N, Q-values: [+0.03 +0.07 −0.06 +0.06 −0.03 +0.07]
State 277, Action D, Q-values: [+0.01 +0.06 −0.06 +0.06 −0.03 +0.07]
State 278, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.07]
State 279, Action D, Q-values: [+0.02 +0.06 −0.05 +0.06 −0.03 +0.07]
State 280, Action D, Q-values: [+0.02 +0.06 −0.06 +0.07 −0.03 +0.07]
State 281, Action D, Q-values: [+0.02 +0.07 −0.06 +0.06 −0.03 +0.07]
State 282, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.07]
State 283, Action D, Q-values: [+0.02 +0.07 −0.06 +0.06 −0.03 +0.07]
State 284, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.07]
State 285, Action D, Q-values: [+0.03 +0.06 −0.06 +0.07 −0.03 +0.07]
State 286, Action D, Q-values: [+0.03 +0.07 −0.06 +0.06 −0.04 +0.08]
State 287, Action D, Q-values: [+0.02 +0.07 −0.05 +0.05 −0.03 +0.07]
State 288, Action D, Q-values: [+0.02 +0.06 −0.06 +0.07 −0.03 +0.08]
State 289, Action D, Q-values: [+0.02 +0.07 −0.06 +0.06 −0.04 +0.07]
State 290, Action N, Q-values: [+0.02 +0.07 −0.06 +0.06 −0.03 +0.07]
State 291, Action D, Q-values: [+0.02 +0.06 −0.06 +0.07 −0.03 +0.08]
State 292, Action D, Q-values: [+0.02 +0.07 −0.06 +0.06 −0.03 +0.07]
State 293, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.06]
State 294, Action D, Q-values: [+0.03 +0.07 −0.06 +0.07 −0.03 +0.08]
State 295, Action D, Q-values: [+0.03 +0.06 −0.06 +0.07 −0.03 +0.07]
State 296, Action W, Q-values: [+0.03 +0.07 −0.06 +0.08 −0.03 +0.07]
State 297, Action W, Q-values: [+0.02 +0.06 −0.07 +0.06 −0.04 +0.06]
State 298, Action D, Q-values: [+0.03 +0.07 −0.06 +0.07 −0.03 +0.07]
State 299, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.07]
State 300, Action D, Q-values: [+0.02 +0.06 −0.06 +0.06 −0.03 +0.07]
State 301, Action D, Q-values: [+0.02 +0.06 −0.05 +0.06 −0.03 +0.08]
```

```
State 302, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 303, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 304, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 305, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 306, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 307, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 308, Action N, Q-values: [+0.03 +0.07 -0.07 +0.06 -0.03 +0.07]
State 309, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.07]
State 310, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.03 +0.07]
State 311, Action N, Q-values: [+0.03 +0.08 -0.07 +0.07 -0.04 +0.07]
State 312, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 313, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 314, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.07]
State 315, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 316, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 317, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 318, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 319, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 320, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.04 +0.07]
State 321, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.04 +0.07]
State 322, Action D, Q-values: [+0.02 +0.07 -0.05 +0.07 -0.02 +0.07]
State 323, Action N, Q-values: [+0.03 +0.07 -0.05 +0.06 -0.03 +0.07]
State 324, Action D, Q-values: [+0.03 +0.07 -0.05 +0.07 -0.02 +0.08]
State 325, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.03 +0.07]
State 326, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 327, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 328, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 329, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 330, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.04 +0.07]
State 331, Action W, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 332, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 333, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 334, Action W, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 335, Action N, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 336, Action D, Q-values: [+0.02 +0.05 -0.06 +0.06 -0.02 +0.07]
State 337, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 338, Action D, Q-values: [+0.02 +0.06 -0.06 +0.05 -0.03 +0.07]
State 339, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.02 +0.08]
State 340, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 341, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 342, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.04 +0.07]
State 343, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.06]
State 344, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 345, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 346, Action D, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.04 +0.07]
State 347, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 348, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 349, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 350, Action W, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 351, Action D, Q-values: [+0.01 +0.07 -0.05 +0.06 -0.03 +0.08]
State 352, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 353, Action D, Q-values: [+0.02 +0.05 -0.07 +0.06 -0.03 +0.07]
State 354, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.07]
State 355, Action N, Q-values: [+0.03 +0.07 -0.05 +0.06 -0.02 +0.06]
State 356, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 357, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.02 +0.07]
State 358, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
```

```
State 359, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 360, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 361, Action D, Q-values: [+0.03 +0.06 -0.06 +0.05 -0.04 +0.07]
State 362, Action D, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.03 +0.07]
State 363, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 364, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 365, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 366, Action W, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 367, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 368, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 369, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.06]
State 370, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.02 +0.07]
State 371, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.03 +0.08]
State 372, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.02 +0.08]
State 373, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 374, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 375, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.02 +0.07]
State 376, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.03 +0.07]
State 377, Action D, Q-values: [+0.03 +0.07 -0.05 +0.07 -0.02 +0.08]
State 378, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 379, Action D, Q-values: [+0.02 +0.07 -0.06 +0.07 -0.03 +0.07]
State 380, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.03 +0.08]
State 381, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 382, Action N, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 383, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 384, Action W, Q-values: [+0.03 +0.07 -0.06 +0.08 -0.03 +0.07]
State 385, Action D, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.04 +0.07]
State 386, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 387, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.02 +0.07]
State 388, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.03 +0.07]
State 389, Action D, Q-values: [+0.02 +0.07 -0.07 +0.07 -0.03 +0.07]
State 390, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.04 +0.07]
State 391, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.04 +0.07]
State 392, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.08]
State 393, Action N, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.04 +0.07]
State 394, Action W, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 395, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 396, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 397, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.02 +0.08]
State 398, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 399, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 400, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.04 +0.07]
State 401, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 402, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.08]
State 403, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 404, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.07]
State 405, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 406, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.06]
State 407, Action W, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.03 +0.06]
State 408, Action D, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.08]
State 409, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 410, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 411, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 412, Action D, Q-values: [+0.02 +0.07 -0.05 +0.07 -0.03 +0.07]
State 413, Action N, Q-values: [+0.03 +0.07 -0.05 +0.06 -0.02 +0.07]
State 414, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 415, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
```

```
State 416, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 417, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.03 +0.07]
State 418, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 419, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.02 +0.07]
State 420, Action W, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.06]
State 421, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.08]
State 422, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.03 +0.07]
State 423, Action D, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.02 +0.07]
State 424, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
State 425, Action D, Q-values: [+0.02 +0.06 -0.07 +0.05 -0.04 +0.07]
State 426, Action D, Q-values: [+0.01 +0.06 -0.07 +0.06 -0.03 +0.08]
State 427, Action W, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.03 +0.07]
State 428, Action D, Q-values: [+0.03 +0.05 -0.06 +0.07 -0.03 +0.08]
State 429, Action D, Q-values: [+0.02 +0.05 -0.06 +0.06 -0.03 +0.08]
State 430, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.04 +0.08]
State 431, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.02 +0.08]
State 432, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 433, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 434, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.04 +0.08]
State 435, Action D, Q-values: [+0.03 +0.07 -0.05 +0.06 -0.02 +0.07]
State 436, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 437, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.03 +0.08]
State 438, Action D, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.03 +0.07]
State 439, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 440, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.02 +0.08]
State 441, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.04 +0.07]
State 442, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.08]
State 443, Action W, Q-values: [+0.03 +0.06 -0.05 +0.07 -0.04 +0.07]
State 444, Action D, Q-values: [+0.02 +0.05 -0.06 +0.06 -0.03 +0.08]
State 445, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 446, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 447, Action D, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.03 +0.08]
State 448, Action W, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 449, Action W, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.06]
State 450, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 451, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 452, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.03 +0.07]
State 453, Action W, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.04 +0.07]
State 454, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 455, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 456, Action D, Q-values: [+0.03 +0.06 -0.06 +0.05 -0.03 +0.07]
State 457, Action D, Q-values: [+0.02 +0.06 -0.05 +0.07 -0.03 +0.08]
State 458, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.07]
State 459, Action D, Q-values: [+0.02 +0.06 -0.05 +0.07 -0.02 +0.08]
State 460, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 461, Action D, Q-values: [+0.03 +0.07 -0.07 +0.05 -0.04 +0.07]
State 462, Action N, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.06]
State 463, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.02 +0.07]
State 464, Action D, Q-values: [+0.02 +0.06 -0.07 +0.07 -0.03 +0.07]
State 465, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 466, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 467, Action D, Q-values: [+0.02 +0.06 -0.06 +0.05 -0.02 +0.07]
State 468, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.04 +0.08]
State 469, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 470, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 471, Action W, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.07]
State 472, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.03 +0.07]
```

```
State 473, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.08]
State 474, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.04 +0.07]
State 475, Action D, Q-values: [+0.01 +0.05 -0.06 +0.06 -0.04 +0.07]
State 476, Action D, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.04 +0.07]
State 477, Action D, Q-values: [+0.02 +0.06 -0.05 +0.06 -0.04 +0.07]
State 478, Action D, Q-values: [+0.02 +0.06 -0.06 +0.07 -0.02 +0.07]
State 479, Action W, Q-values: [+0.03 +0.06 -0.06 +0.07 -0.03 +0.07]
State 480, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.03 +0.07]
State 481, Action D, Q-values: [+0.03 +0.07 -0.06 +0.07 -0.03 +0.07]
State 482, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.06]
State 483, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 484, Action D, Q-values: [+0.03 +0.07 -0.06 +0.06 -0.03 +0.07]
State 485, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.08]
State 486, Action N, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.03 +0.06]
State 487, Action D, Q-values: [+0.02 +0.07 -0.07 +0.06 -0.03 +0.07]
State 488, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.07]
State 489, Action D, Q-values: [+0.02 +0.07 -0.05 +0.06 -0.03 +0.07]
State 490, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 491, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.03 +0.07]
State 492, Action D, Q-values: [+0.03 +0.07 -0.05 +0.07 -0.03 +0.08]
State 493, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.07]
State 494, Action D, Q-values: [+0.02 +0.06 -0.07 +0.06 -0.04 +0.08]
State 495, Action D, Q-values: [+0.03 +0.06 -0.07 +0.07 -0.04 +0.07]
State 496, Action D, Q-values: [+0.02 +0.07 -0.06 +0.06 -0.03 +0.08]
State 497, Action D, Q-values: [+0.03 +0.06 -0.05 +0.06 -0.02 +0.08]
State 498, Action D, Q-values: [+0.02 +0.06 -0.06 +0.06 -0.03 +0.06]
State 499, Action D, Q-values: [+0.03 +0.06 -0.06 +0.06 -0.03 +0.07]
Episode 100/100000, Avg Reward: -777.86, Epsilon: 0.99
Episode 200/100000, Avg Reward: -783.83, Epsilon: 0.99
Episode 300/100000, Avg Reward: -759.55, Epsilon: 0.98
Episode 400/100000, Avg Reward: -767.31, Epsilon: 0.97
Episode 500/100000, Avg Reward: -752.85, Epsilon: 0.97
Episode 600/100000, Avg Reward: -759.51, Epsilon: 0.96
Episode 700/100000, Avg Reward: -724.39, Epsilon: 0.95
Episode 800/100000, Avg Reward: -724.42, Epsilon: 0.95
Episode 900/100000, Avg Reward: -737.55, Epsilon: 0.94
Episode 1000/100000, Avg Reward: -729.67, Epsilon: 0.93
Episode 1100/100000, Avg Reward: -743.21, Epsilon: 0.93
Episode 1200/100000, Avg Reward: -732.25, Epsilon: 0.92
Episode 1300/100000, Avg Reward: -720.31, Epsilon: 0.91
Episode 1400/100000, Avg Reward: -716.30, Epsilon: 0.91
Episode 1500/100000, Avg Reward: -716.35, Epsilon: 0.90
Episode 1600/100000, Avg Reward: -693.17, Epsilon: 0.90
Episode 1700/100000, Avg Reward: -707.46, Epsilon: 0.89
Episode 1800/100000, Avg Reward: -703.07, Epsilon: 0.88
Episode 1900/100000, Avg Reward: -708.82, Epsilon: 0.88
Episode 2000/100000, Avg Reward: -696.71, Epsilon: 0.87
Episode 2100/100000, Avg Reward: -684.13, Epsilon: 0.86
Episode 2200/100000, Avg Reward: -686.57, Epsilon: 0.86
Episode 2300/100000, Avg Reward: -693.39, Epsilon: 0.85
Episode 2400/100000, Avg Reward: -675.96, Epsilon: 0.85
Episode 2500/100000, Avg Reward: -676.33, Epsilon: 0.84
Episode 2600/100000, Avg Reward: -684.13, Epsilon: 0.84
Episode 2700/100000, Avg Reward: -680.59, Epsilon: 0.83
Episode 2800/100000, Avg Reward: -655.38, Epsilon: 0.82
Episode 2900/100000, Avg Reward: -675.52, Epsilon: 0.82
Episode 3000/100000, Avg Reward: -684.70, Epsilon: 0.81
```

```
Episode 3100/100000, Avg Reward: -662.74, Epsilon: 0.81
Episode 3200/100000, Avg Reward: -642.81, Epsilon: 0.80
Episode 3300/100000, Avg Reward: -649.26, Epsilon: 0.80
Episode 3400/100000, Avg Reward: -656.04, Epsilon: 0.79
Episode 3500/100000, Avg Reward: -651.58, Epsilon: 0.79
Episode 3600/100000, Avg Reward: -639.57, Epsilon: 0.78
Episode 3700/100000, Avg Reward: -647.71, Epsilon: 0.77
Episode 3800/100000, Avg Reward: -647.15, Epsilon: 0.77
Episode 3900/100000, Avg Reward: -635.78, Epsilon: 0.76
Episode 4000/100000, Avg Reward: -639.53, Epsilon: 0.76
Episode 4100/100000, Avg Reward: -632.00, Epsilon: 0.75
Episode 4200/100000, Avg Reward: -627.99, Epsilon: 0.75
Episode 4300/100000, Avg Reward: -642.80, Epsilon: 0.74
Episode 4400/100000, Avg Reward: -626.15, Epsilon: 0.74
Episode 4500/100000, Avg Reward: -615.08, Epsilon: 0.73
Episode 4600/100000, Avg Reward: -606.62, Epsilon: 0.73
Episode 4700/100000, Avg Reward: -616.86, Epsilon: 0.72
Episode 4800/100000, Avg Reward: -629.30, Epsilon: 0.72
Episode 4900/100000, Avg Reward: -607.84, Epsilon: 0.71
Episode 5000/100000, Avg Reward: -612.70, Epsilon: 0.71
Episode 5100/100000, Avg Reward: -605.97, Epsilon: 0.70
Episode 5200/100000, Avg Reward: -603.15, Epsilon: 0.70
Episode 5300/100000, Avg Reward: -593.66, Epsilon: 0.69
Episode 5400/100000, Avg Reward: -593.47, Epsilon: 0.69
Episode 5500/100000, Avg Reward: -602.47, Epsilon: 0.68
Episode 5600/100000, Avg Reward: -585.62, Epsilon: 0.68
Episode 5700/100000, Avg Reward: -588.35, Epsilon: 0.67
Episode 5800/100000, Avg Reward: -583.96, Epsilon: 0.67
Episode 5900/100000, Avg Reward: -586.61, Epsilon: 0.67
Episode 6000/100000, Avg Reward: -598.79, Epsilon: 0.66
Episode 6100/100000, Avg Reward: -566.23, Epsilon: 0.66
Episode 6200/100000, Avg Reward: -568.05, Epsilon: 0.65
Episode 6300/100000, Avg Reward: -571.34, Epsilon: 0.65
Episode 6400/100000, Avg Reward: -562.20, Epsilon: 0.64
Episode 6500/100000, Avg Reward: -567.15, Epsilon: 0.64
Episode 6600/100000, Avg Reward: -565.01, Epsilon: 0.63
Episode 6700/100000, Avg Reward: -556.25, Epsilon: 0.63
Episode 6800/100000, Avg Reward: -554.95, Epsilon: 0.63
Episode 6900/100000, Avg Reward: -543.45, Epsilon: 0.62
Episode 7000/100000, Avg Reward: -558.69, Epsilon: 0.62
Episode 7100/100000, Avg Reward: -549.30, Epsilon: 0.61
Episode 7200/100000, Avg Reward: -537.60, Epsilon: 0.61
Episode 7300/100000, Avg Reward: -549.93, Epsilon: 0.60
Episode 7400/100000, Avg Reward: -545.71, Epsilon: 0.60
Episode 7500/100000, Avg Reward: -534.16, Epsilon: 0.60
Episode 7600/100000, Avg Reward: -544.79, Epsilon: 0.59
Episode 7700/100000, Avg Reward: -541.53, Epsilon: 0.59
Episode 7800/100000, Avg Reward: -533.12, Epsilon: 0.58
Episode 7900/100000, Avg Reward: -535.41, Epsilon: 0.58
Episode 8000/100000, Avg Reward: -527.28, Epsilon: 0.58
Episode 8100/100000, Avg Reward: -525.54, Epsilon: 0.57
Episode 8200/100000, Avg Reward: -521.03, Epsilon: 0.57
Episode 8300/100000, Avg Reward: -519.07, Epsilon: 0.56
Episode 8400/100000, Avg Reward: -519.91, Epsilon: 0.56
Episode 8500/100000, Avg Reward: -519.69, Epsilon: 0.56
Episode 8600/100000, Avg Reward: -515.63, Epsilon: 0.55
Episode 8700/100000, Avg Reward: -504.58, Epsilon: 0.55
```

```
Episode 8800/100000, Avg Reward: -510.82, Epsilon: 0.54
Episode 8900/100000, Avg Reward: -512.23, Epsilon: 0.54
Episode 9000/100000, Avg Reward: -510.68, Epsilon: 0.54
Episode 9100/100000, Avg Reward: -512.12, Epsilon: 0.53
Episode 9200/100000, Avg Reward: -503.47, Epsilon: 0.53
Episode 9300/100000, Avg Reward: -493.89, Epsilon: 0.53
Episode 9400/100000, Avg Reward: -487.88, Epsilon: 0.52
Episode 9500/100000, Avg Reward: -503.03, Epsilon: 0.52
Episode 9600/100000, Avg Reward: -501.50, Epsilon: 0.52
Episode 9700/100000, Avg Reward: -487.10, Epsilon: 0.51
Episode 9800/100000, Avg Reward: -493.81, Epsilon: 0.51
Episode 9900/100000, Avg Reward: -494.96, Epsilon: 0.50
Episode 10000/100000, Avg Reward: -488.09, Epsilon: 0.50
Episode 10100/100000, Avg Reward: -477.44, Epsilon: 0.50
Episode 10200/100000, Avg Reward: -490.34, Epsilon: 0.49
Episode 10300/100000, Avg Reward: -468.14, Epsilon: 0.49
Episode 10400/100000, Avg Reward: -471.71, Epsilon: 0.49
Episode 10500/100000, Avg Reward: -472.91, Epsilon: 0.48
Episode 10600/100000, Avg Reward: -481.54, Epsilon: 0.48
Episode 10700/100000, Avg Reward: -477.71, Epsilon: 0.48
Episode 10800/100000, Avg Reward: -465.95, Epsilon: 0.47
Episode 10900/100000, Avg Reward: -477.03, Epsilon: 0.47
Episode 11000/100000, Avg Reward: -471.35, Epsilon: 0.47
Episode 11100/100000, Avg Reward: -472.76, Epsilon: 0.46
Episode 11200/100000, Avg Reward: -465.10, Epsilon: 0.46
Episode 11300/100000, Avg Reward: -462.08, Epsilon: 0.46
Episode 11400/100000, Avg Reward: -470.18, Epsilon: 0.45
Episode 11500/100000, Avg Reward: -453.64, Epsilon: 0.45
Episode 11600/100000, Avg Reward: -457.90, Epsilon: 0.45
Episode 11700/100000, Avg Reward: -470.81, Epsilon: 0.45
Episode 11800/100000, Avg Reward: -459.60, Epsilon: 0.44
Episode 11900/100000, Avg Reward: -459.47, Epsilon: 0.44
Episode 12000/100000, Avg Reward: -438.37, Epsilon: 0.44
Episode 12100/100000, Avg Reward: -453.34, Epsilon: 0.43
Episode 12200/100000, Avg Reward: -440.05, Epsilon: 0.43
Episode 12300/100000, Avg Reward: -449.48, Epsilon: 0.43
Episode 12400/100000, Avg Reward: -451.73, Epsilon: 0.42
Episode 12500/100000, Avg Reward: -443.58, Epsilon: 0.42
Episode 12600/100000, Avg Reward: -443.00, Epsilon: 0.42
Episode 12700/100000, Avg Reward: -456.32, Epsilon: 0.42
Episode 12800/100000, Avg Reward: -436.82, Epsilon: 0.41
Episode 12900/100000, Avg Reward: -441.47, Epsilon: 0.41
Episode 13000/100000, Avg Reward: -432.75, Epsilon: 0.41
Episode 13100/100000, Avg Reward: -431.03, Epsilon: 0.40
Episode 13200/100000, Avg Reward: -435.80, Epsilon: 0.40
Episode 13300/100000, Avg Reward: -424.37, Epsilon: 0.40
Episode 13400/100000, Avg Reward: -426.07, Epsilon: 0.40
Episode 13500/100000, Avg Reward: -432.08, Epsilon: 0.39
Episode 13600/100000, Avg Reward: -434.18, Epsilon: 0.39
Episode 13700/100000, Avg Reward: -421.94, Epsilon: 0.39
Episode 13800/100000, Avg Reward: -424.44, Epsilon: 0.39
Episode 13900/100000, Avg Reward: -421.95, Epsilon: 0.38
Episode 14000/100000, Avg Reward: -416.90, Epsilon: 0.38
Episode 14100/100000, Avg Reward: -422.39, Epsilon: 0.38
Episode 14200/100000, Avg Reward: -415.29, Epsilon: 0.37
Episode 14300/100000, Avg Reward: -410.70, Epsilon: 0.37
Episode 14400/100000, Avg Reward: -417.35, Epsilon: 0.37
```

```
Episode 14500/100000, Avg Reward: -410.96, Epsilon: 0.37
Episode 14600/100000, Avg Reward: -408.89, Epsilon: 0.36
Episode 14700/100000, Avg Reward: -405.83, Epsilon: 0.36
Episode 14800/100000, Avg Reward: -405.16, Epsilon: 0.36
Episode 14900/100000, Avg Reward: -406.32, Epsilon: 0.36
Episode 15000/100000, Avg Reward: -408.71, Epsilon: 0.35
Episode 15100/100000, Avg Reward: -404.75, Epsilon: 0.35
Episode 15200/100000, Avg Reward: -408.44, Epsilon: 0.35
Episode 15300/100000, Avg Reward: -397.16, Epsilon: 0.35
Episode 15400/100000, Avg Reward: -402.40, Epsilon: 0.35
Episode 15500/100000, Avg Reward: -392.35, Epsilon: 0.34
Episode 15600/100000, Avg Reward: -392.06, Epsilon: 0.34
Episode 15700/100000, Avg Reward: -401.98, Epsilon: 0.34
Episode 15800/100000, Avg Reward: -387.53, Epsilon: 0.34
Episode 15900/100000, Avg Reward: -397.10, Epsilon: 0.33
Episode 16000/100000, Avg Reward: -381.95, Epsilon: 0.33
Episode 16100/100000, Avg Reward: -393.23, Epsilon: 0.33
Episode 16200/100000, Avg Reward: -383.87, Epsilon: 0.33
Episode 16300/100000, Avg Reward: -390.03, Epsilon: 0.32
Episode 16400/100000, Avg Reward: -384.41, Epsilon: 0.32
Episode 16500/100000, Avg Reward: -381.62, Epsilon: 0.32
Episode 16600/100000, Avg Reward: -387.74, Epsilon: 0.32
Episode 16700/100000, Avg Reward: -385.94, Epsilon: 0.32
Episode 16800/100000, Avg Reward: -381.62, Epsilon: 0.31
Episode 16900/100000, Avg Reward: -374.15, Epsilon: 0.31
Episode 17000/100000, Avg Reward: -372.72, Epsilon: 0.31
Episode 17100/100000, Avg Reward: -377.84, Epsilon: 0.31
Episode 17200/100000, Avg Reward: -377.12, Epsilon: 0.30
Episode 17300/100000, Avg Reward: -373.34, Epsilon: 0.30
Episode 17400/100000, Avg Reward: -366.32, Epsilon: 0.30
Episode 17500/100000, Avg Reward: -377.30, Epsilon: 0.30
Episode 17600/100000, Avg Reward: -358.01, Epsilon: 0.30
Episode 17700/100000, Avg Reward: -365.24, Epsilon: 0.29
Episode 17800/100000, Avg Reward: -368.66, Epsilon: 0.29
Episode 17900/100000, Avg Reward: -371.90, Epsilon: 0.29
Episode 18000/100000, Avg Reward: -371.45, Epsilon: 0.29
Episode 18100/100000, Avg Reward: -363.17, Epsilon: 0.29
Episode 18200/100000, Avg Reward: -370.91, Epsilon: 0.28
Episode 18300/100000, Avg Reward: -358.40, Epsilon: 0.28
Episode 18400/100000, Avg Reward: -363.44, Epsilon: 0.28
Episode 18500/100000, Avg Reward: -359.75, Epsilon: 0.28
Episode 18600/100000, Avg Reward: -360.02, Epsilon: 0.28
Episode 18700/100000, Avg Reward: -352.73, Epsilon: 0.27
Episode 18800/100000, Avg Reward: -353.72, Epsilon: 0.27
Episode 18900/100000, Avg Reward: -356.24, Epsilon: 0.27
Episode 19000/100000, Avg Reward: -359.57, Epsilon: 0.27
Episode 19100/100000, Avg Reward: -349.13, Epsilon: 0.27
Episode 19200/100000, Avg Reward: -353.81, Epsilon: 0.27
Episode 19300/100000, Avg Reward: -348.50, Epsilon: 0.26
Episode 19400/100000, Avg Reward: -349.85, Epsilon: 0.26
Episode 19500/100000, Avg Reward: -354.98, Epsilon: 0.26
Episode 19600/100000, Avg Reward: -344.18, Epsilon: 0.26
Episode 19700/100000, Avg Reward: -344.81, Epsilon: 0.26
Episode 19800/100000, Avg Reward: -347.60, Epsilon: 0.25
Episode 19900/100000, Avg Reward: -336.71, Epsilon: 0.25
Episode 20000/100000, Avg Reward: -339.77, Epsilon: 0.25
Episode 20100/100000, Avg Reward: -339.95, Epsilon: 0.25
```

```
Episode 20200/100000, Avg Reward: -339.65, Epsilon: 0.25
Episode 20300/100000, Avg Reward: -343.19, Epsilon: 0.25
Episode 20400/100000, Avg Reward: -344.18, Epsilon: 0.24
Episode 20500/100000, Avg Reward: -338.96, Epsilon: 0.24
Episode 20600/100000, Avg Reward: -337.52, Epsilon: 0.24
Episode 20700/100000, Avg Reward: -332.48, Epsilon: 0.24
Episode 20800/100000, Avg Reward: -336.89, Epsilon: 0.24
Episode 20900/100000, Avg Reward: -337.70, Epsilon: 0.24
Episode 21000/100000, Avg Reward: -334.37, Epsilon: 0.23
Episode 21100/100000, Avg Reward: -331.49, Epsilon: 0.23
Episode 21200/100000, Avg Reward: -341.48, Epsilon: 0.23
Episode 21300/100000, Avg Reward: -337.34, Epsilon: 0.23
Episode 21400/100000, Avg Reward: -336.53, Epsilon: 0.23
Episode 21500/100000, Avg Reward: -330.68, Epsilon: 0.23
Episode 21600/100000, Avg Reward: -326.81, Epsilon: 0.22
Episode 21700/100000, Avg Reward: -334.19, Epsilon: 0.22
Episode 21800/100000, Avg Reward: -329.96, Epsilon: 0.22
Episode 21900/100000, Avg Reward: -325.82, Epsilon: 0.22
Episode 22000/100000, Avg Reward: -318.71, Epsilon: 0.22
Episode 22100/100000, Avg Reward: -324.11, Epsilon: 0.22
Episode 22200/100000, Avg Reward: -324.29, Epsilon: 0.22
Episode 22300/100000, Avg Reward: -324.56, Epsilon: 0.21
Episode 22400/100000, Avg Reward: -316.55, Epsilon: 0.21
Episode 22500/100000, Avg Reward: -320.33, Epsilon: 0.21
Episode 22600/100000, Avg Reward: -312.68, Epsilon: 0.21
Episode 22700/100000, Avg Reward: -315.56, Epsilon: 0.21
Episode 22800/100000, Avg Reward: -316.10, Epsilon: 0.21
Episode 22900/100000, Avg Reward: -316.46, Epsilon: 0.21
Episode 23000/100000, Avg Reward: -320.96, Epsilon: 0.20
Episode 23100/100000, Avg Reward: -319.52, Epsilon: 0.20
Episode 23200/100000, Avg Reward: -308.54, Epsilon: 0.20
Episode 23300/100000, Avg Reward: -312.41, Epsilon: 0.20
Episode 23400/100000, Avg Reward: -315.38, Epsilon: 0.20
Episode 23500/100000, Avg Reward: -314.12, Epsilon: 0.20
Episode 23600/100000, Avg Reward: -307.64, Epsilon: 0.20
Episode 23700/100000, Avg Reward: -303.68, Epsilon: 0.19
Episode 23800/100000, Avg Reward: -310.70, Epsilon: 0.19
Episode 23900/100000, Avg Reward: -309.17, Epsilon: 0.19
Episode 24000/100000, Avg Reward: -307.46, Epsilon: 0.19
Episode 24100/100000, Avg Reward: -311.60, Epsilon: 0.19
Episode 24200/100000, Avg Reward: -309.26, Epsilon: 0.19
Episode 24300/100000, Avg Reward: -300.89, Epsilon: 0.19
Episode 24400/100000, Avg Reward: -309.71, Epsilon: 0.19
Episode 24500/100000, Avg Reward: -300.62, Epsilon: 0.18
Episode 24600/100000, Avg Reward: -309.80, Epsilon: 0.18
Episode 24700/100000, Avg Reward: -304.13, Epsilon: 0.18
Episode 24800/100000, Avg Reward: -302.33, Epsilon: 0.18
Episode 24900/100000, Avg Reward: -304.31, Epsilon: 0.18
Episode 25000/100000, Avg Reward: -301.61, Epsilon: 0.18
Episode 25100/100000, Avg Reward: -298.01, Epsilon: 0.18
Episode 25200/100000, Avg Reward: -304.04, Epsilon: 0.18
Episode 25300/100000, Avg Reward: -301.88, Epsilon: 0.17
Episode 25400/100000, Avg Reward: -297.11, Epsilon: 0.17
Episode 25500/100000, Avg Reward: -303.23, Epsilon: 0.17
Episode 25600/100000, Avg Reward: -294.35, Epsilon: 0.17
Episode 25700/100000, Avg Reward: -294.68, Epsilon: 0.17
Episode 25800/100000, Avg Reward: -300.17, Epsilon: 0.17
```

```
Episode 25900/100000, Avg Reward: -296.75, Epsilon: 0.17
Episode 26000/100000, Avg Reward: -295.67, Epsilon: 0.17
Episode 26100/100000, Avg Reward: -293.60, Epsilon: 0.16
Episode 26200/100000, Avg Reward: -295.67, Epsilon: 0.16
Episode 26300/100000, Avg Reward: -293.33, Epsilon: 0.16
Episode 26400/100000, Avg Reward: -293.06, Epsilon: 0.16
Episode 26500/100000, Avg Reward: -294.77, Epsilon: 0.16
Episode 26600/100000, Avg Reward: -293.87, Epsilon: 0.16
Episode 26700/100000, Avg Reward: -290.54, Epsilon: 0.16
Episode 26800/100000, Avg Reward: -288.65, Epsilon: 0.16
Episode 26900/100000, Avg Reward: -287.57, Epsilon: 0.16
Episode 27000/100000, Avg Reward: -288.29, Epsilon: 0.15
Episode 27100/100000, Avg Reward: -288.65, Epsilon: 0.15
Episode 27200/100000, Avg Reward: -289.73, Epsilon: 0.15
Episode 27300/100000, Avg Reward: -287.57, Epsilon: 0.15
Episode 27400/100000, Avg Reward: -289.46, Epsilon: 0.15
Episode 27500/100000, Avg Reward: -281.72, Epsilon: 0.15
Episode 27600/100000, Avg Reward: -288.83, Epsilon: 0.15
Episode 27700/100000, Avg Reward: -285.95, Epsilon: 0.15
Episode 27800/100000, Avg Reward: -280.91, Epsilon: 0.15
Episode 27900/100000, Avg Reward: -281.99, Epsilon: 0.15
Episode 28000/100000, Avg Reward: -287.12, Epsilon: 0.14
Episode 28100/100000, Avg Reward: -282.53, Epsilon: 0.14
Episode 28200/100000, Avg Reward: -282.62, Epsilon: 0.14
Episode 28300/100000, Avg Reward: -282.53, Epsilon: 0.14
Episode 28400/100000, Avg Reward: -283.97, Epsilon: 0.14
Episode 28500/100000, Avg Reward: -277.22, Epsilon: 0.14
Episode 28600/100000, Avg Reward: -272.45, Epsilon: 0.14
Episode 28700/100000, Avg Reward: -281.45, Epsilon: 0.14
Episode 28800/100000, Avg Reward: -275.87, Epsilon: 0.14
Episode 28900/100000, Avg Reward: -277.94, Epsilon: 0.14
Episode 29000/100000, Avg Reward: -277.22, Epsilon: 0.13
Episode 29100/100000, Avg Reward: -272.81, Epsilon: 0.13
Episode 29200/100000, Avg Reward: -274.34, Epsilon: 0.13
Episode 29300/100000, Avg Reward: -277.76, Epsilon: 0.13
Episode 29400/100000, Avg Reward: -271.01, Epsilon: 0.13
Episode 29500/100000, Avg Reward: -275.24, Epsilon: 0.13
Episode 29600/100000, Avg Reward: -270.83, Epsilon: 0.13
Episode 29700/100000, Avg Reward: -271.55, Epsilon: 0.13
Episode 29800/100000, Avg Reward: -274.52, Epsilon: 0.13
Episode 29900/100000, Avg Reward: -273.17, Epsilon: 0.13
Episode 30000/100000, Avg Reward: -269.30, Epsilon: 0.13
Episode 30100/100000, Avg Reward: -272.45, Epsilon: 0.13
Episode 30200/100000, Avg Reward: -270.02, Epsilon: 0.12
Episode 30300/100000, Avg Reward: -272.99, Epsilon: 0.12
Episode 30400/100000, Avg Reward: -266.87, Epsilon: 0.12
Episode 30500/100000, Avg Reward: -269.93, Epsilon: 0.12
Episode 30600/100000, Avg Reward: -268.67, Epsilon: 0.12
Episode 30700/100000, Avg Reward: -268.31, Epsilon: 0.12
Episode 30800/100000, Avg Reward: -269.48, Epsilon: 0.12
Episode 30900/100000, Avg Reward: -267.59, Epsilon: 0.12
Episode 31000/100000, Avg Reward: -272.90, Epsilon: 0.12
Episode 31100/100000, Avg Reward: -266.06, Epsilon: 0.12
Episode 31200/100000, Avg Reward: -262.73, Epsilon: 0.12
Episode 31300/100000, Avg Reward: -263.45, Epsilon: 0.12
Episode 31400/100000, Avg Reward: -262.55, Epsilon: 0.11
Episode 31500/100000, Avg Reward: -265.70, Epsilon: 0.11
```

```
Episode 31600/100000, Avg Reward: -266.42, Epsilon: 0.11
Episode 31700/100000, Avg Reward: -262.55, Epsilon: 0.11
Episode 31800/100000, Avg Reward: -260.30, Epsilon: 0.11
Episode 31900/100000, Avg Reward: -269.48, Epsilon: 0.11
Episode 32000/100000, Avg Reward: -262.55, Epsilon: 0.11
Episode 32100/100000, Avg Reward: -260.48, Epsilon: 0.11
Episode 32200/100000, Avg Reward: -261.20, Epsilon: 0.11
Episode 32300/100000, Avg Reward: -261.29, Epsilon: 0.11
Episode 32400/100000, Avg Reward: -262.55, Epsilon: 0.11
Episode 32500/100000, Avg Reward: -262.46, Epsilon: 0.11
Episode 32600/100000, Avg Reward: -261.56, Epsilon: 0.11
Episode 32700/100000, Avg Reward: -258.14, Epsilon: 0.10
Episode 32800/100000, Avg Reward: -258.05, Epsilon: 0.10
Episode 32900/100000, Avg Reward: -258.05, Epsilon: 0.10
Episode 33000/100000, Avg Reward: -256.07, Epsilon: 0.10
Episode 33100/100000, Avg Reward: -260.48, Epsilon: 0.10
Episode 33200/100000, Avg Reward: -256.52, Epsilon: 0.10
Episode 33300/100000, Avg Reward: -256.79, Epsilon: 0.10
Episode 33400/100000, Avg Reward: -258.32, Epsilon: 0.10
Episode 33500/100000, Avg Reward: -259.31, Epsilon: 0.10
Episode 33600/100000, Avg Reward: -254.81, Epsilon: 0.10
Episode 33700/100000, Avg Reward: -254.00, Epsilon: 0.10
Episode 33800/100000, Avg Reward: -255.71, Epsilon: 0.10
Episode 33900/100000, Avg Reward: -254.18, Epsilon: 0.10
Episode 34000/100000, Avg Reward: -258.68, Epsilon: 0.10
Episode 34100/100000, Avg Reward: -255.17, Epsilon: 0.09
Episode 34200/100000, Avg Reward: -250.94, Epsilon: 0.09
Episode 34300/100000, Avg Reward: -253.64, Epsilon: 0.09
Episode 34400/100000, Avg Reward: -254.99, Epsilon: 0.09
Episode 34500/100000, Avg Reward: -251.93, Epsilon: 0.09
Episode 34600/100000, Avg Reward: -253.28, Epsilon: 0.09
Episode 34700/100000, Avg Reward: -249.23, Epsilon: 0.09
Episode 34800/100000, Avg Reward: -252.74, Epsilon: 0.09
Episode 34900/100000, Avg Reward: -250.40, Epsilon: 0.09
Episode 35000/100000, Avg Reward: -250.58, Epsilon: 0.09
Episode 35100/100000, Avg Reward: -250.94, Epsilon: 0.09
Episode 35200/100000, Avg Reward: -249.68, Epsilon: 0.09
Episode 35300/100000, Avg Reward: -254.63, Epsilon: 0.09
Episode 35400/100000, Avg Reward: -251.93, Epsilon: 0.09
Episode 35500/100000, Avg Reward: -251.57, Epsilon: 0.09
Episode 35600/100000, Avg Reward: -249.77, Epsilon: 0.09
Episode 35700/100000, Avg Reward: -247.43, Epsilon: 0.08
Episode 35800/100000, Avg Reward: -248.87, Epsilon: 0.08
Episode 35900/100000, Avg Reward: -244.73, Epsilon: 0.08
Episode 36000/100000, Avg Reward: -243.38, Epsilon: 0.08
Episode 36100/100000, Avg Reward: -250.04, Epsilon: 0.08
Episode 36200/100000, Avg Reward: -242.39, Epsilon: 0.08
Episode 36300/100000, Avg Reward: -242.48, Epsilon: 0.08
Episode 36400/100000, Avg Reward: -243.72, Epsilon: 0.08
Episode 36500/100000, Avg Reward: -249.23, Epsilon: 0.08
Episode 36600/100000, Avg Reward: -245.63, Epsilon: 0.08
Episode 36700/100000, Avg Reward: -244.82, Epsilon: 0.08
Episode 36800/100000, Avg Reward: -245.09, Epsilon: 0.08
Episode 36900/100000, Avg Reward: -245.00, Epsilon: 0.08
Episode 37000/100000, Avg Reward: -247.34, Epsilon: 0.08
Episode 37100/100000, Avg Reward: -240.14, Epsilon: 0.08
Episode 37200/100000, Avg Reward: -239.42, Epsilon: 0.08
```

```
Episode 37300/100000, Avg Reward: -239.78, Epsilon: 0.08
Episode 37400/100000, Avg Reward: -244.73, Epsilon: 0.08
Episode 37500/100000, Avg Reward: -243.11, Epsilon: 0.07
Episode 37600/100000, Avg Reward: -241.85, Epsilon: 0.07
Episode 37700/100000, Avg Reward: -246.80, Epsilon: 0.07
Episode 37800/100000, Avg Reward: -241.94, Epsilon: 0.07
Episode 37900/100000, Avg Reward: -243.29, Epsilon: 0.07
Episode 38000/100000, Avg Reward: -243.38, Epsilon: 0.07
Episode 38100/100000, Avg Reward: -243.20, Epsilon: 0.07
Episode 38200/100000, Avg Reward: -241.67, Epsilon: 0.07
Episode 38300/100000, Avg Reward: -240.77, Epsilon: 0.07
Episode 38400/100000, Avg Reward: -236.99, Epsilon: 0.07
Episode 38500/100000, Avg Reward: -243.74, Epsilon: 0.07
Episode 38600/100000, Avg Reward: -238.88, Epsilon: 0.07
Episode 38700/100000, Avg Reward: -235.46, Epsilon: 0.07
Episode 38800/100000, Avg Reward: -240.14, Epsilon: 0.07
Episode 38900/100000, Avg Reward: -240.86, Epsilon: 0.07
Episode 39000/100000, Avg Reward: -237.26, Epsilon: 0.07
Episode 39100/100000, Avg Reward: -236.72, Epsilon: 0.07
Episode 39200/100000, Avg Reward: -235.19, Epsilon: 0.07
Episode 39300/100000, Avg Reward: -237.89, Epsilon: 0.07
Episode 39400/100000, Avg Reward: -238.61, Epsilon: 0.07
Episode 39500/100000, Avg Reward: -234.56, Epsilon: 0.07
Episode 39600/100000, Avg Reward: -238.70, Epsilon: 0.06
Episode 39700/100000, Avg Reward: -236.54, Epsilon: 0.06
Episode 39800/100000, Avg Reward: -237.80, Epsilon: 0.06
Episode 39900/100000, Avg Reward: -233.39, Epsilon: 0.06
Episode 40000/100000, Avg Reward: -234.29, Epsilon: 0.06
Episode 40100/100000, Avg Reward: -237.17, Epsilon: 0.06
Episode 40200/100000, Avg Reward: -232.31, Epsilon: 0.06
Episode 40300/100000, Avg Reward: -236.54, Epsilon: 0.06
Episode 40400/100000, Avg Reward: -237.08, Epsilon: 0.06
Episode 40500/100000, Avg Reward: -232.67, Epsilon: 0.06
Episode 40600/100000, Avg Reward: -235.55, Epsilon: 0.06
Episode 40700/100000, Avg Reward: -235.46, Epsilon: 0.06
Episode 40800/100000, Avg Reward: -233.39, Epsilon: 0.06
Episode 40900/100000, Avg Reward: -232.04, Epsilon: 0.06
Episode 41000/100000, Avg Reward: -231.68, Epsilon: 0.06
Episode 41100/100000, Avg Reward: -232.58, Epsilon: 0.06
Episode 41200/100000, Avg Reward: -231.95, Epsilon: 0.06
Episode 41300/100000, Avg Reward: -231.77, Epsilon: 0.06
Episode 41400/100000, Avg Reward: -230.51, Epsilon: 0.06
Episode 41500/100000, Avg Reward: -231.50, Epsilon: 0.06
Episode 41600/100000, Avg Reward: -231.95, Epsilon: 0.06
Episode 41700/100000, Avg Reward: -231.86, Epsilon: 0.06
Episode 41800/100000, Avg Reward: -232.76, Epsilon: 0.06
Episode 41900/100000, Avg Reward: -230.87, Epsilon: 0.06
Episode 42000/100000, Avg Reward: -234.38, Epsilon: 0.05
Episode 42100/100000, Avg Reward: -230.06, Epsilon: 0.05
Episode 42200/100000, Avg Reward: -227.99, Epsilon: 0.05
Episode 42300/100000, Avg Reward: -232.49, Epsilon: 0.05
Episode 42400/100000, Avg Reward: -232.76, Epsilon: 0.05
Episode 42500/100000, Avg Reward: -231.59, Epsilon: 0.05
Episode 42600/100000, Avg Reward: -229.16, Epsilon: 0.05
Episode 42700/100000, Avg Reward: -228.44, Epsilon: 0.05
Episode 42800/100000, Avg Reward: -227.00, Epsilon: 0.05
Episode 42900/100000, Avg Reward: -228.80, Epsilon: 0.05
```

```
Episode 43000/100000, Avg Reward: -230.24, Epsilon: 0.05
Episode 43100/100000, Avg Reward: -229.34, Epsilon: 0.05
Episode 43200/100000, Avg Reward: -233.03, Epsilon: 0.05
Episode 43300/100000, Avg Reward: -226.64, Epsilon: 0.05
Episode 43400/100000, Avg Reward: -227.36, Epsilon: 0.05
Episode 43500/100000, Avg Reward: -229.61, Epsilon: 0.05
Episode 43600/100000, Avg Reward: -228.80, Epsilon: 0.05
Episode 43700/100000, Avg Reward: -231.05, Epsilon: 0.05
Episode 43800/100000, Avg Reward: -227.45, Epsilon: 0.05
Episode 43900/100000, Avg Reward: -230.06, Epsilon: 0.05
Episode 44000/100000, Avg Reward: -223.40, Epsilon: 0.05
Episode 44100/100000, Avg Reward: -226.37, Epsilon: 0.05
Episode 44200/100000, Avg Reward: -231.50, Epsilon: 0.05
Episode 44300/100000, Avg Reward: -227.00, Epsilon: 0.05
Episode 44400/100000, Avg Reward: -227.09, Epsilon: 0.05
Episode 44500/100000, Avg Reward: -226.10, Epsilon: 0.05
Episode 44600/100000, Avg Reward: -226.01, Epsilon: 0.05
Episode 44700/100000, Avg Reward: -224.93, Epsilon: 0.05
Episode 44800/100000, Avg Reward: -224.03, Epsilon: 0.05
Episode 44900/100000, Avg Reward: -226.01, Epsilon: 0.04
Episode 45000/100000, Avg Reward: -229.97, Epsilon: 0.04
Episode 45100/100000, Avg Reward: -224.12, Epsilon: 0.04
Episode 45200/100000, Avg Reward: -227.90, Epsilon: 0.04
Episode 45300/100000, Avg Reward: -227.09, Epsilon: 0.04
Episode 45400/100000, Avg Reward: -223.49, Epsilon: 0.04
Episode 45500/100000, Avg Reward: -225.11, Epsilon: 0.04
Episode 45600/100000, Avg Reward: -222.41, Epsilon: 0.04
Episode 45700/100000, Avg Reward: -220.07, Epsilon: 0.04
Episode 45800/100000, Avg Reward: -222.77, Epsilon: 0.04
Episode 45900/100000, Avg Reward: -225.92, Epsilon: 0.04
Episode 46000/100000, Avg Reward: -224.12, Epsilon: 0.04
Episode 46100/100000, Avg Reward: -222.77, Epsilon: 0.04
Episode 46200/100000, Avg Reward: -223.67, Epsilon: 0.04
Episode 46300/100000, Avg Reward: -222.05, Epsilon: 0.04
Episode 46400/100000, Avg Reward: -225.38, Epsilon: 0.04
Episode 46500/100000, Avg Reward: -220.61, Epsilon: 0.04
Episode 46600/100000, Avg Reward: -222.23, Epsilon: 0.04
Episode 46700/100000, Avg Reward: -223.58, Epsilon: 0.04
Episode 46800/100000, Avg Reward: -224.30, Epsilon: 0.04
Episode 46900/100000, Avg Reward: -221.33, Epsilon: 0.04
Episode 47000/100000, Avg Reward: -221.96, Epsilon: 0.04
Episode 47100/100000, Avg Reward: -221.87, Epsilon: 0.04
Episode 47200/100000, Avg Reward: -222.95, Epsilon: 0.04
Episode 47300/100000, Avg Reward: -222.68, Epsilon: 0.04
Episode 47400/100000, Avg Reward: -222.32, Epsilon: 0.04
Episode 47500/100000, Avg Reward: -220.79, Epsilon: 0.04
Episode 47600/100000, Avg Reward: -220.79, Epsilon: 0.04
Episode 47700/100000, Avg Reward: -222.50, Epsilon: 0.04
Episode 47800/100000, Avg Reward: -221.96, Epsilon: 0.04
Episode 47900/100000, Avg Reward: -220.43, Epsilon: 0.04
Episode 48000/100000, Avg Reward: -224.03, Epsilon: 0.04
Episode 48100/100000, Avg Reward: -221.51, Epsilon: 0.04
Episode 48200/100000, Avg Reward: -220.79, Epsilon: 0.04
Episode 48300/100000, Avg Reward: -221.24, Epsilon: 0.04
Episode 48400/100000, Avg Reward: -220.07, Epsilon: 0.04
Episode 48500/100000, Avg Reward: -218.36, Epsilon: 0.04
Episode 48600/100000, Avg Reward: -219.80, Epsilon: 0.03
```

```
Episode 48700/100000, Avg Reward: -219.44, Epsilon: 0.03
Episode 48800/100000, Avg Reward: -222.23, Epsilon: 0.03
Episode 48900/100000, Avg Reward: -219.26, Epsilon: 0.03
Episode 49000/100000, Avg Reward: -217.64, Epsilon: 0.03
Episode 49100/100000, Avg Reward: -219.08, Epsilon: 0.03
Episode 49200/100000, Avg Reward: -219.71, Epsilon: 0.03
Episode 49300/100000, Avg Reward: -220.07, Epsilon: 0.03
Episode 49400/100000, Avg Reward: -220.61, Epsilon: 0.03
Episode 49500/100000, Avg Reward: -218.18, Epsilon: 0.03
Episode 49600/100000, Avg Reward: -219.44, Epsilon: 0.03
Episode 49700/100000, Avg Reward: -221.42, Epsilon: 0.03
Episode 49800/100000, Avg Reward: -217.28, Epsilon: 0.03
Episode 49900/100000, Avg Reward: -219.71, Epsilon: 0.03
Episode 50000/100000, Avg Reward: -218.63, Epsilon: 0.03
Episode 50100/100000, Avg Reward: -217.46, Epsilon: 0.03
Episode 50200/100000, Avg Reward: -217.64, Epsilon: 0.03
Episode 50300/100000, Avg Reward: -218.99, Epsilon: 0.03
Episode 50400/100000, Avg Reward: -217.82, Epsilon: 0.03
Episode 50500/100000, Avg Reward: -218.00, Epsilon: 0.03
Episode 50600/100000, Avg Reward: -218.18, Epsilon: 0.03
Episode 50700/100000, Avg Reward: -217.55, Epsilon: 0.03
Episode 50800/100000, Avg Reward: -215.66, Epsilon: 0.03
Episode 50900/100000, Avg Reward: -216.38, Epsilon: 0.03
Episode 51000/100000, Avg Reward: -217.55, Epsilon: 0.03
Episode 51100/100000, Avg Reward: -217.55, Epsilon: 0.03
Episode 51200/100000, Avg Reward: -216.38, Epsilon: 0.03
Episode 51300/100000, Avg Reward: -216.65, Epsilon: 0.03
Episode 51400/100000, Avg Reward: -216.56, Epsilon: 0.03
Episode 51500/100000, Avg Reward: -217.01, Epsilon: 0.03
Episode 51600/100000, Avg Reward: -215.93, Epsilon: 0.03
Episode 51700/100000, Avg Reward: -217.01, Epsilon: 0.03
Episode 51800/100000, Avg Reward: -215.75, Epsilon: 0.03
Episode 51900/100000, Avg Reward: -215.66, Epsilon: 0.03
Episode 52000/100000, Avg Reward: -217.37, Epsilon: 0.03
Episode 52100/100000, Avg Reward: -215.93, Epsilon: 0.03
Episode 52200/100000, Avg Reward: -216.02, Epsilon: 0.03
Episode 52300/100000, Avg Reward: -215.12, Epsilon: 0.03
Episode 52400/100000, Avg Reward: -216.47, Epsilon: 0.03
Episode 52500/100000, Avg Reward: -215.93, Epsilon: 0.03
Episode 52600/100000, Avg Reward: -213.95, Epsilon: 0.03
Episode 52700/100000, Avg Reward: -213.32, Epsilon: 0.03
Episode 52800/100000, Avg Reward: -214.40, Epsilon: 0.03
Episode 52900/100000, Avg Reward: -214.76, Epsilon: 0.03
Episode 53000/100000, Avg Reward: -215.39, Epsilon: 0.03
Episode 53100/100000, Avg Reward: -214.85, Epsilon: 0.03
Episode 53200/100000, Avg Reward: -215.57, Epsilon: 0.03
Episode 53300/100000, Avg Reward: -215.93, Epsilon: 0.03
Episode 53400/100000, Avg Reward: -215.93, Epsilon: 0.03
Episode 53500/100000, Avg Reward: -215.12, Epsilon: 0.02
Episode 53600/100000, Avg Reward: -213.23, Epsilon: 0.02
Episode 53700/100000, Avg Reward: -212.96, Epsilon: 0.02
Episode 53800/100000, Avg Reward: -214.04, Epsilon: 0.02
Episode 53900/100000, Avg Reward: -215.84, Epsilon: 0.02
Episode 54000/100000, Avg Reward: -213.41, Epsilon: 0.02
Episode 54100/100000, Avg Reward: -215.30, Epsilon: 0.02
Episode 54200/100000, Avg Reward: -214.49, Epsilon: 0.02
Episode 54300/100000, Avg Reward: -212.42, Epsilon: 0.02
```

```
Episode 54400/100000, Avg Reward: -213.41, Epsilon: 0.02
Episode 54500/100000, Avg Reward: -212.60, Epsilon: 0.02
Episode 54600/100000, Avg Reward: -215.57, Epsilon: 0.02
Episode 54700/100000, Avg Reward: -214.58, Epsilon: 0.02
Episode 54800/100000, Avg Reward: -213.68, Epsilon: 0.02
Episode 54900/100000, Avg Reward: -211.34, Epsilon: 0.02
Episode 55000/100000, Avg Reward: -211.97, Epsilon: 0.02
Episode 55100/100000, Avg Reward: -213.50, Epsilon: 0.02
Episode 55200/100000, Avg Reward: -211.34, Epsilon: 0.02
Episode 55300/100000, Avg Reward: -211.70, Epsilon: 0.02
Episode 55400/100000, Avg Reward: -212.69, Epsilon: 0.02
Episode 55500/100000, Avg Reward: -211.07, Epsilon: 0.02
Episode 55600/100000, Avg Reward: -213.68, Epsilon: 0.02
Episode 55700/100000, Avg Reward: -216.02, Epsilon: 0.02
Episode 55800/100000, Avg Reward: -210.80, Epsilon: 0.02
Episode 55900/100000, Avg Reward: -211.61, Epsilon: 0.02
Episode 56000/100000, Avg Reward: -211.25, Epsilon: 0.02
Episode 56100/100000, Avg Reward: -211.25, Epsilon: 0.02
Episode 56200/100000, Avg Reward: -214.76, Epsilon: 0.02
Episode 56300/100000, Avg Reward: -211.16, Epsilon: 0.02
Episode 56400/100000, Avg Reward: -210.35, Epsilon: 0.02
Episode 56500/100000, Avg Reward: -212.51, Epsilon: 0.02
Episode 56600/100000, Avg Reward: -211.61, Epsilon: 0.02
Episode 56700/100000, Avg Reward: -210.98, Epsilon: 0.02
Episode 56800/100000, Avg Reward: -210.80, Epsilon: 0.02
Episode 56900/100000, Avg Reward: -209.90, Epsilon: 0.02
Episode 57000/100000, Avg Reward: -210.35, Epsilon: 0.02
Episode 57100/100000, Avg Reward: -210.35, Epsilon: 0.02
Episode 57200/100000, Avg Reward: -208.91, Epsilon: 0.02
Episode 57300/100000, Avg Reward: -211.07, Epsilon: 0.02
Episode 57400/100000, Avg Reward: -211.79, Epsilon: 0.02
Episode 57500/100000, Avg Reward: -210.08, Epsilon: 0.02
Episode 57600/100000, Avg Reward: -209.81, Epsilon: 0.02
Episode 57700/100000, Avg Reward: -209.72, Epsilon: 0.02
Episode 57800/100000, Avg Reward: -212.06, Epsilon: 0.02
Episode 57900/100000, Avg Reward: -210.26, Epsilon: 0.02
Episode 58000/100000, Avg Reward: -210.89, Epsilon: 0.02
Episode 58100/100000, Avg Reward: -209.63, Epsilon: 0.02
Episode 58200/100000, Avg Reward: -209.36, Epsilon: 0.02
Episode 58300/100000, Avg Reward: -210.80, Epsilon: 0.02
Episode 58400/100000, Avg Reward: -209.90, Epsilon: 0.02
Episode 58500/100000, Avg Reward: -211.34, Epsilon: 0.02
Episode 58600/100000, Avg Reward: -208.19, Epsilon: 0.02
Episode 58700/100000, Avg Reward: -208.37, Epsilon: 0.02
Episode 58800/100000, Avg Reward: -209.36, Epsilon: 0.02
Episode 58900/100000, Avg Reward: -210.44, Epsilon: 0.02
Episode 59000/100000, Avg Reward: -208.37, Epsilon: 0.02
Episode 59100/100000, Avg Reward: -209.72, Epsilon: 0.02
Episode 59200/100000, Avg Reward: -209.27, Epsilon: 0.02
Episode 59300/100000, Avg Reward: -208.91, Epsilon: 0.02
Episode 59400/100000, Avg Reward: -208.10, Epsilon: 0.02
Episode 59500/100000, Avg Reward: -207.47, Epsilon: 0.02
Episode 59600/100000, Avg Reward: -207.83, Epsilon: 0.02
Episode 59700/100000, Avg Reward: -208.55, Epsilon: 0.02
Episode 59800/100000, Avg Reward: -211.07, Epsilon: 0.02
Episode 59900/100000, Avg Reward: -210.08, Epsilon: 0.02
Episode 60000/100000, Avg Reward: -208.10, Epsilon: 0.02
```

```
Episode 60100/100000, Avg Reward: -209.18, Epsilon: 0.02
Episode 60200/100000, Avg Reward: -208.73, Epsilon: 0.02
Episode 60300/100000, Avg Reward: -208.46, Epsilon: 0.02
Episode 60400/100000, Avg Reward: -208.91, Epsilon: 0.02
Episode 60500/100000, Avg Reward: -209.99, Epsilon: 0.02
Episode 60600/100000, Avg Reward: -210.80, Epsilon: 0.02
Episode 60700/100000, Avg Reward: -209.00, Epsilon: 0.02
Episode 60800/100000, Avg Reward: -208.01, Epsilon: 0.01
Episode 60900/100000, Avg Reward: -208.73, Epsilon: 0.01
Episode 61000/100000, Avg Reward: -208.82, Epsilon: 0.01
Episode 61100/100000, Avg Reward: -208.55, Epsilon: 0.01
Episode 61200/100000, Avg Reward: -208.01, Epsilon: 0.01
Episode 61300/100000, Avg Reward: -207.29, Epsilon: 0.01
Episode 61400/100000, Avg Reward: -208.10, Epsilon: 0.01
Episode 61500/100000, Avg Reward: -208.19, Epsilon: 0.01
Episode 61600/100000, Avg Reward: -207.11, Epsilon: 0.01
Episode 61700/100000, Avg Reward: -207.92, Epsilon: 0.01
Episode 61800/100000, Avg Reward: -207.65, Epsilon: 0.01
Episode 61900/100000, Avg Reward: -208.10, Epsilon: 0.01
Episode 62000/100000, Avg Reward: -208.28, Epsilon: 0.01
Episode 62100/100000, Avg Reward: -208.64, Epsilon: 0.01
Episode 62200/100000, Avg Reward: -207.56, Epsilon: 0.01
Episode 62300/100000, Avg Reward: -208.19, Epsilon: 0.01
Episode 62400/100000, Avg Reward: -206.75, Epsilon: 0.01
Episode 62500/100000, Avg Reward: -207.38, Epsilon: 0.01
Episode 62600/100000, Avg Reward: -206.75, Epsilon: 0.01
Episode 62700/100000, Avg Reward: -207.20, Epsilon: 0.01
Episode 62800/100000, Avg Reward: -206.48, Epsilon: 0.01
Episode 62900/100000, Avg Reward: -208.10, Epsilon: 0.01
Episode 63000/100000, Avg Reward: -208.10, Epsilon: 0.01
Episode 63100/100000, Avg Reward: -207.92, Epsilon: 0.01
Episode 63200/100000, Avg Reward: -207.56, Epsilon: 0.01
Episode 63300/100000, Avg Reward: -207.56, Epsilon: 0.01
Episode 63400/100000, Avg Reward: -207.29, Epsilon: 0.01
Episode 63500/100000, Avg Reward: -206.84, Epsilon: 0.01
Episode 63600/100000, Avg Reward: -208.19, Epsilon: 0.01
Episode 63700/100000, Avg Reward: -207.02, Epsilon: 0.01
Episode 63800/100000, Avg Reward: -206.75, Epsilon: 0.01
Episode 63900/100000, Avg Reward: -206.57, Epsilon: 0.01
Episode 64000/100000, Avg Reward: -206.30, Epsilon: 0.01
Episode 64100/100000, Avg Reward: -207.65, Epsilon: 0.01
Episode 64200/100000, Avg Reward: -207.20, Epsilon: 0.01
Episode 64300/100000, Avg Reward: -205.85, Epsilon: 0.01
Episode 64400/100000, Avg Reward: -206.39, Epsilon: 0.01
Episode 64500/100000, Avg Reward: -206.39, Epsilon: 0.01
Episode 64600/100000, Avg Reward: -206.30, Epsilon: 0.01
Episode 64700/100000, Avg Reward: -206.48, Epsilon: 0.01
Episode 64800/100000, Avg Reward: -207.29, Epsilon: 0.01
Episode 64900/100000, Avg Reward: -204.59, Epsilon: 0.01
Episode 65000/100000, Avg Reward: -206.84, Epsilon: 0.01
Episode 65100/100000, Avg Reward: -205.22, Epsilon: 0.01
Episode 65200/100000, Avg Reward: -206.21, Epsilon: 0.01
Episode 65300/100000, Avg Reward: -205.85, Epsilon: 0.01
Episode 65400/100000, Avg Reward: -205.67, Epsilon: 0.01
Episode 65500/100000, Avg Reward: -206.39, Epsilon: 0.01
Episode 65600/100000, Avg Reward: -206.03, Epsilon: 0.01
Episode 65700/100000, Avg Reward: -206.12, Epsilon: 0.01
```

```
Episode 65800/100000, Avg Reward: -207.83, Epsilon: 0.01
Episode 65900/100000, Avg Reward: -204.86, Epsilon: 0.01
Episode 66000/100000, Avg Reward: -206.66, Epsilon: 0.01
Episode 66100/100000, Avg Reward: -207.56, Epsilon: 0.01
Episode 66200/100000, Avg Reward: -205.49, Epsilon: 0.01
Episode 66300/100000, Avg Reward: -206.21, Epsilon: 0.01
Episode 66400/100000, Avg Reward: -206.21, Epsilon: 0.01
Episode 66500/100000, Avg Reward: -205.31, Epsilon: 0.01
Episode 66600/100000, Avg Reward: -206.21, Epsilon: 0.01
Episode 66700/100000, Avg Reward: -205.85, Epsilon: 0.01
Episode 66800/100000, Avg Reward: -205.31, Epsilon: 0.01
Episode 66900/100000, Avg Reward: -204.86, Epsilon: 0.01
Episode 67000/100000, Avg Reward: -205.58, Epsilon: 0.01
Episode 67100/100000, Avg Reward: -204.95, Epsilon: 0.01
Episode 67200/100000, Avg Reward: -205.67, Epsilon: 0.01
Episode 67300/100000, Avg Reward: -205.76, Epsilon: 0.01
Episode 67400/100000, Avg Reward: -207.29, Epsilon: 0.01
Episode 67500/100000, Avg Reward: -205.76, Epsilon: 0.01
Episode 67600/100000, Avg Reward: -204.68, Epsilon: 0.01
Episode 67700/100000, Avg Reward: -223.40, Epsilon: 0.01
Episode 67800/100000, Avg Reward: -240.77, Epsilon: 0.01
Episode 67900/100000, Avg Reward: -205.13, Epsilon: 0.01
Episode 68000/100000, Avg Reward: -206.39, Epsilon: 0.01
Episode 68100/100000, Avg Reward: -205.31, Epsilon: 0.01
Episode 68200/100000, Avg Reward: -205.76, Epsilon: 0.01
Episode 68300/100000, Avg Reward: -205.76, Epsilon: 0.01
Episode 68400/100000, Avg Reward: -204.59, Epsilon: 0.01
Episode 68500/100000, Avg Reward: -203.96, Epsilon: 0.01
Episode 68600/100000, Avg Reward: -205.31, Epsilon: 0.01
Episode 68700/100000, Avg Reward: -205.40, Epsilon: 0.01
Episode 68800/100000, Avg Reward: -204.95, Epsilon: 0.01
Episode 68900/100000, Avg Reward: -204.77, Epsilon: 0.01
Episode 69000/100000, Avg Reward: -204.86, Epsilon: 0.01
Episode 69100/100000, Avg Reward: -205.58, Epsilon: 0.01
Episode 69200/100000, Avg Reward: -204.50, Epsilon: 0.01
Episode 69300/100000, Avg Reward: -204.32, Epsilon: 0.01
Episode 69400/100000, Avg Reward: -205.13, Epsilon: 0.01
Episode 69500/100000, Avg Reward: -203.96, Epsilon: 0.01
Episode 69600/100000, Avg Reward: -204.14, Epsilon: 0.01
Episode 69700/100000, Avg Reward: -204.23, Epsilon: 0.01
Episode 69800/100000, Avg Reward: -205.04, Epsilon: 0.01
Episode 69900/100000, Avg Reward: -203.96, Epsilon: 0.01
Episode 70000/100000, Avg Reward: -204.95, Epsilon: 0.01
Episode 70100/100000, Avg Reward: -203.96, Epsilon: 0.01
Episode 70200/100000, Avg Reward: -205.13, Epsilon: 0.01
Episode 70300/100000, Avg Reward: -204.95, Epsilon: 0.01
Episode 70400/100000, Avg Reward: -204.14, Epsilon: 0.01
Episode 70500/100000, Avg Reward: -203.60, Epsilon: 0.01
Episode 70600/100000, Avg Reward: -204.05, Epsilon: 0.01
Episode 70700/100000, Avg Reward: -203.96, Epsilon: 0.01
Episode 70800/100000, Avg Reward: -204.05, Epsilon: 0.01
Episode 70900/100000, Avg Reward: -203.51, Epsilon: 0.01
Episode 71000/100000, Avg Reward: -204.14, Epsilon: 0.01
Episode 71100/100000, Avg Reward: -204.86, Epsilon: 0.01
Episode 71200/100000, Avg Reward: -203.78, Epsilon: 0.01
Episode 71300/100000, Avg Reward: -204.50, Epsilon: 0.01
Episode 71400/100000, Avg Reward: -203.87, Epsilon: 0.01
```

```
Episode 71500/100000, Avg Reward: -204.05, Epsilon: 0.01
Episode 71600/100000, Avg Reward: -203.87, Epsilon: 0.01
Episode 71700/100000, Avg Reward: -203.78, Epsilon: 0.01
Episode 71800/100000, Avg Reward: -204.41, Epsilon: 0.01
Episode 71900/100000, Avg Reward: -204.50, Epsilon: 0.01
Episode 72000/100000, Avg Reward: -204.14, Epsilon: 0.01
Episode 72100/100000, Avg Reward: -203.69, Epsilon: 0.01
Episode 72200/100000, Avg Reward: -204.50, Epsilon: 0.01
Episode 72300/100000, Avg Reward: -203.78, Epsilon: 0.01
Episode 72400/100000, Avg Reward: -202.88, Epsilon: 0.01
Episode 72500/100000, Avg Reward: -203.69, Epsilon: 0.01
Episode 72600/100000, Avg Reward: -204.23, Epsilon: 0.01
Episode 72700/100000, Avg Reward: -203.69, Epsilon: 0.01
Episode 72800/100000, Avg Reward: -202.97, Epsilon: 0.01
Episode 72900/100000, Avg Reward: -204.41, Epsilon: 0.01
Episode 73000/100000, Avg Reward: -203.87, Epsilon: 0.01
Episode 73100/100000, Avg Reward: -203.78, Epsilon: 0.01
Episode 73200/100000, Avg Reward: -202.52, Epsilon: 0.01
Episode 73300/100000, Avg Reward: -203.33, Epsilon: 0.01
Episode 73400/100000, Avg Reward: -203.51, Epsilon: 0.01
Episode 73500/100000, Avg Reward: -203.60, Epsilon: 0.01
Episode 73600/100000, Avg Reward: -203.42, Epsilon: 0.01
Episode 73700/100000, Avg Reward: -204.32, Epsilon: 0.01
Episode 73800/100000, Avg Reward: -203.06, Epsilon: 0.01
Episode 73900/100000, Avg Reward: -204.68, Epsilon: 0.01
Episode 74000/100000, Avg Reward: -204.23, Epsilon: 0.01
Episode 74100/100000, Avg Reward: -203.96, Epsilon: 0.01
Episode 74200/100000, Avg Reward: -203.33, Epsilon: 0.01
Episode 74300/100000, Avg Reward: -203.24, Epsilon: 0.01
Episode 74400/100000, Avg Reward: -204.05, Epsilon: 0.01
Episode 74500/100000, Avg Reward: -202.43, Epsilon: 0.01
Episode 74600/100000, Avg Reward: -204.23, Epsilon: 0.01
Episode 74700/100000, Avg Reward: -203.33, Epsilon: 0.01
Episode 74800/100000, Avg Reward: -203.24, Epsilon: 0.01
Episode 74900/100000, Avg Reward: -202.52, Epsilon: 0.01
Episode 75000/100000, Avg Reward: -202.88, Epsilon: 0.01
Episode 75100/100000, Avg Reward: -202.43, Epsilon: 0.01
Episode 75200/100000, Avg Reward: -203.51, Epsilon: 0.01
Episode 75300/100000, Avg Reward: -203.15, Epsilon: 0.01
Episode 75400/100000, Avg Reward: -203.24, Epsilon: 0.01
Episode 75500/100000, Avg Reward: -203.96, Epsilon: 0.01
Episode 75600/100000, Avg Reward: -203.33, Epsilon: 0.01
Episode 75700/100000, Avg Reward: -202.79, Epsilon: 0.01
Episode 75800/100000, Avg Reward: -203.24, Epsilon: 0.01
Episode 75900/100000, Avg Reward: -203.24, Epsilon: 0.01
Episode 76000/100000, Avg Reward: -202.52, Epsilon: 0.01
Episode 76100/100000, Avg Reward: -203.78, Epsilon: 0.01
Episode 76200/100000, Avg Reward: -203.33, Epsilon: 0.01
Episode 76300/100000, Avg Reward: -202.79, Epsilon: 0.01
Episode 76400/100000, Avg Reward: -202.61, Epsilon: 0.01
Episode 76500/100000, Avg Reward: -202.70, Epsilon: 0.01
Episode 76600/100000, Avg Reward: -202.61, Epsilon: 0.01
Episode 76700/100000, Avg Reward: -203.24, Epsilon: 0.01
Episode 76800/100000, Avg Reward: -202.97, Epsilon: 0.00
Episode 76900/100000, Avg Reward: -202.43, Epsilon: 0.00
Episode 77000/100000, Avg Reward: -202.70, Epsilon: 0.00
Episode 77100/100000, Avg Reward: -202.88, Epsilon: 0.00
```

```
Episode 77200/100000, Avg Reward: -202.70, Epsilon: 0.00
Episode 77300/100000, Avg Reward: -202.61, Epsilon: 0.00
Episode 77400/100000, Avg Reward: -202.88, Epsilon: 0.00
Episode 77500/100000, Avg Reward: -203.78, Epsilon: 0.00
Episode 77600/100000, Avg Reward: -291.62, Epsilon: 0.00
Episode 77700/100000, Avg Reward: -202.79, Epsilon: 0.00
Episode 77800/100000, Avg Reward: -202.61, Epsilon: 0.00
Episode 77900/100000, Avg Reward: -203.51, Epsilon: 0.00
Episode 78000/100000, Avg Reward: -203.42, Epsilon: 0.00
Episode 78100/100000, Avg Reward: -202.88, Epsilon: 0.00
Episode 78200/100000, Avg Reward: -202.52, Epsilon: 0.00
Episode 78300/100000, Avg Reward: -202.25, Epsilon: 0.00
Episode 78400/100000, Avg Reward: -203.06, Epsilon: 0.00
Episode 78500/100000, Avg Reward: -202.52, Epsilon: 0.00
Episode 78600/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 78700/100000, Avg Reward: -202.16, Epsilon: 0.00
Episode 78800/100000, Avg Reward: -201.71, Epsilon: 0.00
Episode 78900/100000, Avg Reward: -202.70, Epsilon: 0.00
Episode 79000/100000, Avg Reward: -202.79, Epsilon: 0.00
Episode 79100/100000, Avg Reward: -202.16, Epsilon: 0.00
Episode 79200/100000, Avg Reward: -202.34, Epsilon: 0.00
Episode 79300/100000, Avg Reward: -202.16, Epsilon: 0.00
Episode 79400/100000, Avg Reward: -202.70, Epsilon: 0.00
Episode 79500/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 79600/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 79700/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 79800/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 79900/100000, Avg Reward: -202.16, Epsilon: 0.00
Episode 80000/100000, Avg Reward: -202.52, Epsilon: 0.00
Episode 80100/100000, Avg Reward: -202.88, Epsilon: 0.00
Episode 80200/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 80300/100000, Avg Reward: -202.79, Epsilon: 0.00
Episode 80400/100000, Avg Reward: -202.52, Epsilon: 0.00
Episode 80500/100000, Avg Reward: -202.16, Epsilon: 0.00
Episode 80600/100000, Avg Reward: -202.43, Epsilon: 0.00
Episode 80700/100000, Avg Reward: -202.61, Epsilon: 0.00
Episode 80800/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 80900/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 81000/100000, Avg Reward: -201.89, Epsilon: 0.00
Episode 81100/100000, Avg Reward: -201.71, Epsilon: 0.00
Episode 81200/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 81300/100000, Avg Reward: -202.70, Epsilon: 0.00
Episode 81400/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 81500/100000, Avg Reward: -201.71, Epsilon: 0.00
Episode 81600/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 81700/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 81800/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 81900/100000, Avg Reward: -201.80, Epsilon: 0.00
Episode 82000/100000, Avg Reward: -202.16, Epsilon: 0.00
Episode 82100/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 82200/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 82300/100000, Avg Reward: -202.43, Epsilon: 0.00
Episode 82400/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 82500/100000, Avg Reward: -273.71, Epsilon: 0.00
Episode 82600/100000, Avg Reward: -202.25, Epsilon: 0.00
Episode 82700/100000, Avg Reward: -202.61, Epsilon: 0.00
Episode 82800/100000, Avg Reward: -201.71, Epsilon: 0.00
```

```
Episode 82900/100000, Avg Reward: -201.71, Epsilon: 0.00
Episode 83000/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 83100/100000, Avg Reward: -202.25, Epsilon: 0.00
Episode 83200/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 83300/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 83400/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 83500/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 83600/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 83700/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 83800/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 83900/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 84000/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 84100/100000, Avg Reward: -201.71, Epsilon: 0.00
Episode 84200/100000, Avg Reward: -202.25, Epsilon: 0.00
Episode 84300/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 84400/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 84500/100000, Avg Reward: -201.44, Epsilon: 0.00
Episode 84600/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 84700/100000, Avg Reward: -202.07, Epsilon: 0.00
Episode 84800/100000, Avg Reward: -202.52, Epsilon: 0.00
Episode 84900/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 85000/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 85100/100000, Avg Reward: -201.89, Epsilon: 0.00
Episode 85200/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 85300/100000, Avg Reward: -201.71, Epsilon: 0.00
Episode 85400/100000, Avg Reward: -201.80, Epsilon: 0.00
Episode 85500/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 85600/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 85700/100000, Avg Reward: -201.44, Epsilon: 0.00
Episode 85800/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 85900/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 86000/100000, Avg Reward: -201.44, Epsilon: 0.00
Episode 86100/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 86200/100000, Avg Reward: -202.25, Epsilon: 0.00
Episode 86300/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 86400/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 86500/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 86600/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 86700/100000, Avg Reward: -202.16, Epsilon: 0.00
Episode 86800/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 86900/100000, Avg Reward: -236.90, Epsilon: 0.00
Episode 87000/100000, Avg Reward: -219.89, Epsilon: 0.00
Episode 87100/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 87200/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 87300/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 87400/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 87500/100000, Avg Reward: -201.89, Epsilon: 0.00
Episode 87600/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 87700/100000, Avg Reward: -201.44, Epsilon: 0.00
Episode 87800/100000, Avg Reward: -201.44, Epsilon: 0.00
Episode 87900/100000, Avg Reward: -219.17, Epsilon: 0.00
Episode 88000/100000, Avg Reward: -201.98, Epsilon: 0.00
Episode 88100/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 88200/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 88300/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 88400/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 88500/100000, Avg Reward: -200.99, Epsilon: 0.00
```

```
Episode 88600/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 88700/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 88800/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 88900/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 89000/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 89100/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 89200/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 89300/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 89400/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 89500/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 89600/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 89700/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 89800/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 89900/100000, Avg Reward: -201.44, Epsilon: 0.00
Episode 90000/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 90100/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 90200/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 90300/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 90400/100000, Avg Reward: -201.35, Epsilon: 0.00
Episode 90500/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 90600/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 90700/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 90800/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 90900/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 91000/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 91100/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 91200/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 91300/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 91400/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 91500/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 91600/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 91700/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 91800/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 91900/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 92000/100000, Avg Reward: -201.62, Epsilon: 0.00
Episode 92100/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 92200/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 92300/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 92400/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 92500/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 92600/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 92700/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 92800/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 92900/100000, Avg Reward: -201.26, Epsilon: 0.00
Episode 93000/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 93100/100000, Avg Reward: -201.71, Epsilon: 0.00
Episode 93200/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 93300/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 93400/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 93500/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 93600/100000, Avg Reward: -218.54, Epsilon: 0.00
Episode 93700/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 93800/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 93900/100000, Avg Reward: -201.17, Epsilon: 0.00
Episode 94000/100000, Avg Reward: -200.27, Epsilon: 0.00
Episode 94100/100000, Avg Reward: -200.36, Epsilon: 0.00
Episode 94200/100000, Avg Reward: -200.72, Epsilon: 0.00
```

```
Episode 94300/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 94400/100000, Avg Reward: -201.44, Epsilon: 0.00
Episode 94500/100000, Avg Reward: -201.53, Epsilon: 0.00
Episode 94600/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 94700/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 94800/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 94900/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 95000/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 95100/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 95200/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 95300/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 95400/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 95500/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 95600/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 95700/100000, Avg Reward: -200.36, Epsilon: 0.00
Episode 95800/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 95900/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 96000/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 96100/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 96200/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 96300/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 96400/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 96500/100000, Avg Reward: -218.36, Epsilon: 0.00
Episode 96600/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 96700/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 96800/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 96900/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 97000/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 97100/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 97200/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 97300/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 97400/100000, Avg Reward: -201.08, Epsilon: 0.00
Episode 97500/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 97600/100000, Avg Reward: -200.45, Epsilon: 0.00
Episode 97700/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 97800/100000, Avg Reward: -200.45, Epsilon: 0.00
Episode 97900/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 98000/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 98100/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 98200/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 98300/100000, Avg Reward: -200.45, Epsilon: 0.00
Episode 98400/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 98500/100000, Avg Reward: -200.90, Epsilon: 0.00
Episode 98600/100000, Avg Reward: -200.45, Epsilon: 0.00
Episode 98700/100000, Avg Reward: -200.72, Epsilon: 0.00
Episode 98800/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 98900/100000, Avg Reward: -200.99, Epsilon: 0.00
Episode 99000/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 99100/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 99200/100000, Avg Reward: -200.45, Epsilon: 0.00
Episode 99300/100000, Avg Reward: -200.54, Epsilon: 0.00
Episode 99400/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 99500/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 99600/100000, Avg Reward: -200.45, Epsilon: 0.00
Episode 99700/100000, Avg Reward: -200.63, Epsilon: 0.00
Episode 99800/100000, Avg Reward: -200.81, Epsilon: 0.00
Episode 99900/100000, Avg Reward: -200.72, Epsilon: 0.00
```

```
Episode 100000/100000, Avg Reward: -200.90, Epsilon: 0.00
!!!!!!!!!!!!!!!!!Trained and Stored the model successfully!!!!!!!!!!!!!!!!!
CPU times: user 2h 9min 59s, sys: 5min 10s, total: 2h 15min 10s
Wall time: 59min 31s
```
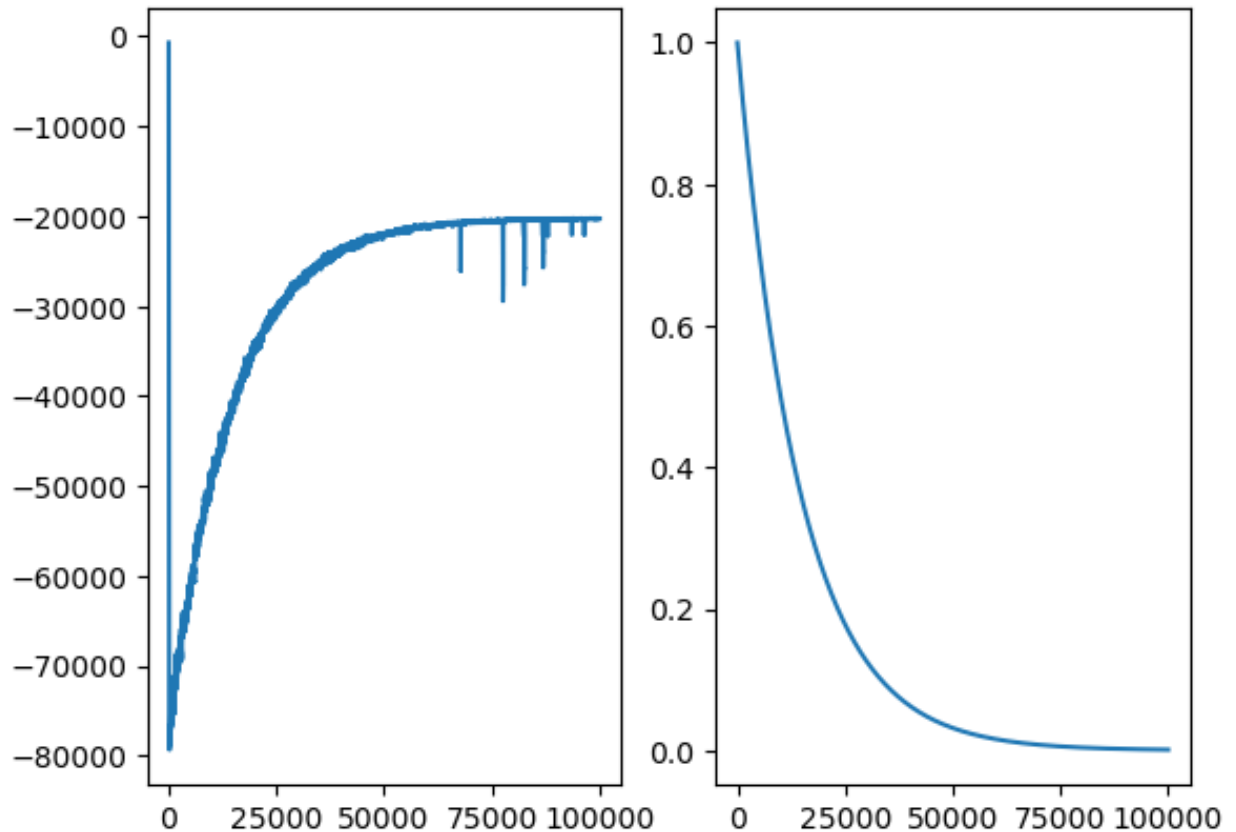


```
In [38]: print(my_dqn_agent_controlled_training_rewards )

         [-704. -722. -866. ... -200. -200. -200.]

In [39]: controlled_training = {
             "my_q_agent_controlled_training_rewards": list(my_q_agent_controlled_tr
             "my_dqn_agent_controlled_training_rewards": list(my_dqn_agent_controlle
         }

In [40]: print(type(my_q_agent_controlled_training_rewards))
         print(type(my_dqn_agent_controlled_training_rewards))

         <class 'numpy.ndarray'>
         <class 'numpy.ndarray'>

In [41]: with open("controlled_training.json", "w") as f:
             json.dump(controlled_training, f)

         # r = {'is_claimed': 'True', 'rating': 3.5}
         # r = json.dumps(controlled_training)
         # loaded_r = json.loads(r)

In [42]: with open("controlled_training.json", "r") as f:
             controlled_training_dict = json.load(f)

         # controlled_training_dict = json.loads(jsonString)
```

Compare the training reward

```
In [44]: x = range(1,100000 + 1)

         plt.figure(figsize=(10, 6))
         plt.plot(x, controlled_training_dict["my_dqn_agent_controlled_training_re
         plt.plot(x, controlled_training_dict["my_q_agent_controlled_training_rewa

         # titles and labels
         plt.title('Training Comparison (Controlled Parameters)')
         plt.xlabel('Time Episodes')
         plt.ylabel('Average Reward')
         plt.legend()
         # display
         plt.grid(True)
         plt.show()
```
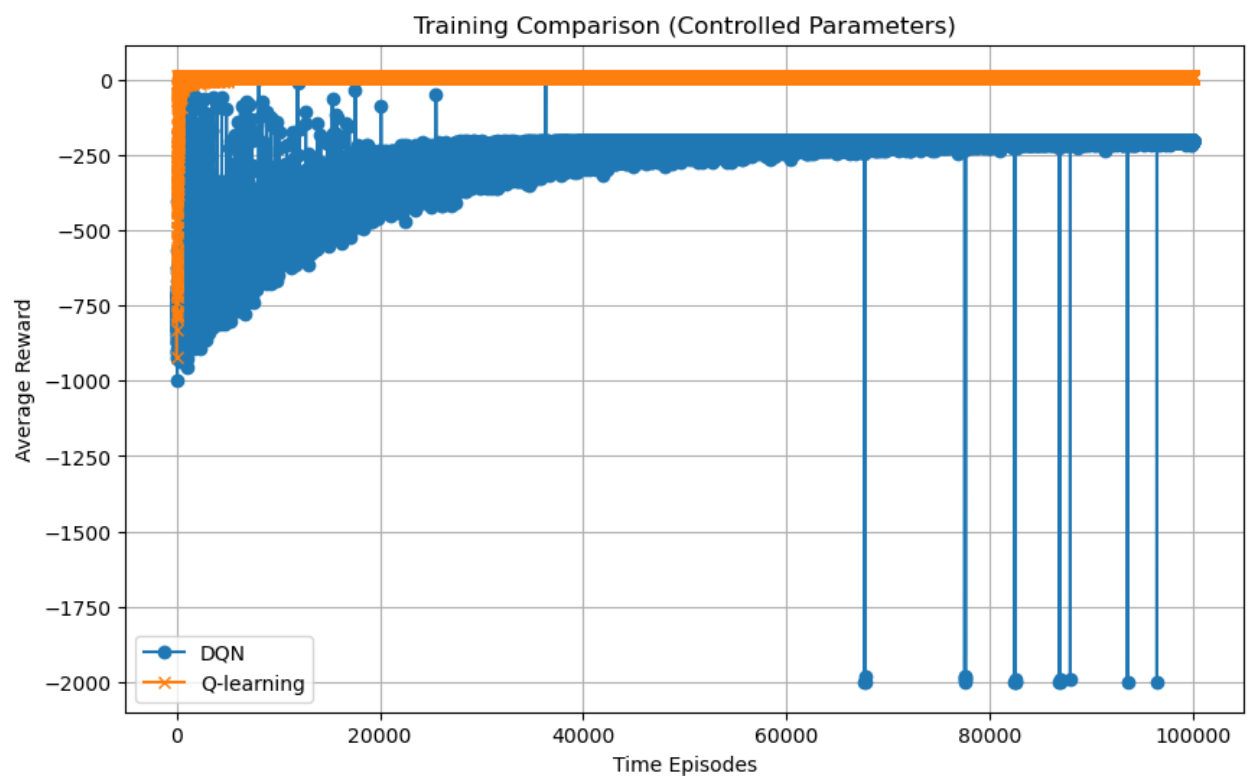

Prediction with the Controlled models

Q Learning Agent with Controlled Parameters

```
In [47]: my_q_agent_controlled_avg_reward_lst = my_q_agent_controlled.test(test_ep
```

```
========================================================
EPISODE  0
Score of total rewards: 6
========================================================
EPISODE  1
Score of total rewards: 8
========================================================
EPISODE  2
Score of total rewards: 9
========================================================
EPISODE  3
Score of total rewards: 9
```

```
========================================================
EPISODE  4
Score of total rewards: 11
========================================================
EPISODE  5
Score of total rewards: 4
========================================================
EPISODE  6
Score of total rewards: 6
========================================================
EPISODE  7
Score of total rewards: 5
========================================================
EPISODE  8
Score of total rewards: 5
========================================================
EPISODE  9
Score of total rewards: 13
========================================================
EPISODE  10
Score of total rewards: 7
========================================================
EPISODE  11
Score of total rewards: 11
========================================================
EPISODE  12
Score of total rewards: 4
========================================================
EPISODE  13
Score of total rewards: 11
========================================================
EPISODE  14
Score of total rewards: 7
========================================================
EPISODE  15
Score of total rewards: 8
========================================================
EPISODE  16
Score of total rewards: 7
========================================================
EPISODE  17
Score of total rewards: 4
========================================================
EPISODE  18
Score of total rewards: 6
========================================================
EPISODE  19
Score of total rewards: 12
========================================================
EPISODE  20
Score of total rewards: 7
========================================================
EPISODE  21
Score of total rewards: 4
========================================================
EPISODE  22
Score of total rewards: 10
```

```
========================================================
EPISODE  23
Score of total rewards: 7
========================================================
EPISODE  24
Score of total rewards: 12
========================================================
EPISODE  25
Score of total rewards: 5
========================================================
EPISODE  26
Score of total rewards: 5
========================================================
EPISODE  27
Score of total rewards: 7
========================================================
EPISODE  28
Score of total rewards: 7
========================================================
EPISODE  29
Score of total rewards: 6
========================================================
EPISODE  30
Score of total rewards: 11
========================================================
EPISODE  31
Score of total rewards: 11
========================================================
EPISODE  32
Score of total rewards: 12
========================================================
EPISODE  33
Score of total rewards: 5
========================================================
EPISODE  34
Score of total rewards: 10
========================================================
EPISODE  35
Score of total rewards: 8
========================================================
EPISODE  36
Score of total rewards: 10
========================================================
EPISODE  37
Score of total rewards: 11
========================================================
EPISODE  38
Score of total rewards: 7
========================================================
EPISODE  39
Score of total rewards: 12
========================================================
EPISODE  40
Score of total rewards: 11
========================================================
EPISODE  41
Score of total rewards: 9
```

```
========================================================
EPISODE  42
Score of total rewards: 9
========================================================
EPISODE  43
Score of total rewards: 11
========================================================
EPISODE  44
Score of total rewards: 9
========================================================
EPISODE  45
Score of total rewards: 9
========================================================
EPISODE  46
Score of total rewards: 6
========================================================
EPISODE  47
Score of total rewards: 7
========================================================
EPISODE  48
Score of total rewards: 5
========================================================
EPISODE  49
Score of total rewards: 9
========================================================
EPISODE  50
Score of total rewards: 6
========================================================
EPISODE  51
Score of total rewards: 7
========================================================
EPISODE  52
Score of total rewards: 7
========================================================
EPISODE  53
Score of total rewards: 7
========================================================
EPISODE  54
Score of total rewards: 4
========================================================
EPISODE  55
Score of total rewards: 9
========================================================
EPISODE  56
Score of total rewards: 13
========================================================
EPISODE  57
Score of total rewards: 4
========================================================
EPISODE  58
Score of total rewards: 11
========================================================
EPISODE  59
Score of total rewards: 5
========================================================
EPISODE  60
Score of total rewards: 8
```

```
========================================================
EPISODE  61
Score of total rewards: 7
========================================================
EPISODE  62
Score of total rewards: 7
========================================================
EPISODE  63
Score of total rewards: 7
========================================================
EPISODE  64
Score of total rewards: 5
========================================================
EPISODE  65
Score of total rewards: 6
========================================================
EPISODE  66
Score of total rewards: 6
========================================================
EPISODE  67
Score of total rewards: 4
========================================================
EPISODE  68
Score of total rewards: 6
========================================================
EPISODE  69
Score of total rewards: 11
========================================================
EPISODE  70
Score of total rewards: 8
========================================================
EPISODE  71
Score of total rewards: 7
========================================================
EPISODE  72
Score of total rewards: 4
========================================================
EPISODE  73
Score of total rewards: 6
========================================================
EPISODE  74
Score of total rewards: 11
========================================================
EPISODE  75
Score of total rewards: 7
========================================================
EPISODE  76
Score of total rewards: 3
========================================================
EPISODE  77
Score of total rewards: 9
========================================================
EPISODE  78
Score of total rewards: 9
========================================================
EPISODE  79
Score of total rewards: 9
```

```
============================================================
EPISODE  80
Score of total rewards: 3
============================================================
EPISODE  81
Score of total rewards: 8
============================================================
EPISODE  82
Score of total rewards: 10
============================================================
EPISODE  83
Score of total rewards: 12
============================================================
EPISODE  84
Score of total rewards: 7
============================================================
EPISODE  85
Score of total rewards: 4
============================================================
EPISODE  86
Score of total rewards: 8
============================================================
EPISODE  87
Score of total rewards: 9
============================================================
EPISODE  88
Score of total rewards: 8
============================================================
EPISODE  89
Score of total rewards: 8
============================================================
EPISODE  90
Score of total rewards: 8
============================================================
EPISODE  91
Score of total rewards: 10
============================================================
EPISODE  92
Score of total rewards: 9
============================================================
EPISODE  93
Score of total rewards: 8
============================================================
EPISODE  94
Score of total rewards: 7
============================================================
EPISODE  95
Score of total rewards: 8
============================================================
EPISODE  96
Score of total rewards: 10
============================================================
EPISODE  97
Score of total rewards: 5
============================================================
EPISODE  98
Score of total rewards: 11
```

```
========================================================
EPISODE  99
Score of total rewards: 4
========================================================
EPISODE  100
Score of total rewards: 7
========================================================
EPISODE  101
Score of total rewards: 10
========================================================
EPISODE  102
Score of total rewards: 7
========================================================
EPISODE  103
Score of total rewards: 8
========================================================
EPISODE  104
Score of total rewards: 9
========================================================
EPISODE  105
Score of total rewards: 4
========================================================
EPISODE  106
Score of total rewards: 9
========================================================
EPISODE  107
Score of total rewards: 10
========================================================
EPISODE  108
Score of total rewards: 9
========================================================
EPISODE  109
Score of total rewards: 10
========================================================
EPISODE  110
Score of total rewards: 6
========================================================
EPISODE  111
Score of total rewards: 8
========================================================
EPISODE  112
Score of total rewards: 8
========================================================
EPISODE  113
Score of total rewards: 12
========================================================
EPISODE  114
Score of total rewards: 7
========================================================
EPISODE  115
Score of total rewards: 5
========================================================
EPISODE  116
Score of total rewards: 10
========================================================
EPISODE  117
Score of total rewards: 4
```

```
========================================================
EPISODE  118
Score of total rewards: 9
========================================================
EPISODE  119
Score of total rewards: 8
========================================================
EPISODE  120
Score of total rewards: 7
========================================================
EPISODE  121
Score of total rewards: 12
========================================================
EPISODE  122
Score of total rewards: 5
========================================================
EPISODE  123
Score of total rewards: 9
========================================================
EPISODE  124
Score of total rewards: 10
========================================================
EPISODE  125
Score of total rewards: 8
========================================================
EPISODE  126
Score of total rewards: 9
========================================================
EPISODE  127
Score of total rewards: 9
========================================================
EPISODE  128
Score of total rewards: 11
========================================================
EPISODE  129
Score of total rewards: 7
========================================================
EPISODE  130
Score of total rewards: 10
========================================================
EPISODE  131
Score of total rewards: 10
========================================================
EPISODE  132
Score of total rewards: 10
========================================================
EPISODE  133
Score of total rewards: 7
========================================================
EPISODE  134
Score of total rewards: 5
========================================================
EPISODE  135
Score of total rewards: 9
========================================================
EPISODE  136
Score of total rewards: 6
```

```
========================================================
EPISODE  137
Score of total rewards: 9
========================================================
EPISODE  138
Score of total rewards: 7
========================================================
EPISODE  139
Score of total rewards: 13
========================================================
EPISODE  140
Score of total rewards: 9
========================================================
EPISODE  141
Score of total rewards: 10
========================================================
EPISODE  142
Score of total rewards: 6
========================================================
EPISODE  143
Score of total rewards: 8
========================================================
EPISODE  144
Score of total rewards: 10
========================================================
EPISODE  145
Score of total rewards: 6
========================================================
EPISODE  146
Score of total rewards: 8
========================================================
EPISODE  147
Score of total rewards: 10
========================================================
EPISODE  148
Score of total rewards: 13
========================================================
EPISODE  149
Score of total rewards: 4
========================================================
EPISODE  150
Score of total rewards: 5
========================================================
EPISODE  151
Score of total rewards: 4
========================================================
EPISODE  152
Score of total rewards: 8
========================================================
EPISODE  153
Score of total rewards: 6
========================================================
EPISODE  154
Score of total rewards: 14
========================================================
EPISODE  155
Score of total rewards: 5
```

```
========================================================
EPISODE  156
Score of total rewards: 9
========================================================
EPISODE  157
Score of total rewards: 9
========================================================
EPISODE  158
Score of total rewards: 11
========================================================
EPISODE  159
Score of total rewards: 13
========================================================
EPISODE  160
Score of total rewards: 5
========================================================
EPISODE  161
Score of total rewards: 10
========================================================
EPISODE  162
Score of total rewards: 3
========================================================
EPISODE  163
Score of total rewards: 3
========================================================
EPISODE  164
Score of total rewards: 10
========================================================
EPISODE  165
Score of total rewards: 7
========================================================
EPISODE  166
Score of total rewards: 10
========================================================
EPISODE  167
Score of total rewards: 9
========================================================
EPISODE  168
Score of total rewards: 6
========================================================
EPISODE  169
Score of total rewards: 5
========================================================
EPISODE  170
Score of total rewards: 7
========================================================
EPISODE  171
Score of total rewards: 11
========================================================
EPISODE  172
Score of total rewards: 10
========================================================
EPISODE  173
Score of total rewards: 10
========================================================
EPISODE  174
Score of total rewards: 10
```

```
============================================================
EPISODE  175
Score of total rewards: 6
============================================================
EPISODE  176
Score of total rewards: 15
============================================================
EPISODE  177
Score of total rewards: 7
============================================================
EPISODE  178
Score of total rewards: 8
============================================================
EPISODE  179
Score of total rewards: 10
============================================================
EPISODE  180
Score of total rewards: 10
============================================================
EPISODE  181
Score of total rewards: 7
============================================================
EPISODE  182
Score of total rewards: 10
============================================================
EPISODE  183
Score of total rewards: 5
============================================================
EPISODE  184
Score of total rewards: 3
============================================================
EPISODE  185
Score of total rewards: 5
============================================================
EPISODE  186
Score of total rewards: 8
============================================================
EPISODE  187
Score of total rewards: 7
============================================================
EPISODE  188
Score of total rewards: 6
============================================================
EPISODE  189
Score of total rewards: 5
============================================================
EPISODE  190
Score of total rewards: 8
============================================================
EPISODE  191
Score of total rewards: 12
============================================================
EPISODE  192
Score of total rewards: 8
============================================================
EPISODE  193
Score of total rewards: 14
```

```
========================================================
EPISODE  194
Score of total rewards: 11
========================================================
EPISODE  195
Score of total rewards: 12
========================================================
EPISODE  196
Score of total rewards: 9
========================================================
EPISODE  197
Score of total rewards: 12
========================================================
EPISODE  198
Score of total rewards: 8
========================================================
EPISODE  199
Score of total rewards: 9
========================================================
EPISODE  200
Score of total rewards: 15
========================================================
EPISODE  201
Score of total rewards: 6
========================================================
EPISODE  202
Score of total rewards: 7
========================================================
EPISODE  203
Score of total rewards: 6
========================================================
EPISODE  204
Score of total rewards: 9
========================================================
EPISODE  205
Score of total rewards: 5
========================================================
EPISODE  206
Score of total rewards: 10
========================================================
EPISODE  207
Score of total rewards: 5
========================================================
EPISODE  208
Score of total rewards: 4
========================================================
EPISODE  209
Score of total rewards: 11
========================================================
EPISODE  210
Score of total rewards: 10
========================================================
EPISODE  211
Score of total rewards: 7
========================================================
EPISODE  212
Score of total rewards: 9
```

```
============================================================
EPISODE  213
Score of total rewards: 10
============================================================
EPISODE  214
Score of total rewards: 13
============================================================
EPISODE  215
Score of total rewards: 8
============================================================
EPISODE  216
Score of total rewards: 5
============================================================
EPISODE  217
Score of total rewards: 5
============================================================
EPISODE  218
Score of total rewards: 9
============================================================
EPISODE  219
Score of total rewards: 4
============================================================
EPISODE  220
Score of total rewards: 5
============================================================
EPISODE  221
Score of total rewards: 8
============================================================
EPISODE  222
Score of total rewards: 11
============================================================
EPISODE  223
Score of total rewards: 10
============================================================
EPISODE  224
Score of total rewards: 8
============================================================
EPISODE  225
Score of total rewards: 5
============================================================
EPISODE  226
Score of total rewards: 8
============================================================
EPISODE  227
Score of total rewards: 11
============================================================
EPISODE  228
Score of total rewards: 8
============================================================
EPISODE  229
Score of total rewards: 7
============================================================
EPISODE  230
Score of total rewards: 4
============================================================
EPISODE  231
Score of total rewards: 11
```

```
========================================================
EPISODE  232
Score of total rewards: 7
========================================================
EPISODE  233
Score of total rewards: 3
========================================================
EPISODE  234
Score of total rewards: 9
========================================================
EPISODE  235
Score of total rewards: 10
========================================================
EPISODE  236
Score of total rewards: 13
========================================================
EPISODE  237
Score of total rewards: 8
========================================================
EPISODE  238
Score of total rewards: 6
========================================================
EPISODE  239
Score of total rewards: 10
========================================================
EPISODE  240
Score of total rewards: 7
========================================================
EPISODE  241
Score of total rewards: 9
========================================================
EPISODE  242
Score of total rewards: 10
========================================================
EPISODE  243
Score of total rewards: 10
========================================================
EPISODE  244
Score of total rewards: 7
========================================================
EPISODE  245
Score of total rewards: 9
========================================================
EPISODE  246
Score of total rewards: 6
========================================================
EPISODE  247
Score of total rewards: 5
========================================================
EPISODE  248
Score of total rewards: 7
========================================================
EPISODE  249
Score of total rewards: 14
========================================================
EPISODE  250
Score of total rewards: 4
```

```
========================================================
EPISODE  251
Score of total rewards: 13
========================================================
EPISODE  252
Score of total rewards: 9
========================================================
EPISODE  253
Score of total rewards: 4
========================================================
EPISODE  254
Score of total rewards: 9
========================================================
EPISODE  255
Score of total rewards: 10
========================================================
EPISODE  256
Score of total rewards: 5
========================================================
EPISODE  257
Score of total rewards: 4
========================================================
EPISODE  258
Score of total rewards: 14
========================================================
EPISODE  259
Score of total rewards: 8
========================================================
EPISODE  260
Score of total rewards: 8
========================================================
EPISODE  261
Score of total rewards: 9
========================================================
EPISODE  262
Score of total rewards: 11
========================================================
EPISODE  263
Score of total rewards: 9
========================================================
EPISODE  264
Score of total rewards: 14
========================================================
EPISODE  265
Score of total rewards: 9
========================================================
EPISODE  266
Score of total rewards: 6
========================================================
EPISODE  267
Score of total rewards: 8
========================================================
EPISODE  268
Score of total rewards: 5
========================================================
EPISODE  269
Score of total rewards: 5
```

```
========================================================
EPISODE  270
Score of total rewards: 4
========================================================
EPISODE  271
Score of total rewards: 6
========================================================
EPISODE  272
Score of total rewards: 6
========================================================
EPISODE  273
Score of total rewards: 6
========================================================
EPISODE  274
Score of total rewards: 7
========================================================
EPISODE  275
Score of total rewards: 9
========================================================
EPISODE  276
Score of total rewards: 7
========================================================
EPISODE  277
Score of total rewards: 6
========================================================
EPISODE  278
Score of total rewards: 9
========================================================
EPISODE  279
Score of total rewards: 4
========================================================
EPISODE  280
Score of total rewards: 5
========================================================
EPISODE  281
Score of total rewards: 11
========================================================
EPISODE  282
Score of total rewards: 5
========================================================
EPISODE  283
Score of total rewards: 8
========================================================
EPISODE  284
Score of total rewards: 9
========================================================
EPISODE  285
Score of total rewards: 4
========================================================
EPISODE  286
Score of total rewards: 7
========================================================
EPISODE  287
Score of total rewards: 8
========================================================
EPISODE  288
Score of total rewards: 13
```

```
========================================================
EPISODE  289
Score of total rewards: 8
========================================================
EPISODE  290
Score of total rewards: 11
========================================================
EPISODE  291
Score of total rewards: 7
========================================================
EPISODE  292
Score of total rewards: 6
========================================================
EPISODE  293
Score of total rewards: 9
========================================================
EPISODE  294
Score of total rewards: 9
========================================================
EPISODE  295
Score of total rewards: 8
========================================================
EPISODE  296
Score of total rewards: 9
========================================================
EPISODE  297
Score of total rewards: 13
========================================================
EPISODE  298
Score of total rewards: 6
========================================================
EPISODE  299
Score of total rewards: 15
========================================================
EPISODE  300
Score of total rewards: 9
========================================================
EPISODE  301
Score of total rewards: 7
========================================================
EPISODE  302
Score of total rewards: 11
========================================================
EPISODE  303
Score of total rewards: 10
========================================================
EPISODE  304
Score of total rewards: 6
========================================================
EPISODE  305
Score of total rewards: 10
========================================================
EPISODE  306
Score of total rewards: 7
========================================================
EPISODE  307
Score of total rewards: 11
```

```
============================================================
EPISODE  308
Score of total rewards: 8
============================================================
EPISODE  309
Score of total rewards: 5
============================================================
EPISODE  310
Score of total rewards: 8
============================================================
EPISODE  311
Score of total rewards: 11
============================================================
EPISODE  312
Score of total rewards: 7
============================================================
EPISODE  313
Score of total rewards: 5
============================================================
EPISODE  314
Score of total rewards: 6
============================================================
EPISODE  315
Score of total rewards: 8
============================================================
EPISODE  316
Score of total rewards: 4
============================================================
EPISODE  317
Score of total rewards: 8
============================================================
EPISODE  318
Score of total rewards: 9
============================================================
EPISODE  319
Score of total rewards: 10
============================================================
EPISODE  320
Score of total rewards: 4
============================================================
EPISODE  321
Score of total rewards: 6
============================================================
EPISODE  322
Score of total rewards: 5
============================================================
EPISODE  323
Score of total rewards: 13
============================================================
EPISODE  324
Score of total rewards: 15
============================================================
EPISODE  325
Score of total rewards: 11
============================================================
EPISODE  326
Score of total rewards: 10
```

```
========================================================
EPISODE  327
Score of total rewards: 9
========================================================
EPISODE  328
Score of total rewards: 10
========================================================
EPISODE  329
Score of total rewards: 6
========================================================
EPISODE  330
Score of total rewards: 7
========================================================
EPISODE  331
Score of total rewards: 4
========================================================
EPISODE  332
Score of total rewards: 7
========================================================
EPISODE  333
Score of total rewards: 14
========================================================
EPISODE  334
Score of total rewards: 7
========================================================
EPISODE  335
Score of total rewards: 8
========================================================
EPISODE  336
Score of total rewards: 4
========================================================
EPISODE  337
Score of total rewards: 5
========================================================
EPISODE  338
Score of total rewards: 7
========================================================
EPISODE  339
Score of total rewards: 12
========================================================
EPISODE  340
Score of total rewards: 4
========================================================
EPISODE  341
Score of total rewards: 4
========================================================
EPISODE  342
Score of total rewards: 9
========================================================
EPISODE  343
Score of total rewards: 9
========================================================
EPISODE  344
Score of total rewards: 7
========================================================
EPISODE  345
Score of total rewards: 6
```

```
========================================================
EPISODE  346
Score of total rewards: 14
========================================================
EPISODE  347
Score of total rewards: 10
========================================================
EPISODE  348
Score of total rewards: 9
========================================================
EPISODE  349
Score of total rewards: 6
========================================================
EPISODE  350
Score of total rewards: 6
========================================================
EPISODE  351
Score of total rewards: 10
========================================================
EPISODE  352
Score of total rewards: 3
========================================================
EPISODE  353
Score of total rewards: 11
========================================================
EPISODE  354
Score of total rewards: 4
========================================================
EPISODE  355
Score of total rewards: 5
========================================================
EPISODE  356
Score of total rewards: 8
========================================================
EPISODE  357
Score of total rewards: 5
========================================================
EPISODE  358
Score of total rewards: 10
========================================================
EPISODE  359
Score of total rewards: 10
========================================================
EPISODE  360
Score of total rewards: 8
========================================================
EPISODE  361
Score of total rewards: 9
========================================================
EPISODE  362
Score of total rewards: 6
========================================================
EPISODE  363
Score of total rewards: 11
========================================================
EPISODE  364
Score of total rewards: 11
```

```
========================================================
EPISODE  365
Score of total rewards: 13
========================================================
EPISODE  366
Score of total rewards: 5
========================================================
EPISODE  367
Score of total rewards: 10
========================================================
EPISODE  368
Score of total rewards: 6
========================================================
EPISODE  369
Score of total rewards: 5
========================================================
EPISODE  370
Score of total rewards: 6
========================================================
EPISODE  371
Score of total rewards: 8
========================================================
EPISODE  372
Score of total rewards: 6
========================================================
EPISODE  373
Score of total rewards: 6
========================================================
EPISODE  374
Score of total rewards: 9
========================================================
EPISODE  375
Score of total rewards: 6
========================================================
EPISODE  376
Score of total rewards: 10
========================================================
EPISODE  377
Score of total rewards: 6
========================================================
EPISODE  378
Score of total rewards: 10
========================================================
EPISODE  379
Score of total rewards: 9
========================================================
EPISODE  380
Score of total rewards: 11
========================================================
EPISODE  381
Score of total rewards: 7
========================================================
EPISODE  382
Score of total rewards: 12
========================================================
EPISODE  383
Score of total rewards: 8
```

```
============================================================
EPISODE  384
Score of total rewards: 9
============================================================
EPISODE  385
Score of total rewards: 9
============================================================
EPISODE  386
Score of total rewards: 8
============================================================
EPISODE  387
Score of total rewards: 8
============================================================
EPISODE  388
Score of total rewards: 10
============================================================
EPISODE  389
Score of total rewards: 8
============================================================
EPISODE  390
Score of total rewards: 11
============================================================
EPISODE  391
Score of total rewards: 7
============================================================
EPISODE  392
Score of total rewards: 5
============================================================
EPISODE  393
Score of total rewards: 8
============================================================
EPISODE  394
Score of total rewards: 6
============================================================
EPISODE  395
Score of total rewards: 10
============================================================
EPISODE  396
Score of total rewards: 7
============================================================
EPISODE  397
Score of total rewards: 13
============================================================
EPISODE  398
Score of total rewards: 11
============================================================
EPISODE  399
Score of total rewards: 9
============================================================
EPISODE  400
Score of total rewards: 9
============================================================
EPISODE  401
Score of total rewards: 9
============================================================
EPISODE  402
Score of total rewards: 11
```

```
========================================================
EPISODE  403
Score of total rewards: 9
========================================================
EPISODE  404
Score of total rewards: 12
========================================================
EPISODE  405
Score of total rewards: 6
========================================================
EPISODE  406
Score of total rewards: 9
========================================================
EPISODE  407
Score of total rewards: 7
========================================================
EPISODE  408
Score of total rewards: 5
========================================================
EPISODE  409
Score of total rewards: 8
========================================================
EPISODE  410
Score of total rewards: 9
========================================================
EPISODE  411
Score of total rewards: 10
========================================================
EPISODE  412
Score of total rewards: 6
========================================================
EPISODE  413
Score of total rewards: 9
========================================================
EPISODE  414
Score of total rewards: 8
========================================================
EPISODE  415
Score of total rewards: 5
========================================================
EPISODE  416
Score of total rewards: 5
========================================================
EPISODE  417
Score of total rewards: 5
========================================================
EPISODE  418
Score of total rewards: 12
========================================================
EPISODE  419
Score of total rewards: 10
========================================================
EPISODE  420
Score of total rewards: 6
========================================================
EPISODE  421
Score of total rewards: 7
```

```
============================================================
EPISODE  422
Score of total rewards: 6
============================================================
EPISODE  423
Score of total rewards: 3
============================================================
EPISODE  424
Score of total rewards: 8
============================================================
EPISODE  425
Score of total rewards: 12
============================================================
EPISODE  426
Score of total rewards: 11
============================================================
EPISODE  427
Score of total rewards: 6
============================================================
EPISODE  428
Score of total rewards: 10
============================================================
EPISODE  429
Score of total rewards: 9
============================================================
EPISODE  430
Score of total rewards: 8
============================================================
EPISODE  431
Score of total rewards: 6
============================================================
EPISODE  432
Score of total rewards: 8
============================================================
EPISODE  433
Score of total rewards: 7
============================================================
EPISODE  434
Score of total rewards: 4
============================================================
EPISODE  435
Score of total rewards: 14
============================================================
EPISODE  436
Score of total rewards: 10
============================================================
EPISODE  437
Score of total rewards: 8
============================================================
EPISODE  438
Score of total rewards: 14
============================================================
EPISODE  439
Score of total rewards: 6
============================================================
EPISODE  440
Score of total rewards: 6
```

```
============================================================
EPISODE  441
Score of total rewards: 4
============================================================
EPISODE  442
Score of total rewards: 13
============================================================
EPISODE  443
Score of total rewards: 5
============================================================
EPISODE  444
Score of total rewards: 7
============================================================
EPISODE  445
Score of total rewards: 12
============================================================
EPISODE  446
Score of total rewards: 9
============================================================
EPISODE  447
Score of total rewards: 10
============================================================
EPISODE  448
Score of total rewards: 7
============================================================
EPISODE  449
Score of total rewards: 12
============================================================
EPISODE  450
Score of total rewards: 10
============================================================
EPISODE  451
Score of total rewards: 7
============================================================
EPISODE  452
Score of total rewards: 5
============================================================
EPISODE  453
Score of total rewards: 8
============================================================
EPISODE  454
Score of total rewards: 5
============================================================
EPISODE  455
Score of total rewards: 7
============================================================
EPISODE  456
Score of total rewards: 9
============================================================
EPISODE  457
Score of total rewards: 11
============================================================
EPISODE  458
Score of total rewards: 7
============================================================
EPISODE  459
Score of total rewards: 9
```

```
========================================================
EPISODE  460
Score of total rewards: 9
========================================================
EPISODE  461
Score of total rewards: 8
========================================================
EPISODE  462
Score of total rewards: 10
========================================================
EPISODE  463
Score of total rewards: 9
========================================================
EPISODE  464
Score of total rewards: 6
========================================================
EPISODE  465
Score of total rewards: 6
========================================================
EPISODE  466
Score of total rewards: 5
========================================================
EPISODE  467
Score of total rewards: 5
========================================================
EPISODE  468
Score of total rewards: 9
========================================================
EPISODE  469
Score of total rewards: 10
========================================================
EPISODE  470
Score of total rewards: 10
========================================================
EPISODE  471
Score of total rewards: 9
========================================================
EPISODE  472
Score of total rewards: 10
========================================================
EPISODE  473
Score of total rewards: 10
========================================================
EPISODE  474
Score of total rewards: 3
========================================================
EPISODE  475
Score of total rewards: 8
========================================================
EPISODE  476
Score of total rewards: 11
========================================================
EPISODE  477
Score of total rewards: 8
========================================================
EPISODE  478
Score of total rewards: 7
```

```
========================================================
EPISODE  479
Score of total rewards: 7
========================================================
EPISODE  480
Score of total rewards: 8
========================================================
EPISODE  481
Score of total rewards: 9
========================================================
EPISODE  482
Score of total rewards: 15
========================================================
EPISODE  483
Score of total rewards: 5
========================================================
EPISODE  484
Score of total rewards: 14
========================================================
EPISODE  485
Score of total rewards: 5
========================================================
EPISODE  486
Score of total rewards: 11
========================================================
EPISODE  487
Score of total rewards: 6
========================================================
EPISODE  488
Score of total rewards: 6
========================================================
EPISODE  489
Score of total rewards: 5
========================================================
EPISODE  490
Score of total rewards: 11
========================================================
EPISODE  491
Score of total rewards: 11
========================================================
EPISODE  492
Score of total rewards: 10
========================================================
EPISODE  493
Score of total rewards: 15
========================================================
EPISODE  494
Score of total rewards: 5
========================================================
EPISODE  495
Score of total rewards: 7
========================================================
EPISODE  496
Score of total rewards: 11
========================================================
EPISODE  497
Score of total rewards: 14
```

```
========================================================
EPISODE  498
Score of total rewards: 9
========================================================
EPISODE  499
Score of total rewards: 3
========================================================
EPISODE  500
Score of total rewards: 11
========================================================
EPISODE  501
Score of total rewards: 10
========================================================
EPISODE  502
Score of total rewards: 8
========================================================
EPISODE  503
Score of total rewards: 13
========================================================
EPISODE  504
Score of total rewards: 3
========================================================
EPISODE  505
Score of total rewards: 5
========================================================
EPISODE  506
Score of total rewards: 4
========================================================
EPISODE  507
Score of total rewards: 10
========================================================
EPISODE  508
Score of total rewards: 5
========================================================
EPISODE  509
Score of total rewards: 11
========================================================
EPISODE  510
Score of total rewards: 10
========================================================
EPISODE  511
Score of total rewards: 8
========================================================
EPISODE  512
Score of total rewards: 6
========================================================
EPISODE  513
Score of total rewards: 8
========================================================
EPISODE  514
Score of total rewards: 8
========================================================
EPISODE  515
Score of total rewards: 8
========================================================
EPISODE  516
Score of total rewards: 7
```

```
========================================================
EPISODE  517
Score of total rewards: 7
========================================================
EPISODE  518
Score of total rewards: 3
========================================================
EPISODE  519
Score of total rewards: 3
========================================================
EPISODE  520
Score of total rewards: 9
========================================================
EPISODE  521
Score of total rewards: 4
========================================================
EPISODE  522
Score of total rewards: 6
========================================================
EPISODE  523
Score of total rewards: 6
========================================================
EPISODE  524
Score of total rewards: 9
========================================================
EPISODE  525
Score of total rewards: 8
========================================================
EPISODE  526
Score of total rewards: 13
========================================================
EPISODE  527
Score of total rewards: 5
========================================================
EPISODE  528
Score of total rewards: 6
========================================================
EPISODE  529
Score of total rewards: 5
========================================================
EPISODE  530
Score of total rewards: 12
========================================================
EPISODE  531
Score of total rewards: 4
========================================================
EPISODE  532
Score of total rewards: 4
========================================================
EPISODE  533
Score of total rewards: 8
========================================================
EPISODE  534
Score of total rewards: 7
========================================================
EPISODE  535
Score of total rewards: 4
```

```
========================================================
EPISODE  536
Score of total rewards: 5
========================================================
EPISODE  537
Score of total rewards: 6
========================================================
EPISODE  538
Score of total rewards: 14
========================================================
EPISODE  539
Score of total rewards: 7
========================================================
EPISODE  540
Score of total rewards: 8
========================================================
EPISODE  541
Score of total rewards: 12
========================================================
EPISODE  542
Score of total rewards: 7
========================================================
EPISODE  543
Score of total rewards: 7
========================================================
EPISODE  544
Score of total rewards: 7
========================================================
EPISODE  545
Score of total rewards: 3
========================================================
EPISODE  546
Score of total rewards: 11
========================================================
EPISODE  547
Score of total rewards: 10
========================================================
EPISODE  548
Score of total rewards: 9
========================================================
EPISODE  549
Score of total rewards: 10
========================================================
EPISODE  550
Score of total rewards: 10
========================================================
EPISODE  551
Score of total rewards: 10
========================================================
EPISODE  552
Score of total rewards: 5
========================================================
EPISODE  553
Score of total rewards: 5
========================================================
EPISODE  554
Score of total rewards: 4
```

```
============================================================
EPISODE  555
Score of total rewards: 8
============================================================
EPISODE  556
Score of total rewards: 4
============================================================
EPISODE  557
Score of total rewards: 5
============================================================
EPISODE  558
Score of total rewards: 5
============================================================
EPISODE  559
Score of total rewards: 6
============================================================
EPISODE  560
Score of total rewards: 11
============================================================
EPISODE  561
Score of total rewards: 12
============================================================
EPISODE  562
Score of total rewards: 7
============================================================
EPISODE  563
Score of total rewards: 8
============================================================
EPISODE  564
Score of total rewards: 8
============================================================
EPISODE  565
Score of total rewards: 11
============================================================
EPISODE  566
Score of total rewards: 5
============================================================
EPISODE  567
Score of total rewards: 8
============================================================
EPISODE  568
Score of total rewards: 12
============================================================
EPISODE  569
Score of total rewards: 10
============================================================
EPISODE  570
Score of total rewards: 10
============================================================
EPISODE  571
Score of total rewards: 5
============================================================
EPISODE  572
Score of total rewards: 5
============================================================
EPISODE  573
Score of total rewards: 7
```

```
========================================================
EPISODE  574
Score of total rewards: 9
========================================================
EPISODE  575
Score of total rewards: 10
========================================================
EPISODE  576
Score of total rewards: 10
========================================================
EPISODE  577
Score of total rewards: 6
========================================================
EPISODE  578
Score of total rewards: 8
========================================================
EPISODE  579
Score of total rewards: 6
========================================================
EPISODE  580
Score of total rewards: 7
========================================================
EPISODE  581
Score of total rewards: 6
========================================================
EPISODE  582
Score of total rewards: 3
========================================================
EPISODE  583
Score of total rewards: 6
========================================================
EPISODE  584
Score of total rewards: 11
========================================================
EPISODE  585
Score of total rewards: 7
========================================================
EPISODE  586
Score of total rewards: 7
========================================================
EPISODE  587
Score of total rewards: 7
========================================================
EPISODE  588
Score of total rewards: 8
========================================================
EPISODE  589
Score of total rewards: 11
========================================================
EPISODE  590
Score of total rewards: 9
========================================================
EPISODE  591
Score of total rewards: 11
========================================================
EPISODE  592
Score of total rewards: 8
```

```
========================================================
EPISODE  593
Score of total rewards: 4
========================================================
EPISODE  594
Score of total rewards: 9
========================================================
EPISODE  595
Score of total rewards: 5
========================================================
EPISODE  596
Score of total rewards: 8
========================================================
EPISODE  597
Score of total rewards: 7
========================================================
EPISODE  598
Score of total rewards: 9
========================================================
EPISODE  599
Score of total rewards: 8
========================================================
EPISODE  600
Score of total rewards: 10
========================================================
EPISODE  601
Score of total rewards: 11
========================================================
EPISODE  602
Score of total rewards: 11
========================================================
EPISODE  603
Score of total rewards: 10
========================================================
EPISODE  604
Score of total rewards: 9
========================================================
EPISODE  605
Score of total rewards: 12
========================================================
EPISODE  606
Score of total rewards: 13
========================================================
EPISODE  607
Score of total rewards: 7
========================================================
EPISODE  608
Score of total rewards: 10
========================================================
EPISODE  609
Score of total rewards: 7
========================================================
EPISODE  610
Score of total rewards: 5
========================================================
EPISODE  611
Score of total rewards: 7
```

```
============================================================
EPISODE  612
Score of total rewards: 5
============================================================
EPISODE  613
Score of total rewards: 10
============================================================
EPISODE  614
Score of total rewards: 10
============================================================
EPISODE  615
Score of total rewards: 9
============================================================
EPISODE  616
Score of total rewards: 7
============================================================
EPISODE  617
Score of total rewards: 10
============================================================
EPISODE  618
Score of total rewards: 10
============================================================
EPISODE  619
Score of total rewards: 7
============================================================
EPISODE  620
Score of total rewards: 11
============================================================
EPISODE  621
Score of total rewards: 10
============================================================
EPISODE  622
Score of total rewards: 9
============================================================
EPISODE  623
Score of total rewards: 11
============================================================
EPISODE  624
Score of total rewards: 10
============================================================
EPISODE  625
Score of total rewards: 10
============================================================
EPISODE  626
Score of total rewards: 7
============================================================
EPISODE  627
Score of total rewards: 5
============================================================
EPISODE  628
Score of total rewards: 6
============================================================
EPISODE  629
Score of total rewards: 10
============================================================
EPISODE  630
Score of total rewards: 10
```

```
========================================================
EPISODE  631
Score of total rewards: 6
========================================================
EPISODE  632
Score of total rewards: 12
========================================================
EPISODE  633
Score of total rewards: 9
========================================================
EPISODE  634
Score of total rewards: 8
========================================================
EPISODE  635
Score of total rewards: 9
========================================================
EPISODE  636
Score of total rewards: 8
========================================================
EPISODE  637
Score of total rewards: 10
========================================================
EPISODE  638
Score of total rewards: 14
========================================================
EPISODE  639
Score of total rewards: 5
========================================================
EPISODE  640
Score of total rewards: 7
========================================================
EPISODE  641
Score of total rewards: 6
========================================================
EPISODE  642
Score of total rewards: 7
========================================================
EPISODE  643
Score of total rewards: 6
========================================================
EPISODE  644
Score of total rewards: 9
========================================================
EPISODE  645
Score of total rewards: 7
========================================================
EPISODE  646
Score of total rewards: 9
========================================================
EPISODE  647
Score of total rewards: 11
========================================================
EPISODE  648
Score of total rewards: 9
========================================================
EPISODE  649
Score of total rewards: 8
```

```
========================================================
EPISODE  650
Score of total rewards: 12
========================================================
EPISODE  651
Score of total rewards: 11
========================================================
EPISODE  652
Score of total rewards: 6
========================================================
EPISODE  653
Score of total rewards: 11
========================================================
EPISODE  654
Score of total rewards: 6
========================================================
EPISODE  655
Score of total rewards: 9
========================================================
EPISODE  656
Score of total rewards: 5
========================================================
EPISODE  657
Score of total rewards: 9
========================================================
EPISODE  658
Score of total rewards: 11
========================================================
EPISODE  659
Score of total rewards: 11
========================================================
EPISODE  660
Score of total rewards: 8
========================================================
EPISODE  661
Score of total rewards: 7
========================================================
EPISODE  662
Score of total rewards: 4
========================================================
EPISODE  663
Score of total rewards: 6
========================================================
EPISODE  664
Score of total rewards: 13
========================================================
EPISODE  665
Score of total rewards: 13
========================================================
EPISODE  666
Score of total rewards: 10
========================================================
EPISODE  667
Score of total rewards: 7
========================================================
EPISODE  668
Score of total rewards: 12
```

```
============================================================
EPISODE  669
Score of total rewards: 15
============================================================
EPISODE  670
Score of total rewards: 12
============================================================
EPISODE  671
Score of total rewards: 8
============================================================
EPISODE  672
Score of total rewards: 10
============================================================
EPISODE  673
Score of total rewards: 8
============================================================
EPISODE  674
Score of total rewards: 4
============================================================
EPISODE  675
Score of total rewards: 8
============================================================
EPISODE  676
Score of total rewards: 10
============================================================
EPISODE  677
Score of total rewards: 6
============================================================
EPISODE  678
Score of total rewards: 9
============================================================
EPISODE  679
Score of total rewards: 7
============================================================
EPISODE  680
Score of total rewards: 5
============================================================
EPISODE  681
Score of total rewards: 7
============================================================
EPISODE  682
Score of total rewards: 8
============================================================
EPISODE  683
Score of total rewards: 5
============================================================
EPISODE  684
Score of total rewards: 6
============================================================
EPISODE  685
Score of total rewards: 3
============================================================
EPISODE  686
Score of total rewards: 7
============================================================
EPISODE  687
Score of total rewards: 5
```

```
========================================================
EPISODE  688
Score of total rewards: 7
========================================================
EPISODE  689
Score of total rewards: 8
========================================================
EPISODE  690
Score of total rewards: 11
========================================================
EPISODE  691
Score of total rewards: 5
========================================================
EPISODE  692
Score of total rewards: 9
========================================================
EPISODE  693
Score of total rewards: 10
========================================================
EPISODE  694
Score of total rewards: 6
========================================================
EPISODE  695
Score of total rewards: 8
========================================================
EPISODE  696
Score of total rewards: 7
========================================================
EPISODE  697
Score of total rewards: 4
========================================================
EPISODE  698
Score of total rewards: 6
========================================================
EPISODE  699
Score of total rewards: 6
========================================================
EPISODE  700
Score of total rewards: 12
========================================================
EPISODE  701
Score of total rewards: 9
========================================================
EPISODE  702
Score of total rewards: 9
========================================================
EPISODE  703
Score of total rewards: 8
========================================================
EPISODE  704
Score of total rewards: 7
========================================================
EPISODE  705
Score of total rewards: 6
========================================================
EPISODE  706
Score of total rewards: 3
```

```
========================================================
EPISODE  707
Score of total rewards: 10
========================================================
EPISODE  708
Score of total rewards: 8
========================================================
EPISODE  709
Score of total rewards: 9
========================================================
EPISODE  710
Score of total rewards: 11
========================================================
EPISODE  711
Score of total rewards: 8
========================================================
EPISODE  712
Score of total rewards: 10
========================================================
EPISODE  713
Score of total rewards: 7
========================================================
EPISODE  714
Score of total rewards: 4
========================================================
EPISODE  715
Score of total rewards: 15
========================================================
EPISODE  716
Score of total rewards: 8
========================================================
EPISODE  717
Score of total rewards: 12
========================================================
EPISODE  718
Score of total rewards: 11
========================================================
EPISODE  719
Score of total rewards: 7
========================================================
EPISODE  720
Score of total rewards: 11
========================================================
EPISODE  721
Score of total rewards: 10
========================================================
EPISODE  722
Score of total rewards: 5
========================================================
EPISODE  723
Score of total rewards: 6
========================================================
EPISODE  724
Score of total rewards: 7
========================================================
EPISODE  725
Score of total rewards: 8
```

```
============================================================
EPISODE  726
Score of total rewards: 8
============================================================
EPISODE  727
Score of total rewards: 6
============================================================
EPISODE  728
Score of total rewards: 9
============================================================
EPISODE  729
Score of total rewards: 8
============================================================
EPISODE  730
Score of total rewards: 10
============================================================
EPISODE  731
Score of total rewards: 7
============================================================
EPISODE  732
Score of total rewards: 6
============================================================
EPISODE  733
Score of total rewards: 6
============================================================
EPISODE  734
Score of total rewards: 13
============================================================
EPISODE  735
Score of total rewards: 12
============================================================
EPISODE  736
Score of total rewards: 8
============================================================
EPISODE  737
Score of total rewards: 3
============================================================
EPISODE  738
Score of total rewards: 10
============================================================
EPISODE  739
Score of total rewards: 7
============================================================
EPISODE  740
Score of total rewards: 8
============================================================
EPISODE  741
Score of total rewards: 10
============================================================
EPISODE  742
Score of total rewards: 8
============================================================
EPISODE  743
Score of total rewards: 5
============================================================
EPISODE  744
Score of total rewards: 10
```

```
========================================================
EPISODE  745
Score of total rewards: 9
========================================================
EPISODE  746
Score of total rewards: 8
========================================================
EPISODE  747
Score of total rewards: 7
========================================================
EPISODE  748
Score of total rewards: 12
========================================================
EPISODE  749
Score of total rewards: 12
========================================================
EPISODE  750
Score of total rewards: 8
========================================================
EPISODE  751
Score of total rewards: 10
========================================================
EPISODE  752
Score of total rewards: 9
========================================================
EPISODE  753
Score of total rewards: 7
========================================================
EPISODE  754
Score of total rewards: 4
========================================================
EPISODE  755
Score of total rewards: 10
========================================================
EPISODE  756
Score of total rewards: 9
========================================================
EPISODE  757
Score of total rewards: 10
========================================================
EPISODE  758
Score of total rewards: 7
========================================================
EPISODE  759
Score of total rewards: 5
========================================================
EPISODE  760
Score of total rewards: 6
========================================================
EPISODE  761
Score of total rewards: 10
========================================================
EPISODE  762
Score of total rewards: 5
========================================================
EPISODE  763
Score of total rewards: 5
```

```
========================================================
EPISODE  764
Score of total rewards: 10
========================================================
EPISODE  765
Score of total rewards: 5
========================================================
EPISODE  766
Score of total rewards: 8
========================================================
EPISODE  767
Score of total rewards: 6
========================================================
EPISODE  768
Score of total rewards: 7
========================================================
EPISODE  769
Score of total rewards: 5
========================================================
EPISODE  770
Score of total rewards: 13
========================================================
EPISODE  771
Score of total rewards: 8
========================================================
EPISODE  772
Score of total rewards: 6
========================================================
EPISODE  773
Score of total rewards: 9
========================================================
EPISODE  774
Score of total rewards: 9
========================================================
EPISODE  775
Score of total rewards: 14
========================================================
EPISODE  776
Score of total rewards: 7
========================================================
EPISODE  777
Score of total rewards: 9
========================================================
EPISODE  778
Score of total rewards: 7
========================================================
EPISODE  779
Score of total rewards: 10
========================================================
EPISODE  780
Score of total rewards: 11
========================================================
EPISODE  781
Score of total rewards: 7
========================================================
EPISODE  782
Score of total rewards: 7
```

```
========================================================
EPISODE  783
Score of total rewards: 7
========================================================
EPISODE  784
Score of total rewards: 5
========================================================
EPISODE  785
Score of total rewards: 10
========================================================
EPISODE  786
Score of total rewards: 13
========================================================
EPISODE  787
Score of total rewards: 6
========================================================
EPISODE  788
Score of total rewards: 5
========================================================
EPISODE  789
Score of total rewards: 8
========================================================
EPISODE  790
Score of total rewards: 8
========================================================
EPISODE  791
Score of total rewards: 7
========================================================
EPISODE  792
Score of total rewards: 8
========================================================
EPISODE  793
Score of total rewards: 4
========================================================
EPISODE  794
Score of total rewards: 9
========================================================
EPISODE  795
Score of total rewards: 6
========================================================
EPISODE  796
Score of total rewards: 9
========================================================
EPISODE  797
Score of total rewards: 3
========================================================
EPISODE  798
Score of total rewards: 10
========================================================
EPISODE  799
Score of total rewards: 9
========================================================
EPISODE  800
Score of total rewards: 6
========================================================
EPISODE  801
Score of total rewards: 11
```

```
========================================================
EPISODE  802
Score of total rewards: 10
========================================================
EPISODE  803
Score of total rewards: 3
========================================================
EPISODE  804
Score of total rewards: 10
========================================================
EPISODE  805
Score of total rewards: 3
========================================================
EPISODE  806
Score of total rewards: 9
========================================================
EPISODE  807
Score of total rewards: 13
========================================================
EPISODE  808
Score of total rewards: 7
========================================================
EPISODE  809
Score of total rewards: 3
========================================================
EPISODE  810
Score of total rewards: 10
========================================================
EPISODE  811
Score of total rewards: 5
========================================================
EPISODE  812
Score of total rewards: 11
========================================================
EPISODE  813
Score of total rewards: 6
========================================================
EPISODE  814
Score of total rewards: 6
========================================================
EPISODE  815
Score of total rewards: 7
========================================================
EPISODE  816
Score of total rewards: 8
========================================================
EPISODE  817
Score of total rewards: 4
========================================================
EPISODE  818
Score of total rewards: 7
========================================================
EPISODE  819
Score of total rewards: 11
========================================================
EPISODE  820
Score of total rewards: 8
```

```
========================================================
EPISODE  821
Score of total rewards: 11
========================================================
EPISODE  822
Score of total rewards: 11
========================================================
EPISODE  823
Score of total rewards: 7
========================================================
EPISODE  824
Score of total rewards: 6
========================================================
EPISODE  825
Score of total rewards: 4
========================================================
EPISODE  826
Score of total rewards: 9
========================================================
EPISODE  827
Score of total rewards: 4
========================================================
EPISODE  828
Score of total rewards: 7
========================================================
EPISODE  829
Score of total rewards: 3
========================================================
EPISODE  830
Score of total rewards: 7
========================================================
EPISODE  831
Score of total rewards: 5
========================================================
EPISODE  832
Score of total rewards: 8
========================================================
EPISODE  833
Score of total rewards: 9
========================================================
EPISODE  834
Score of total rewards: 10
========================================================
EPISODE  835
Score of total rewards: 9
========================================================
EPISODE  836
Score of total rewards: 8
========================================================
EPISODE  837
Score of total rewards: 10
========================================================
EPISODE  838
Score of total rewards: 11
========================================================
EPISODE  839
Score of total rewards: 7
```

```
========================================================
EPISODE  840
Score of total rewards: 8
========================================================
EPISODE  841
Score of total rewards: 4
========================================================
EPISODE  842
Score of total rewards: 12
========================================================
EPISODE  843
Score of total rewards: 7
========================================================
EPISODE  844
Score of total rewards: 13
========================================================
EPISODE  845
Score of total rewards: 6
========================================================
EPISODE  846
Score of total rewards: 9
========================================================
EPISODE  847
Score of total rewards: 8
========================================================
EPISODE  848
Score of total rewards: 7
========================================================
EPISODE  849
Score of total rewards: 5
========================================================
EPISODE  850
Score of total rewards: 8
========================================================
EPISODE  851
Score of total rewards: 11
========================================================
EPISODE  852
Score of total rewards: 7
========================================================
EPISODE  853
Score of total rewards: 9
========================================================
EPISODE  854
Score of total rewards: 11
========================================================
EPISODE  855
Score of total rewards: 3
========================================================
EPISODE  856
Score of total rewards: 4
========================================================
EPISODE  857
Score of total rewards: 7
========================================================
EPISODE  858
Score of total rewards: 9
```

```
========================================================
EPISODE  859
Score of total rewards: 3
========================================================
EPISODE  860
Score of total rewards: 8
========================================================
EPISODE  861
Score of total rewards: 6
========================================================
EPISODE  862
Score of total rewards: 8
========================================================
EPISODE  863
Score of total rewards: 9
========================================================
EPISODE  864
Score of total rewards: 4
========================================================
EPISODE  865
Score of total rewards: 6
========================================================
EPISODE  866
Score of total rewards: 4
========================================================
EPISODE  867
Score of total rewards: 9
========================================================
EPISODE  868
Score of total rewards: 7
========================================================
EPISODE  869
Score of total rewards: 9
========================================================
EPISODE  870
Score of total rewards: 7
========================================================
EPISODE  871
Score of total rewards: 5
========================================================
EPISODE  872
Score of total rewards: 6
========================================================
EPISODE  873
Score of total rewards: 10
========================================================
EPISODE  874
Score of total rewards: 8
========================================================
EPISODE  875
Score of total rewards: 12
========================================================
EPISODE  876
Score of total rewards: 9
========================================================
EPISODE  877
Score of total rewards: 4
```

```
========================================================
EPISODE  878
Score of total rewards: 7
========================================================
EPISODE  879
Score of total rewards: 5
========================================================
EPISODE  880
Score of total rewards: 6
========================================================
EPISODE  881
Score of total rewards: 3
========================================================
EPISODE  882
Score of total rewards: 6
========================================================
EPISODE  883
Score of total rewards: 10
========================================================
EPISODE  884
Score of total rewards: 13
========================================================
EPISODE  885
Score of total rewards: 9
========================================================
EPISODE  886
Score of total rewards: 7
========================================================
EPISODE  887
Score of total rewards: 7
========================================================
EPISODE  888
Score of total rewards: 4
========================================================
EPISODE  889
Score of total rewards: 10
========================================================
EPISODE  890
Score of total rewards: 9
========================================================
EPISODE  891
Score of total rewards: 8
========================================================
EPISODE  892
Score of total rewards: 7
========================================================
EPISODE  893
Score of total rewards: 5
========================================================
EPISODE  894
Score of total rewards: 6
========================================================
EPISODE  895
Score of total rewards: 10
========================================================
EPISODE  896
Score of total rewards: 6
```

```
========================================================
EPISODE  897
Score of total rewards: 8
========================================================
EPISODE  898
Score of total rewards: 8
========================================================
EPISODE  899
Score of total rewards: 9
========================================================
EPISODE  900
Score of total rewards: 12
========================================================
EPISODE  901
Score of total rewards: 10
========================================================
EPISODE  902
Score of total rewards: 8
========================================================
EPISODE  903
Score of total rewards: 7
========================================================
EPISODE  904
Score of total rewards: 6
========================================================
EPISODE  905
Score of total rewards: 8
========================================================
EPISODE  906
Score of total rewards: 6
========================================================
EPISODE  907
Score of total rewards: 4
========================================================
EPISODE  908
Score of total rewards: 6
========================================================
EPISODE  909
Score of total rewards: 6
========================================================
EPISODE  910
Score of total rewards: 7
========================================================
EPISODE  911
Score of total rewards: 12
========================================================
EPISODE  912
Score of total rewards: 11
========================================================
EPISODE  913
Score of total rewards: 8
========================================================
EPISODE  914
Score of total rewards: 6
========================================================
EPISODE  915
Score of total rewards: 11
```

```
========================================================
EPISODE  916
Score of total rewards: 9
========================================================
EPISODE  917
Score of total rewards: 5
========================================================
EPISODE  918
Score of total rewards: 6
========================================================
EPISODE  919
Score of total rewards: 6
========================================================
EPISODE  920
Score of total rewards: 4
========================================================
EPISODE  921
Score of total rewards: 5
========================================================
EPISODE  922
Score of total rewards: 11
========================================================
EPISODE  923
Score of total rewards: 3
========================================================
EPISODE  924
Score of total rewards: 13
========================================================
EPISODE  925
Score of total rewards: 10
========================================================
EPISODE  926
Score of total rewards: 9
========================================================
EPISODE  927
Score of total rewards: 12
========================================================
EPISODE  928
Score of total rewards: 15
========================================================
EPISODE  929
Score of total rewards: 6
========================================================
EPISODE  930
Score of total rewards: 7
========================================================
EPISODE  931
Score of total rewards: 12
========================================================
EPISODE  932
Score of total rewards: 4
========================================================
EPISODE  933
Score of total rewards: 5
========================================================
EPISODE  934
Score of total rewards: 7
```

```
========================================================
EPISODE  935
Score of total rewards: 4
========================================================
EPISODE  936
Score of total rewards: 8
========================================================
EPISODE  937
Score of total rewards: 11
========================================================
EPISODE  938
Score of total rewards: 7
========================================================
EPISODE  939
Score of total rewards: 8
========================================================
EPISODE  940
Score of total rewards: 13
========================================================
EPISODE  941
Score of total rewards: 6
========================================================
EPISODE  942
Score of total rewards: 9
========================================================
EPISODE  943
Score of total rewards: 7
========================================================
EPISODE  944
Score of total rewards: 7
========================================================
EPISODE  945
Score of total rewards: 7
========================================================
EPISODE  946
Score of total rewards: 8
========================================================
EPISODE  947
Score of total rewards: 9
========================================================
EPISODE  948
Score of total rewards: 12
========================================================
EPISODE  949
Score of total rewards: 7
========================================================
EPISODE  950
Score of total rewards: 9
========================================================
EPISODE  951
Score of total rewards: 9
========================================================
EPISODE  952
Score of total rewards: 9
========================================================
EPISODE  953
Score of total rewards: 10
```

```
============================================================
EPISODE  954
Score of total rewards: 4
============================================================
EPISODE  955
Score of total rewards: 15
============================================================
EPISODE  956
Score of total rewards: 7
============================================================
EPISODE  957
Score of total rewards: 4
============================================================
EPISODE  958
Score of total rewards: 7
============================================================
EPISODE  959
Score of total rewards: 8
============================================================
EPISODE  960
Score of total rewards: 11
============================================================
EPISODE  961
Score of total rewards: 10
============================================================
EPISODE  962
Score of total rewards: 10
============================================================
EPISODE  963
Score of total rewards: 8
============================================================
EPISODE  964
Score of total rewards: 8
============================================================
EPISODE  965
Score of total rewards: 3
============================================================
EPISODE  966
Score of total rewards: 8
============================================================
EPISODE  967
Score of total rewards: 4
============================================================
EPISODE  968
Score of total rewards: 14
============================================================
EPISODE  969
Score of total rewards: 7
============================================================
EPISODE  970
Score of total rewards: 7
============================================================
EPISODE  971
Score of total rewards: 6
============================================================
EPISODE  972
Score of total rewards: 10
```

```
========================================================
EPISODE  973
Score of total rewards: 7
========================================================
EPISODE  974
Score of total rewards: 7
========================================================
EPISODE  975
Score of total rewards: 6
========================================================
EPISODE  976
Score of total rewards: 11
========================================================
EPISODE  977
Score of total rewards: 5
========================================================
EPISODE  978
Score of total rewards: 11
========================================================
EPISODE  979
Score of total rewards: 7
========================================================
EPISODE  980
Score of total rewards: 7
========================================================
EPISODE  981
Score of total rewards: 9
========================================================
EPISODE  982
Score of total rewards: 7
========================================================
EPISODE  983
Score of total rewards: 5
========================================================
EPISODE  984
Score of total rewards: 6
========================================================
EPISODE  985
Score of total rewards: 5
========================================================
EPISODE  986
Score of total rewards: 10
========================================================
EPISODE  987
Score of total rewards: 7
========================================================
EPISODE  988
Score of total rewards: 5
========================================================
EPISODE  989
Score of total rewards: 8
========================================================
EPISODE  990
Score of total rewards: 11
========================================================
EPISODE  991
Score of total rewards: 12
```

```
========================================================
EPISODE  992
Score of total rewards: 8
========================================================
EPISODE  993
Score of total rewards: 10
========================================================
EPISODE  994
Score of total rewards: 8
========================================================
EPISODE  995
Score of total rewards: 6
========================================================
EPISODE  996
Score of total rewards: 8
========================================================
EPISODE  997
Score of total rewards: 9
========================================================
EPISODE  998
Score of total rewards: 9
========================================================
EPISODE  999
Score of total rewards: 7
Average score over time: 8.007
```

In [48]:
```python
# only to have a look of the test play result
my_q_agent_controlled.test(test_episodes=5, display=True)
```


Taxi-v3 | Episode: 4 | Step: 14
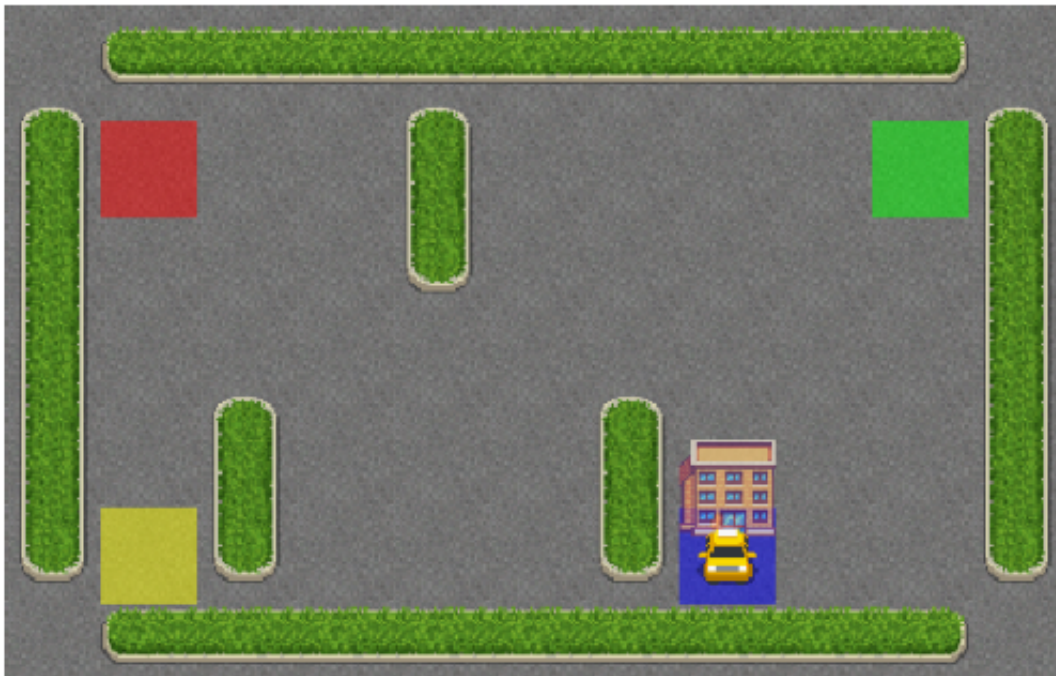
```
Score of total rewards: 7
Average score over time: 6.8
```

Out[48]: [7, 7, 4, 9, 7]
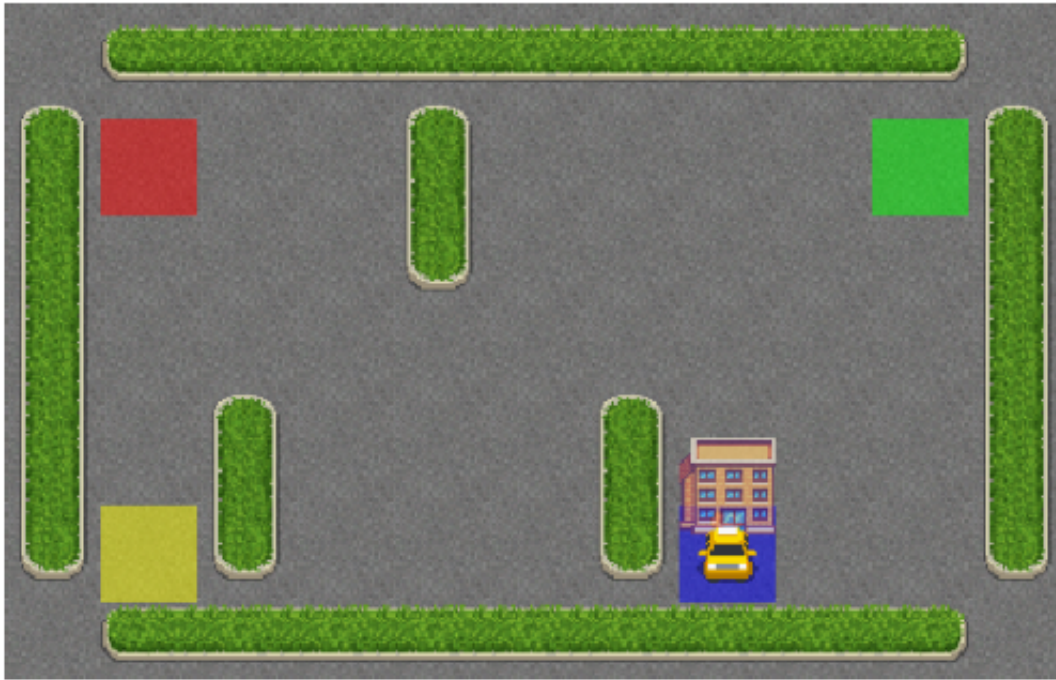
```
In [49]: # save to GIF
         my_q_agent_controlled.test_to_gif(filename="taxi_q_controlled",episodes=5
```

```
episodes: 0, acton: 0, state: 183
episodes: 0, acton: 3, state: 283
episodes: 0, acton: 3, state: 263
episodes: 0, acton: 3, state: 243
episodes: 0, acton: 1, state: 223
episodes: 0, acton: 3, state: 123
episodes: 0, acton: 1, state: 103
episodes: 0, acton: 4, state: 3
episodes: 0, acton: 2, state: 19
episodes: 0, acton: 0, state: 39
episodes: 0, acton: 0, state: 139
episodes: 0, acton: 2, state: 239
episodes: 0, acton: 2, state: 259
episodes: 0, acton: 0, state: 279
episodes: 0, acton: 0, state: 379
episodes: 0, acton: 5, state: 479
episodes: 1, acton: 1, state: 486
episodes: 1, acton: 1, state: 386
episodes: 1, acton: 1, state: 286
episodes: 1, acton: 1, state: 186
episodes: 1, acton: 4, state: 86
episodes: 1, acton: 3, state: 98
episodes: 1, acton: 0, state: 78
episodes: 1, acton: 0, state: 178
episodes: 1, acton: 3, state: 278
episodes: 1, acton: 3, state: 258
episodes: 1, acton: 3, state: 238
episodes: 1, acton: 0, state: 218
episodes: 1, acton: 0, state: 318
episodes: 1, acton: 5, state: 418
episodes: 2, acton: 0, state: 68
episodes: 2, acton: 3, state: 168
```

```
episodes: 2, acton: 0, state: 148
episodes: 2, acton: 3, state: 248
episodes: 2, acton: 3, state: 228
episodes: 2, acton: 0, state: 208
episodes: 2, acton: 0, state: 308
episodes: 2, acton: 4, state: 408
episodes: 2, acton: 1, state: 416
episodes: 2, acton: 1, state: 316
episodes: 2, acton: 1, state: 216
episodes: 2, acton: 1, state: 116
episodes: 2, acton: 5, state: 16
episodes: 3, acton: 2, state: 167
episodes: 3, acton: 1, state: 187
episodes: 3, acton: 4, state: 87
episodes: 3, acton: 0, state: 99
episodes: 3, acton: 0, state: 199
episodes: 3, acton: 0, state: 299
episodes: 3, acton: 3, state: 399
episodes: 3, acton: 0, state: 379
episodes: 3, acton: 5, state: 479
episodes: 4, acton: 2, state: 214
episodes: 4, acton: 2, state: 234
episodes: 4, acton: 2, state: 254
episodes: 4, acton: 0, state: 274
episodes: 4, acton: 0, state: 374
episodes: 4, acton: 4, state: 474
episodes: 4, acton: 1, state: 478
episodes: 4, acton: 1, state: 378
episodes: 4, acton: 3, state: 278
episodes: 4, acton: 3, state: 258
episodes: 4, acton: 3, state: 238
episodes: 4, acton: 0, state: 218
episodes: 4, acton: 0, state: 318
episodes: 4, acton: 5, state: 418
Saving gif...Done!
```

In [50]:
```python
# display the captured GIF
display.Image(data=open("./gifs/taxi_q_controlled.gif",'rb').read(), form
# "taxi_q_controlled"
```

Out[50]:   <IPython.core.display.Image object>

If the above display function is not working, here is the previously prepared taxi_q_controlled.gif

Score=2.00, Episode=0, Step=2
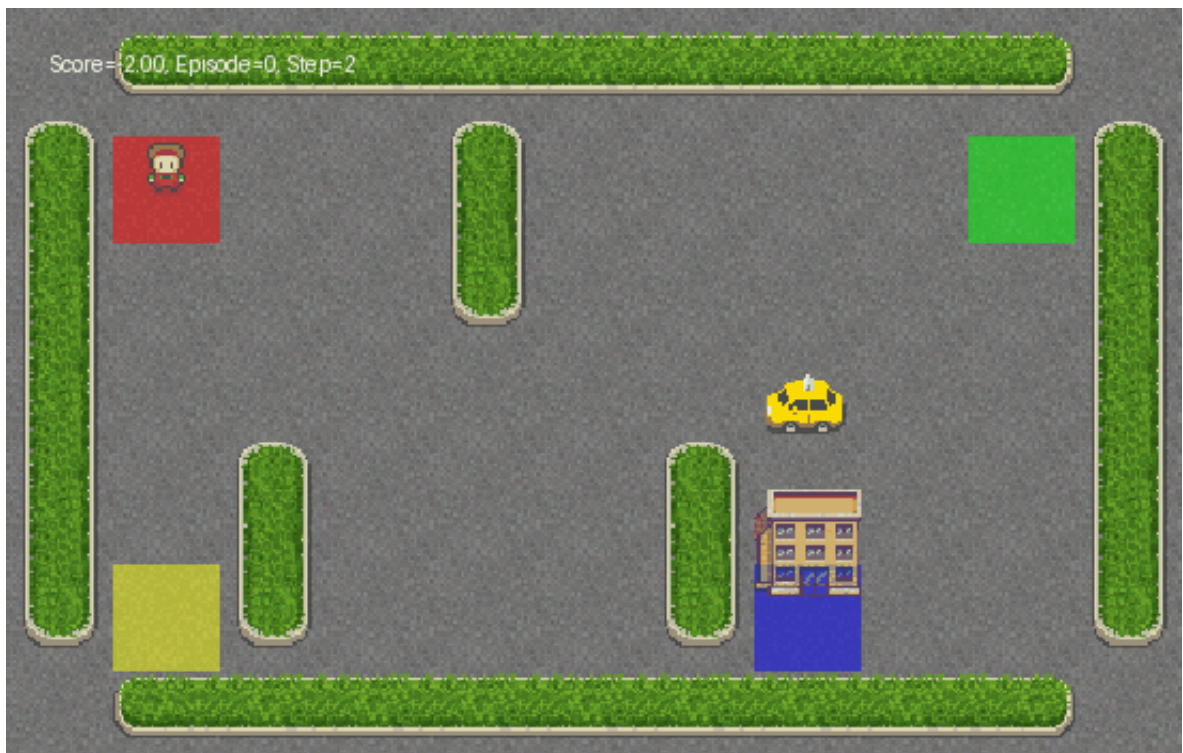
DQN Agent with Controlled Parameters

In [53]:
```
%%time
my_dqn_agent_controlled_avg_reward_lst = my_dqn_agent_controlled.test(tes
```

```
Trained Policy Network Model from the .pt file:
State 000, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 001, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 002, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 003, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 004, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 005, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 006, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 007, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 008, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 009, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 010, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 011, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 012, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 013, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 014, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 015, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 016, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 017, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 018, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 019, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 020, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 021, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 022, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 023, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 024, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 025, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 026, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 027, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 028, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 029, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 030, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 031, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 032, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 033, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 034, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 035, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 036, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 037, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 038, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 039, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 040, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 041, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 042, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 043, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 044, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 045, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 046, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 047, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 048, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 049, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 050, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 051, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 052, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 053, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 054, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 055, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 056, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 057, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 058, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 059, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 060, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 061, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 062, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 063, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 064, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 065, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 066, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 067, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 068, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 069, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 070, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 071, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 072, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 073, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 074, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 075, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 076, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 077, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 078, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 079, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 080, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 081, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 082, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 083, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 084, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 085, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 086, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 087, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 088, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 089, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 090, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 091, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 092, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 093, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 094, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 095, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 096, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 097, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 098, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 099, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 100, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 101, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 102, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 103, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 104, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 105, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 106, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 107, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 108, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 109, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 110, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 111, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 112, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 113, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 114, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 115, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 116, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 117, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 118, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 119, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 120, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 121, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 122, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 123, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 124, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 125, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 126, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 127, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 128, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 129, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 130, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 131, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 132, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 133, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 134, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 135, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 136, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 137, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 138, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 139, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 140, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 141, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 142, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 143, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 144, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 145, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 146, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 147, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 148, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 149, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 150, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 151, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 152, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 153, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 154, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 155, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 156, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 157, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 158, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 159, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 160, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 161, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 162, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 163, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 164, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 165, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 166, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 167, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 168, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 169, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 170, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 171, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 172, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 173, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 174, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 175, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 176, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 177, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 178, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 179, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 180, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 181, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 182, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 183, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 184, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 185, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 186, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 187, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 188, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 189, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 190, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 191, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 192, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 193, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 194, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 195, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 196, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 197, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 198, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 199, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 200, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 201, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 202, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 203, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 204, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 205, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 206, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 207, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 208, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 209, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 210, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 211, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 212, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 213, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 214, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 215, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 216, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 217, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 218, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 219, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 220, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 221, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 222, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 223, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 224, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 225, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 226, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 227, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 228, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 229, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 230, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 231, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 232, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 233, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 234, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 235, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 236, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 237, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 238, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 239, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 240, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 241, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 242, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 243, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 244, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 245, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 246, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 247, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 248, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 249, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 250, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 251, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 252, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 253, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 254, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 255, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 256, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 257, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 258, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 259, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 260, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 261, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 262, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 263, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 264, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 265, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 266, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 267, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 268, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 269, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 270, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 271, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 272, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 273, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 274, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 275, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 276, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 277, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 278, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 279, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 280, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 281, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 282, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 283, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 284, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 285, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 286, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 287, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 288, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 289, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 290, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 291, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 292, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 293, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 294, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 295, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 296, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 297, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 298, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 299, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 300, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 301, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 302, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 303, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 304, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 305, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 306, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 307, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 308, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 309, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 310, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 311, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 312, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 313, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 314, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 315, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 316, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 317, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 318, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 319, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 320, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 321, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 322, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 323, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 324, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 325, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 326, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 327, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 328, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 329, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 330, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 331, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 332, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 333, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 334, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 335, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 336, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 337, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 338, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 339, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 340, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 341, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 342, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 343, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 344, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 345, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 346, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 347, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 348, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 349, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 350, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 351, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 352, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 353, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 354, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 355, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 356, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 357, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 358, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 359, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 360, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 361, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 362, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 363, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 364, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 365, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 366, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 367, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 368, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 369, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 370, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 371, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 372, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 373, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 374, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 375, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 376, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 377, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 378, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 379, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 380, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 381, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 382, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 383, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 384, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 385, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 386, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 387, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 388, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 389, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 390, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 391, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 392, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 393, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 394, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 395, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 396, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 397, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 398, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 399, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 400, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 401, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 402, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 403, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 404, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 405, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 406, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 407, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 408, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 409, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 410, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 411, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 412, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 413, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 414, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 415, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 416, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 417, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 418, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 419, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 420, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 421, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 422, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 423, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 424, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 425, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 426, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 427, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 428, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 429, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 430, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 431, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 432, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 433, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 434, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 435, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 436, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 437, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 438, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 439, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 440, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 441, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 442, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 443, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 444, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 445, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 446, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 447, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 448, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 449, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 450, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 451, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 452, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 453, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 454, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 455, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 456, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 457, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 458, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 459, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 460, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 461, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 462, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 463, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 464, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 465, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 466, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 467, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 468, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 469, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 470, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 471, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 472, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 473, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 474, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 475, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 476, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 477, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 478, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 479, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 480, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 481, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 482, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 483, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 484, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 485, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 486, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 487, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 488, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 489, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 490, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 491, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 492, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 493, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 494, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 495, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 496, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 497, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 498, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 499, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
Episode 1: Total Reward = -200
Episode 2: Total Reward = -200
Episode 3: Total Reward = -200
Episode 4: Total Reward = -200
Episode 5: Total Reward = -200
Episode 6: Total Reward = -200
Episode 7: Total Reward = -200
Episode 8: Total Reward = -200
Episode 9: Total Reward = -200
Episode 10: Total Reward = -200
Episode 11: Total Reward = -200
Episode 12: Total Reward = -200
Episode 13: Total Reward = -200
Episode 14: Total Reward = -200
Episode 15: Total Reward = -200
Episode 16: Total Reward = -200
Episode 17: Total Reward = -200
Episode 18: Total Reward = -200
Episode 19: Total Reward = -200
Episode 20: Total Reward = -200
Episode 21: Total Reward = -200
Episode 22: Total Reward = -200
Episode 23: Total Reward = -200
Episode 24: Total Reward = -200
Episode 25: Total Reward = -200
Episode 26: Total Reward = -200
Episode 27: Total Reward = -200
Episode 28: Total Reward = -200
Episode 29: Total Reward = -200
Episode 30: Total Reward = -200
Episode 31: Total Reward = -200
Episode 32: Total Reward = -200
Episode 33: Total Reward = -200
Episode 34: Total Reward = -200
Episode 35: Total Reward = -200
Episode 36: Total Reward = -200
Episode 37: Total Reward = -200
Episode 38: Total Reward = -200
Episode 39: Total Reward = -200
Episode 40: Total Reward = -200
Episode 41: Total Reward = -200
Episode 42: Total Reward = -200
```

```
Episode 43: Total Reward = -200
Episode 44: Total Reward = -200
Episode 45: Total Reward = -200
Episode 46: Total Reward = -200
Episode 47: Total Reward = -200
Episode 48: Total Reward = -200
Episode 49: Total Reward = -200
Episode 50: Total Reward = -200
Episode 51: Total Reward = -200
Episode 52: Total Reward = -200
Episode 53: Total Reward = -200
Episode 54: Total Reward = -200
Episode 55: Total Reward = -200
Episode 56: Total Reward = -200
Episode 57: Total Reward = -200
Episode 58: Total Reward = -200
Episode 59: Total Reward = -200
Episode 60: Total Reward = -200
Episode 61: Total Reward = -200
Episode 62: Total Reward = -200
Episode 63: Total Reward = -200
Episode 64: Total Reward = -200
Episode 65: Total Reward = -200
Episode 66: Total Reward = -200
Episode 67: Total Reward = -200
Episode 68: Total Reward = -200
Episode 69: Total Reward = -200
Episode 70: Total Reward = -200
Episode 71: Total Reward = -200
Episode 72: Total Reward = -200
Episode 73: Total Reward = -200
Episode 74: Total Reward = -200
Episode 75: Total Reward = -200
Episode 76: Total Reward = -200
Episode 77: Total Reward = -200
Episode 78: Total Reward = -200
Episode 79: Total Reward = -200
Episode 80: Total Reward = -200
Episode 81: Total Reward = -200
Episode 82: Total Reward = -200
Episode 83: Total Reward = -200
Episode 84: Total Reward = -200
Episode 85: Total Reward = -200
Episode 86: Total Reward = -200
Episode 87: Total Reward = -200
Episode 88: Total Reward = -200
Episode 89: Total Reward = -200
Episode 90: Total Reward = -200
Episode 91: Total Reward = -200
Episode 92: Total Reward = -200
Episode 93: Total Reward = -200
Episode 94: Total Reward = -200
Episode 95: Total Reward = -200
Episode 96: Total Reward = -200
Episode 97: Total Reward = -200
Episode 98: Total Reward = -200
Episode 99: Total Reward = -200
```

```
Episode 100: Total Reward = -200
Episode 101: Total Reward = -200
Episode 102: Total Reward = -200
Episode 103: Total Reward = -200
Episode 104: Total Reward = -200
Episode 105: Total Reward = -200
Episode 106: Total Reward = -200
Episode 107: Total Reward = -200
Episode 108: Total Reward = -200
Episode 109: Total Reward = -200
Episode 110: Total Reward = -200
Episode 111: Total Reward = -200
Episode 112: Total Reward = -200
Episode 113: Total Reward = -200
Episode 114: Total Reward = -200
Episode 115: Total Reward = -200
Episode 116: Total Reward = -200
Episode 117: Total Reward = -200
Episode 118: Total Reward = -200
Episode 119: Total Reward = -200
Episode 120: Total Reward = -200
Episode 121: Total Reward = -200
Episode 122: Total Reward = -200
Episode 123: Total Reward = -200
Episode 124: Total Reward = -200
Episode 125: Total Reward = -200
Episode 126: Total Reward = -200
Episode 127: Total Reward = -200
Episode 128: Total Reward = -200
Episode 129: Total Reward = -200
Episode 130: Total Reward = -200
Episode 131: Total Reward = -200
Episode 132: Total Reward = -200
Episode 133: Total Reward = -200
Episode 134: Total Reward = -200
Episode 135: Total Reward = -200
Episode 136: Total Reward = -200
Episode 137: Total Reward = -200
Episode 138: Total Reward = -200
Episode 139: Total Reward = -200
Episode 140: Total Reward = -200
Episode 141: Total Reward = -200
Episode 142: Total Reward = -200
Episode 143: Total Reward = -200
Episode 144: Total Reward = -200
Episode 145: Total Reward = -200
Episode 146: Total Reward = -200
Episode 147: Total Reward = -200
Episode 148: Total Reward = -200
Episode 149: Total Reward = -200
Episode 150: Total Reward = -200
Episode 151: Total Reward = -200
Episode 152: Total Reward = -200
Episode 153: Total Reward = -200
Episode 154: Total Reward = -200
Episode 155: Total Reward = -200
Episode 156: Total Reward = -200
```

```
Episode 157: Total Reward = -200
Episode 158: Total Reward = -200
Episode 159: Total Reward = -200
Episode 160: Total Reward = -200
Episode 161: Total Reward = -200
Episode 162: Total Reward = -200
Episode 163: Total Reward = -200
Episode 164: Total Reward = -200
Episode 165: Total Reward = -200
Episode 166: Total Reward = -200
Episode 167: Total Reward = -200
Episode 168: Total Reward = -200
Episode 169: Total Reward = -200
Episode 170: Total Reward = -200
Episode 171: Total Reward = -200
Episode 172: Total Reward = -200
Episode 173: Total Reward = -200
Episode 174: Total Reward = -200
Episode 175: Total Reward = -200
Episode 176: Total Reward = -200
Episode 177: Total Reward = -200
Episode 178: Total Reward = -200
Episode 179: Total Reward = -200
Episode 180: Total Reward = -200
Episode 181: Total Reward = -200
Episode 182: Total Reward = -200
Episode 183: Total Reward = -200
Episode 184: Total Reward = -200
Episode 185: Total Reward = -200
Episode 186: Total Reward = -200
Episode 187: Total Reward = -200
Episode 188: Total Reward = -200
Episode 189: Total Reward = -200
Episode 190: Total Reward = -200
Episode 191: Total Reward = -200
Episode 192: Total Reward = -200
Episode 193: Total Reward = -200
Episode 194: Total Reward = -200
Episode 195: Total Reward = -200
Episode 196: Total Reward = -200
Episode 197: Total Reward = -200
Episode 198: Total Reward = -200
Episode 199: Total Reward = -200
Episode 200: Total Reward = -200
Episode 201: Total Reward = -200
Episode 202: Total Reward = -200
Episode 203: Total Reward = -200
Episode 204: Total Reward = -200
Episode 205: Total Reward = -200
Episode 206: Total Reward = -200
Episode 207: Total Reward = -200
Episode 208: Total Reward = -200
Episode 209: Total Reward = -200
Episode 210: Total Reward = -200
Episode 211: Total Reward = -200
Episode 212: Total Reward = -200
Episode 213: Total Reward = -200
```

```
Episode 214: Total Reward = -200
Episode 215: Total Reward = -200
Episode 216: Total Reward = -200
Episode 217: Total Reward = -200
Episode 218: Total Reward = -200
Episode 219: Total Reward = -200
Episode 220: Total Reward = -200
Episode 221: Total Reward = -200
Episode 222: Total Reward = -200
Episode 223: Total Reward = -200
Episode 224: Total Reward = -200
Episode 225: Total Reward = -200
Episode 226: Total Reward = -200
Episode 227: Total Reward = -200
Episode 228: Total Reward = -200
Episode 229: Total Reward = -200
Episode 230: Total Reward = -200
Episode 231: Total Reward = -200
Episode 232: Total Reward = -200
Episode 233: Total Reward = -200
Episode 234: Total Reward = -200
Episode 235: Total Reward = -200
Episode 236: Total Reward = -200
Episode 237: Total Reward = -200
Episode 238: Total Reward = -200
Episode 239: Total Reward = -200
Episode 240: Total Reward = -200
Episode 241: Total Reward = -200
Episode 242: Total Reward = -200
Episode 243: Total Reward = -200
Episode 244: Total Reward = -200
Episode 245: Total Reward = -200
Episode 246: Total Reward = -200
Episode 247: Total Reward = -200
Episode 248: Total Reward = -200
Episode 249: Total Reward = -200
Episode 250: Total Reward = -200
Episode 251: Total Reward = -200
Episode 252: Total Reward = -200
Episode 253: Total Reward = -200
Episode 254: Total Reward = -200
Episode 255: Total Reward = -200
Episode 256: Total Reward = -200
Episode 257: Total Reward = -200
Episode 258: Total Reward = -200
Episode 259: Total Reward = -200
Episode 260: Total Reward = -200
Episode 261: Total Reward = -200
Episode 262: Total Reward = -200
Episode 263: Total Reward = -200
Episode 264: Total Reward = -200
Episode 265: Total Reward = -200
Episode 266: Total Reward = -200
Episode 267: Total Reward = -200
Episode 268: Total Reward = -200
Episode 269: Total Reward = -200
Episode 270: Total Reward = -200
```

```
Episode 271: Total Reward = -200
Episode 272: Total Reward = -200
Episode 273: Total Reward = -200
Episode 274: Total Reward = -200
Episode 275: Total Reward = -200
Episode 276: Total Reward = -200
Episode 277: Total Reward = -200
Episode 278: Total Reward = -200
Episode 279: Total Reward = -200
Episode 280: Total Reward = -200
Episode 281: Total Reward = -200
Episode 282: Total Reward = -200
Episode 283: Total Reward = -200
Episode 284: Total Reward = -200
Episode 285: Total Reward = -200
Episode 286: Total Reward = -200
Episode 287: Total Reward = -200
Episode 288: Total Reward = -200
Episode 289: Total Reward = -200
Episode 290: Total Reward = -200
Episode 291: Total Reward = -200
Episode 292: Total Reward = -200
Episode 293: Total Reward = -200
Episode 294: Total Reward = -200
Episode 295: Total Reward = -200
Episode 296: Total Reward = -200
Episode 297: Total Reward = -200
Episode 298: Total Reward = -200
Episode 299: Total Reward = -200
Episode 300: Total Reward = -200
Episode 301: Total Reward = -200
Episode 302: Total Reward = -200
Episode 303: Total Reward = -200
Episode 304: Total Reward = -200
Episode 305: Total Reward = -200
Episode 306: Total Reward = -200
Episode 307: Total Reward = -200
Episode 308: Total Reward = -200
Episode 309: Total Reward = -200
Episode 310: Total Reward = -200
Episode 311: Total Reward = -200
Episode 312: Total Reward = -200
Episode 313: Total Reward = -200
Episode 314: Total Reward = -200
Episode 315: Total Reward = -200
Episode 316: Total Reward = -200
Episode 317: Total Reward = -200
Episode 318: Total Reward = -200
Episode 319: Total Reward = -200
Episode 320: Total Reward = -200
Episode 321: Total Reward = -200
Episode 322: Total Reward = -200
Episode 323: Total Reward = -200
Episode 324: Total Reward = -200
Episode 325: Total Reward = -200
Episode 326: Total Reward = -200
Episode 327: Total Reward = -200
```

```
Episode 328: Total Reward = -200
Episode 329: Total Reward = -200
Episode 330: Total Reward = -200
Episode 331: Total Reward = -200
Episode 332: Total Reward = -200
Episode 333: Total Reward = -200
Episode 334: Total Reward = -200
Episode 335: Total Reward = -200
Episode 336: Total Reward = -200
Episode 337: Total Reward = -200
Episode 338: Total Reward = -200
Episode 339: Total Reward = -200
Episode 340: Total Reward = -200
Episode 341: Total Reward = -200
Episode 342: Total Reward = -200
Episode 343: Total Reward = -200
Episode 344: Total Reward = -200
Episode 345: Total Reward = -200
Episode 346: Total Reward = -200
Episode 347: Total Reward = -200
Episode 348: Total Reward = -200
Episode 349: Total Reward = -200
Episode 350: Total Reward = -200
Episode 351: Total Reward = -200
Episode 352: Total Reward = -200
Episode 353: Total Reward = -200
Episode 354: Total Reward = -200
Episode 355: Total Reward = -200
Episode 356: Total Reward = -200
Episode 357: Total Reward = -200
Episode 358: Total Reward = -200
Episode 359: Total Reward = -200
Episode 360: Total Reward = -200
Episode 361: Total Reward = -200
Episode 362: Total Reward = -200
Episode 363: Total Reward = -200
Episode 364: Total Reward = -200
Episode 365: Total Reward = -200
Episode 366: Total Reward = -200
Episode 367: Total Reward = -200
Episode 368: Total Reward = -200
Episode 369: Total Reward = -200
Episode 370: Total Reward = -200
Episode 371: Total Reward = -200
Episode 372: Total Reward = -200
Episode 373: Total Reward = -200
Episode 374: Total Reward = -200
Episode 375: Total Reward = -200
Episode 376: Total Reward = -200
Episode 377: Total Reward = -200
Episode 378: Total Reward = -200
Episode 379: Total Reward = -200
Episode 380: Total Reward = -200
Episode 381: Total Reward = -200
Episode 382: Total Reward = -200
Episode 383: Total Reward = -200
Episode 384: Total Reward = -200
```

```
Episode 385: Total Reward = -200
Episode 386: Total Reward = -200
Episode 387: Total Reward = -200
Episode 388: Total Reward = -200
Episode 389: Total Reward = -200
Episode 390: Total Reward = -200
Episode 391: Total Reward = -200
Episode 392: Total Reward = -200
Episode 393: Total Reward = -200
Episode 394: Total Reward = -200
Episode 395: Total Reward = -200
Episode 396: Total Reward = -200
Episode 397: Total Reward = -200
Episode 398: Total Reward = -200
Episode 399: Total Reward = -200
Episode 400: Total Reward = -200
Episode 401: Total Reward = -200
Episode 402: Total Reward = -200
Episode 403: Total Reward = -200
Episode 404: Total Reward = -200
Episode 405: Total Reward = -200
Episode 406: Total Reward = -200
Episode 407: Total Reward = -200
Episode 408: Total Reward = -200
Episode 409: Total Reward = -200
Episode 410: Total Reward = -200
Episode 411: Total Reward = -200
Episode 412: Total Reward = -200
Episode 413: Total Reward = -200
Episode 414: Total Reward = -200
Episode 415: Total Reward = -200
Episode 416: Total Reward = -200
Episode 417: Total Reward = -200
Episode 418: Total Reward = -200
Episode 419: Total Reward = -200
Episode 420: Total Reward = -200
Episode 421: Total Reward = -200
Episode 422: Total Reward = -200
Episode 423: Total Reward = -200
Episode 424: Total Reward = -200
Episode 425: Total Reward = -200
Episode 426: Total Reward = -200
Episode 427: Total Reward = -200
Episode 428: Total Reward = -200
Episode 429: Total Reward = -200
Episode 430: Total Reward = -200
Episode 431: Total Reward = -200
Episode 432: Total Reward = -200
Episode 433: Total Reward = -200
Episode 434: Total Reward = -200
Episode 435: Total Reward = -200
Episode 436: Total Reward = -200
Episode 437: Total Reward = -200
Episode 438: Total Reward = -200
Episode 439: Total Reward = -200
Episode 440: Total Reward = -200
Episode 441: Total Reward = -200
```

```
Episode 442: Total Reward = -200
Episode 443: Total Reward = -200
Episode 444: Total Reward = -200
Episode 445: Total Reward = -200
Episode 446: Total Reward = -200
Episode 447: Total Reward = -200
Episode 448: Total Reward = -200
Episode 449: Total Reward = -200
Episode 450: Total Reward = -200
Episode 451: Total Reward = -200
Episode 452: Total Reward = -200
Episode 453: Total Reward = -200
Episode 454: Total Reward = -200
Episode 455: Total Reward = -200
Episode 456: Total Reward = -200
Episode 457: Total Reward = -200
Episode 458: Total Reward = -200
Episode 459: Total Reward = -200
Episode 460: Total Reward = -200
Episode 461: Total Reward = -200
Episode 462: Total Reward = -200
Episode 463: Total Reward = -200
Episode 464: Total Reward = -200
Episode 465: Total Reward = -200
Episode 466: Total Reward = -200
Episode 467: Total Reward = -200
Episode 468: Total Reward = -200
Episode 469: Total Reward = -200
Episode 470: Total Reward = -200
Episode 471: Total Reward = -200
Episode 472: Total Reward = -200
Episode 473: Total Reward = -200
Episode 474: Total Reward = -200
Episode 475: Total Reward = -200
Episode 476: Total Reward = -200
Episode 477: Total Reward = -200
Episode 478: Total Reward = -200
Episode 479: Total Reward = -200
Episode 480: Total Reward = -200
Episode 481: Total Reward = -200
Episode 482: Total Reward = -200
Episode 483: Total Reward = -200
Episode 484: Total Reward = -200
Episode 485: Total Reward = -200
Episode 486: Total Reward = -200
Episode 487: Total Reward = -200
Episode 488: Total Reward = -200
Episode 489: Total Reward = -200
Episode 490: Total Reward = -200
Episode 491: Total Reward = -200
Episode 492: Total Reward = -200
Episode 493: Total Reward = -200
Episode 494: Total Reward = -200
Episode 495: Total Reward = -200
Episode 496: Total Reward = -200
Episode 497: Total Reward = -200
Episode 498: Total Reward = -200
```

```
Episode 499: Total Reward = −200
Episode 500: Total Reward = −200
Episode 501: Total Reward = −200
Episode 502: Total Reward = −200
Episode 503: Total Reward = −200
Episode 504: Total Reward = −200
Episode 505: Total Reward = −200
Episode 506: Total Reward = −200
Episode 507: Total Reward = −200
Episode 508: Total Reward = −200
Episode 509: Total Reward = −200
Episode 510: Total Reward = −200
Episode 511: Total Reward = −200
Episode 512: Total Reward = −200
Episode 513: Total Reward = −200
Episode 514: Total Reward = −200
Episode 515: Total Reward = −200
Episode 516: Total Reward = −200
Episode 517: Total Reward = −200
Episode 518: Total Reward = −200
Episode 519: Total Reward = −200
Episode 520: Total Reward = −200
Episode 521: Total Reward = −200
Episode 522: Total Reward = −200
Episode 523: Total Reward = −200
Episode 524: Total Reward = −200
Episode 525: Total Reward = −200
Episode 526: Total Reward = −200
Episode 527: Total Reward = −200
Episode 528: Total Reward = −200
Episode 529: Total Reward = −200
Episode 530: Total Reward = −200
Episode 531: Total Reward = −200
Episode 532: Total Reward = −200
Episode 533: Total Reward = −200
Episode 534: Total Reward = −200
Episode 535: Total Reward = −200
Episode 536: Total Reward = −200
Episode 537: Total Reward = −200
Episode 538: Total Reward = −200
Episode 539: Total Reward = −200
Episode 540: Total Reward = −200
Episode 541: Total Reward = −200
Episode 542: Total Reward = −200
Episode 543: Total Reward = −200
Episode 544: Total Reward = −200
Episode 545: Total Reward = −200
Episode 546: Total Reward = −200
Episode 547: Total Reward = −200
Episode 548: Total Reward = −200
Episode 549: Total Reward = −200
Episode 550: Total Reward = −200
Episode 551: Total Reward = −200
Episode 552: Total Reward = −200
Episode 553: Total Reward = −200
Episode 554: Total Reward = −200
Episode 555: Total Reward = −200
```

```
Episode 556: Total Reward = -200
Episode 557: Total Reward = -200
Episode 558: Total Reward = -200
Episode 559: Total Reward = -200
Episode 560: Total Reward = -200
Episode 561: Total Reward = -200
Episode 562: Total Reward = -200
Episode 563: Total Reward = -200
Episode 564: Total Reward = -200
Episode 565: Total Reward = -200
Episode 566: Total Reward = -200
Episode 567: Total Reward = -200
Episode 568: Total Reward = -200
Episode 569: Total Reward = -200
Episode 570: Total Reward = -200
Episode 571: Total Reward = -200
Episode 572: Total Reward = -200
Episode 573: Total Reward = -200
Episode 574: Total Reward = -200
Episode 575: Total Reward = -200
Episode 576: Total Reward = -200
Episode 577: Total Reward = -200
Episode 578: Total Reward = -200
Episode 579: Total Reward = -200
Episode 580: Total Reward = -200
Episode 581: Total Reward = -200
Episode 582: Total Reward = -200
Episode 583: Total Reward = -200
Episode 584: Total Reward = -200
Episode 585: Total Reward = -200
Episode 586: Total Reward = -200
Episode 587: Total Reward = -200
Episode 588: Total Reward = -200
Episode 589: Total Reward = -200
Episode 590: Total Reward = -200
Episode 591: Total Reward = -200
Episode 592: Total Reward = -200
Episode 593: Total Reward = -200
Episode 594: Total Reward = -200
Episode 595: Total Reward = -200
Episode 596: Total Reward = -200
Episode 597: Total Reward = -200
Episode 598: Total Reward = -200
Episode 599: Total Reward = -200
Episode 600: Total Reward = -200
Episode 601: Total Reward = -200
Episode 602: Total Reward = -200
Episode 603: Total Reward = -200
Episode 604: Total Reward = -200
Episode 605: Total Reward = -200
Episode 606: Total Reward = -200
Episode 607: Total Reward = -200
Episode 608: Total Reward = -200
Episode 609: Total Reward = -200
Episode 610: Total Reward = -200
Episode 611: Total Reward = -200
Episode 612: Total Reward = -200
```

```
Episode 613: Total Reward = -200
Episode 614: Total Reward = -200
Episode 615: Total Reward = -200
Episode 616: Total Reward = -200
Episode 617: Total Reward = -200
Episode 618: Total Reward = -200
Episode 619: Total Reward = -200
Episode 620: Total Reward = -200
Episode 621: Total Reward = -200
Episode 622: Total Reward = -200
Episode 623: Total Reward = -200
Episode 624: Total Reward = -200
Episode 625: Total Reward = -200
Episode 626: Total Reward = -200
Episode 627: Total Reward = -200
Episode 628: Total Reward = -200
Episode 629: Total Reward = -200
Episode 630: Total Reward = -200
Episode 631: Total Reward = -200
Episode 632: Total Reward = -200
Episode 633: Total Reward = -200
Episode 634: Total Reward = -200
Episode 635: Total Reward = -200
Episode 636: Total Reward = -200
Episode 637: Total Reward = -200
Episode 638: Total Reward = -200
Episode 639: Total Reward = -200
Episode 640: Total Reward = -200
Episode 641: Total Reward = -200
Episode 642: Total Reward = -200
Episode 643: Total Reward = -200
Episode 644: Total Reward = -200
Episode 645: Total Reward = -200
Episode 646: Total Reward = -200
Episode 647: Total Reward = -200
Episode 648: Total Reward = -200
Episode 649: Total Reward = -200
Episode 650: Total Reward = -200
Episode 651: Total Reward = -200
Episode 652: Total Reward = -200
Episode 653: Total Reward = -200
Episode 654: Total Reward = -200
Episode 655: Total Reward = -200
Episode 656: Total Reward = -200
Episode 657: Total Reward = -200
Episode 658: Total Reward = -200
Episode 659: Total Reward = -200
Episode 660: Total Reward = -200
Episode 661: Total Reward = -200
Episode 662: Total Reward = -200
Episode 663: Total Reward = -200
Episode 664: Total Reward = -200
Episode 665: Total Reward = -200
Episode 666: Total Reward = -200
Episode 667: Total Reward = -200
Episode 668: Total Reward = -200
Episode 669: Total Reward = -200
```

```
Episode 670: Total Reward = -200
Episode 671: Total Reward = -200
Episode 672: Total Reward = -200
Episode 673: Total Reward = -200
Episode 674: Total Reward = -200
Episode 675: Total Reward = -200
Episode 676: Total Reward = -200
Episode 677: Total Reward = -200
Episode 678: Total Reward = -200
Episode 679: Total Reward = -200
Episode 680: Total Reward = -200
Episode 681: Total Reward = -200
Episode 682: Total Reward = -200
Episode 683: Total Reward = -200
Episode 684: Total Reward = -200
Episode 685: Total Reward = -200
Episode 686: Total Reward = -200
Episode 687: Total Reward = -200
Episode 688: Total Reward = -200
Episode 689: Total Reward = -200
Episode 690: Total Reward = -200
Episode 691: Total Reward = -200
Episode 692: Total Reward = -200
Episode 693: Total Reward = -200
Episode 694: Total Reward = -200
Episode 695: Total Reward = -200
Episode 696: Total Reward = -200
Episode 697: Total Reward = -200
Episode 698: Total Reward = -200
Episode 699: Total Reward = -200
Episode 700: Total Reward = -200
Episode 701: Total Reward = -200
Episode 702: Total Reward = -200
Episode 703: Total Reward = -200
Episode 704: Total Reward = -200
Episode 705: Total Reward = -200
Episode 706: Total Reward = -200
Episode 707: Total Reward = -200
Episode 708: Total Reward = -200
Episode 709: Total Reward = -200
Episode 710: Total Reward = -200
Episode 711: Total Reward = -200
Episode 712: Total Reward = -200
Episode 713: Total Reward = -200
Episode 714: Total Reward = -200
Episode 715: Total Reward = -200
Episode 716: Total Reward = -200
Episode 717: Total Reward = -200
Episode 718: Total Reward = -200
Episode 719: Total Reward = -200
Episode 720: Total Reward = -200
Episode 721: Total Reward = -200
Episode 722: Total Reward = -200
Episode 723: Total Reward = -200
Episode 724: Total Reward = -200
Episode 725: Total Reward = -200
Episode 726: Total Reward = -200
```

```
Episode 727: Total Reward = -200
Episode 728: Total Reward = -200
Episode 729: Total Reward = -200
Episode 730: Total Reward = -200
Episode 731: Total Reward = -200
Episode 732: Total Reward = -200
Episode 733: Total Reward = -200
Episode 734: Total Reward = -200
Episode 735: Total Reward = -200
Episode 736: Total Reward = -200
Episode 737: Total Reward = -200
Episode 738: Total Reward = -200
Episode 739: Total Reward = -200
Episode 740: Total Reward = -200
Episode 741: Total Reward = -200
Episode 742: Total Reward = -200
Episode 743: Total Reward = -200
Episode 744: Total Reward = -200
Episode 745: Total Reward = -200
Episode 746: Total Reward = -200
Episode 747: Total Reward = -200
Episode 748: Total Reward = -200
Episode 749: Total Reward = -200
Episode 750: Total Reward = -200
Episode 751: Total Reward = -200
Episode 752: Total Reward = -200
Episode 753: Total Reward = -200
Episode 754: Total Reward = -200
Episode 755: Total Reward = -200
Episode 756: Total Reward = -200
Episode 757: Total Reward = -200
Episode 758: Total Reward = -200
Episode 759: Total Reward = -200
Episode 760: Total Reward = -200
Episode 761: Total Reward = -200
Episode 762: Total Reward = -200
Episode 763: Total Reward = -200
Episode 764: Total Reward = -200
Episode 765: Total Reward = -200
Episode 766: Total Reward = -200
Episode 767: Total Reward = -200
Episode 768: Total Reward = -200
Episode 769: Total Reward = -200
Episode 770: Total Reward = -200
Episode 771: Total Reward = -200
Episode 772: Total Reward = -200
Episode 773: Total Reward = -200
Episode 774: Total Reward = -200
Episode 775: Total Reward = -200
Episode 776: Total Reward = -200
Episode 777: Total Reward = -200
Episode 778: Total Reward = -200
Episode 779: Total Reward = -200
Episode 780: Total Reward = -200
Episode 781: Total Reward = -200
Episode 782: Total Reward = -200
Episode 783: Total Reward = -200
```

```
Episode 784: Total Reward = -200
Episode 785: Total Reward = -200
Episode 786: Total Reward = -200
Episode 787: Total Reward = -200
Episode 788: Total Reward = -200
Episode 789: Total Reward = -200
Episode 790: Total Reward = -200
Episode 791: Total Reward = -200
Episode 792: Total Reward = -200
Episode 793: Total Reward = -200
Episode 794: Total Reward = -200
Episode 795: Total Reward = -200
Episode 796: Total Reward = -200
Episode 797: Total Reward = -200
Episode 798: Total Reward = -200
Episode 799: Total Reward = -200
Episode 800: Total Reward = -200
Episode 801: Total Reward = -200
Episode 802: Total Reward = -200
Episode 803: Total Reward = -200
Episode 804: Total Reward = -200
Episode 805: Total Reward = -200
Episode 806: Total Reward = -200
Episode 807: Total Reward = -200
Episode 808: Total Reward = -200
Episode 809: Total Reward = -200
Episode 810: Total Reward = -200
Episode 811: Total Reward = -200
Episode 812: Total Reward = -200
Episode 813: Total Reward = -200
Episode 814: Total Reward = -200
Episode 815: Total Reward = -200
Episode 816: Total Reward = -200
Episode 817: Total Reward = -200
Episode 818: Total Reward = -200
Episode 819: Total Reward = -200
Episode 820: Total Reward = -200
Episode 821: Total Reward = -200
Episode 822: Total Reward = -200
Episode 823: Total Reward = -200
Episode 824: Total Reward = -200
Episode 825: Total Reward = -200
Episode 826: Total Reward = -200
Episode 827: Total Reward = -200
Episode 828: Total Reward = -200
Episode 829: Total Reward = -200
Episode 830: Total Reward = -200
Episode 831: Total Reward = -200
Episode 832: Total Reward = -200
Episode 833: Total Reward = -200
Episode 834: Total Reward = -200
Episode 835: Total Reward = -200
Episode 836: Total Reward = -200
Episode 837: Total Reward = -200
Episode 838: Total Reward = -200
Episode 839: Total Reward = -200
Episode 840: Total Reward = -200
```

```
Episode 841: Total Reward = -200
Episode 842: Total Reward = -200
Episode 843: Total Reward = -200
Episode 844: Total Reward = -200
Episode 845: Total Reward = -200
Episode 846: Total Reward = -200
Episode 847: Total Reward = -200
Episode 848: Total Reward = -200
Episode 849: Total Reward = -200
Episode 850: Total Reward = -200
Episode 851: Total Reward = -200
Episode 852: Total Reward = -200
Episode 853: Total Reward = -200
Episode 854: Total Reward = -200
Episode 855: Total Reward = -200
Episode 856: Total Reward = -200
Episode 857: Total Reward = -200
Episode 858: Total Reward = -200
Episode 859: Total Reward = -200
Episode 860: Total Reward = -200
Episode 861: Total Reward = -200
Episode 862: Total Reward = -200
Episode 863: Total Reward = -200
Episode 864: Total Reward = -200
Episode 865: Total Reward = -200
Episode 866: Total Reward = -200
Episode 867: Total Reward = -200
Episode 868: Total Reward = -200
Episode 869: Total Reward = -200
Episode 870: Total Reward = -200
Episode 871: Total Reward = -200
Episode 872: Total Reward = -200
Episode 873: Total Reward = -200
Episode 874: Total Reward = -200
Episode 875: Total Reward = -200
Episode 876: Total Reward = -200
Episode 877: Total Reward = -200
Episode 878: Total Reward = -200
Episode 879: Total Reward = -200
Episode 880: Total Reward = -200
Episode 881: Total Reward = -200
Episode 882: Total Reward = -200
Episode 883: Total Reward = -200
Episode 884: Total Reward = -200
Episode 885: Total Reward = -200
Episode 886: Total Reward = -200
Episode 887: Total Reward = -200
Episode 888: Total Reward = -200
Episode 889: Total Reward = -200
Episode 890: Total Reward = -200
Episode 891: Total Reward = -200
Episode 892: Total Reward = -200
Episode 893: Total Reward = -200
Episode 894: Total Reward = -200
Episode 895: Total Reward = -200
Episode 896: Total Reward = -200
Episode 897: Total Reward = -200
```

```
Episode 898: Total Reward = -200
Episode 899: Total Reward = -200
Episode 900: Total Reward = -200
Episode 901: Total Reward = -200
Episode 902: Total Reward = -200
Episode 903: Total Reward = -200
Episode 904: Total Reward = -200
Episode 905: Total Reward = -200
Episode 906: Total Reward = -200
Episode 907: Total Reward = -200
Episode 908: Total Reward = -200
Episode 909: Total Reward = -200
Episode 910: Total Reward = -200
Episode 911: Total Reward = -200
Episode 912: Total Reward = -200
Episode 913: Total Reward = -200
Episode 914: Total Reward = -200
Episode 915: Total Reward = -200
Episode 916: Total Reward = -200
Episode 917: Total Reward = -200
Episode 918: Total Reward = -200
Episode 919: Total Reward = -200
Episode 920: Total Reward = -200
Episode 921: Total Reward = -200
Episode 922: Total Reward = -200
Episode 923: Total Reward = -200
Episode 924: Total Reward = -200
Episode 925: Total Reward = -200
Episode 926: Total Reward = -200
Episode 927: Total Reward = -200
Episode 928: Total Reward = -200
Episode 929: Total Reward = -200
Episode 930: Total Reward = -200
Episode 931: Total Reward = -200
Episode 932: Total Reward = -200
Episode 933: Total Reward = -200
Episode 934: Total Reward = -200
Episode 935: Total Reward = -200
Episode 936: Total Reward = -200
Episode 937: Total Reward = -200
Episode 938: Total Reward = -200
Episode 939: Total Reward = -200
Episode 940: Total Reward = -200
Episode 941: Total Reward = -200
Episode 942: Total Reward = -200
Episode 943: Total Reward = -200
Episode 944: Total Reward = -200
Episode 945: Total Reward = -200
Episode 946: Total Reward = -200
Episode 947: Total Reward = -200
Episode 948: Total Reward = -200
Episode 949: Total Reward = -200
Episode 950: Total Reward = -200
Episode 951: Total Reward = -200
Episode 952: Total Reward = -200
Episode 953: Total Reward = -200
Episode 954: Total Reward = -200
```

```
Episode 955: Total Reward = -200
Episode 956: Total Reward = -200
Episode 957: Total Reward = -200
Episode 958: Total Reward = -200
Episode 959: Total Reward = -200
Episode 960: Total Reward = -200
Episode 961: Total Reward = -200
Episode 962: Total Reward = -200
Episode 963: Total Reward = -200
Episode 964: Total Reward = -200
Episode 965: Total Reward = -200
Episode 966: Total Reward = -200
Episode 967: Total Reward = -200
Episode 968: Total Reward = -200
Episode 969: Total Reward = -200
Episode 970: Total Reward = -200
Episode 971: Total Reward = -200
Episode 972: Total Reward = -200
Episode 973: Total Reward = -200
Episode 974: Total Reward = -200
Episode 975: Total Reward = -200
Episode 976: Total Reward = -200
Episode 977: Total Reward = -200
Episode 978: Total Reward = -200
Episode 979: Total Reward = -200
Episode 980: Total Reward = -200
Episode 981: Total Reward = -200
Episode 982: Total Reward = -200
Episode 983: Total Reward = -200
Episode 984: Total Reward = -200
Episode 985: Total Reward = -200
Episode 986: Total Reward = -200
Episode 987: Total Reward = -200
Episode 988: Total Reward = -200
Episode 989: Total Reward = -200
Episode 990: Total Reward = -200
Episode 991: Total Reward = -200
Episode 992: Total Reward = -200
Episode 993: Total Reward = -200
Episode 994: Total Reward = -200
Episode 995: Total Reward = -200
Episode 996: Total Reward = -200
Episode 997: Total Reward = -200
Episode 998: Total Reward = -200
Episode 999: Total Reward = -200
Episode 1000: Total Reward = -200
Average Reward across all Episodes: = -200.0
CPU times: user 8.11 s, sys: 84.4 ms, total: 8.19 s
Wall time: 7.67 s
```

In [54]:
```python
# # only to have a look of the test play result
my_dqn_agent_controlled.test(test_episodes=5, display=True)
```

Taxi-v3 | Episode: 4 | Step: 200
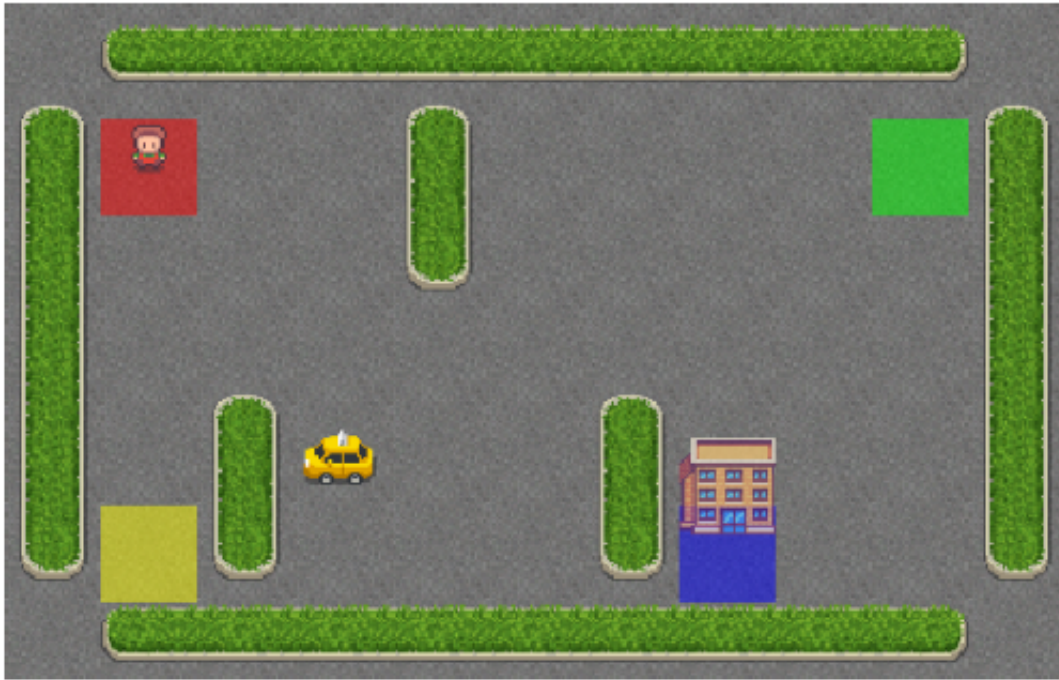
Episode 5: Total Reward = -200
Average Reward across all Episodes: = -200.0

Out[54]: [-200, -200, -200, -200, -200]

Taxi-v3 | Episode: 4 | Step: 200

In [55]: `my_dqn_agent_controlled.test_to_gif(filename="taxi_dqn_controlled",episod`

```
===== Trained Policy ===== :
State 000, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 001, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 002, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 003, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 004, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 005, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 006, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 007, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 008, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 009, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 010, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 011, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 012, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 013, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 014, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 015, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 016, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 017, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 018, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 019, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 020, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 021, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 022, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 023, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 024, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 025, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 026, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 027, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 028, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 029, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 030, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 031, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 032, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 033, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 034, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 035, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 036, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 037, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 038, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 039, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 040, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 041, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 042, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 043, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 044, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 045, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 046, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 047, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 048, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 049, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 050, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 051, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 052, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 053, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 054, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 055, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 056, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 057, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 058, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 059, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 060, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 061, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 062, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 063, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 064, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 065, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 066, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 067, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 068, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 069, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 070, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 071, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 072, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 073, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 074, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 075, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 076, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 077, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 078, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 079, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 080, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 081, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 082, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 083, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 084, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 085, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 086, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 087, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 088, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 089, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 090, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 091, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 092, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 093, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 094, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 095, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 096, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 097, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 098, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 099, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 100, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 101, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 102, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 103, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 104, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 105, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 106, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 107, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 108, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 109, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 110, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 111, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 112, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 113, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 114, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 115, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 116, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 117, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 118, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 119, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 120, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 121, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 122, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 123, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 124, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 125, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 126, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 127, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 128, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 129, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 130, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 131, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 132, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 133, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 134, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 135, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 136, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 137, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 138, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 139, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 140, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 141, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 142, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 143, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 144, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 145, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 146, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 147, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 148, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 149, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 150, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 151, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 152, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 153, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 154, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 155, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 156, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 157, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 158, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 159, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 160, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 161, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 162, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 163, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 164, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 165, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 166, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 167, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 168, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 169, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 170, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 171, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 172, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 173, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 174, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 175, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 176, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 177, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 178, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 179, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 180, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 181, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 182, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 183, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 184, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 185, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 186, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 187, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 188, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 189, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 190, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 191, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 192, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 193, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 194, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 195, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 196, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 197, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 198, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 199, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 200, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 201, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 202, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 203, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 204, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 205, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 206, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 207, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 208, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 209, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 210, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 211, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 212, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 213, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 214, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 215, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 216, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 217, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 218, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 219, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 220, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 221, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 222, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 223, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 224, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 225, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 226, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 227, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 228, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 229, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 230, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 231, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 232, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 233, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 234, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 235, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 236, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 237, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 238, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 239, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 240, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 241, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 242, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 243, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 244, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 245, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 246, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 247, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 248, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 249, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 250, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 251, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 252, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 253, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 254, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 255, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 256, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 257, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 258, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 259, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 260, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 261, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 262, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 263, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 264, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 265, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 266, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 267, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 268, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 269, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 270, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 271, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 272, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 273, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 274, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 275, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 276, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 277, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 278, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 279, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 280, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 281, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 282, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 283, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 284, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 285, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 286, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 287, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 288, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 289, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 290, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 291, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 292, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 293, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 294, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 295, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 296, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 297, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 298, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 299, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 300, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 301, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 302, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 303, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 304, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 305, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 306, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 307, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 308, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 309, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 310, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 311, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 312, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 313, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 314, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 315, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 316, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 317, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 318, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 319, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 320, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 321, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 322, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 323, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 324, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 325, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 326, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 327, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 328, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 329, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 330, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 331, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 332, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 333, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 334, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 335, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 336, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 337, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 338, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 339, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 340, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 341, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 342, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 343, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 344, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 345, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 346, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 347, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 348, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 349, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 350, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 351, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 352, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 353, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 354, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 355, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 356, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 357, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 358, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 359, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 360, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 361, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 362, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 363, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 364, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 365, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 366, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 367, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 368, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 369, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 370, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 371, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 372, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 373, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 374, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 375, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 376, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 377, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 378, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 379, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 380, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 381, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 382, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 383, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 384, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 385, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 386, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 387, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 388, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 389, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 390, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 391, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 392, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 393, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 394, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 395, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 396, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 397, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 398, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 399, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 400, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 401, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 402, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 403, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 404, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 405, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 406, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 407, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 408, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 409, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 410, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 411, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 412, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 413, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 414, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 415, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 416, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 417, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 418, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 419, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 420, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 421, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 422, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 423, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 424, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 425, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 426, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 427, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 428, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 429, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 430, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 431, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 432, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 433, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 434, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 435, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 436, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 437, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 438, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 439, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 440, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 441, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 442, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 443, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 444, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 445, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 446, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 447, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 448, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 449, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 450, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 451, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 452, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 453, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 454, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 455, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 456, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 457, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 458, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 459, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 460, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 461, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 462, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 463, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
```

```
State 464, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 465, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 466, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 467, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 468, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 469, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 470, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 471, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 472, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 473, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 474, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 475, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 476, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 477, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 478, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 479, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 480, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 481, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 482, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 483, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 484, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 485, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 486, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 487, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 488, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 489, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 490, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 491, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 492, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 493, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 494, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 495, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 496, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 497, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 498, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
State 499, Action W, Q-values: [-20.00 -20.00 -20.00 -20.00 -29.00 -28.75]
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
```

```
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
```

```
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
```

```
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
```

```
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
episodes: 0, acton: 3, state: 363
Episode 1: Total Reward = -200
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
```

```
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
```

```
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
```

```
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
episodes: 1, acton: 3, state: 308
Episode 2: Total Reward = -200
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
```

```
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
```

```
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
```

```
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
```

```
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
episodes: 2, acton: 3, state: 327
Episode 3: Total Reward = -200
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
```

```
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
```

```
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
```

```
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
episodes: 3, acton: 3, state: 474
Episode 4: Total Reward = -200
episodes: 4, acton: 3, state: 134
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
```

```
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
```

```
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
```

```
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
```

```
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
episodes: 4, acton: 3, state: 114
Episode 5: Total Reward = -200
Saving gif...Done!
Average Reward across all Episodes: = -200.0
```

In [56]:
```python
# display the captured GIF
display.Image(data=open("./gifs/taxi_dqn_controlled.gif",'rb').read(), fo
# "taxi_q_controlled"
```

Out[56]: `<IPython.core.display.Image object>`

If the above display function is not working, here is the previously prepared taxi_dqn_controlled.gif
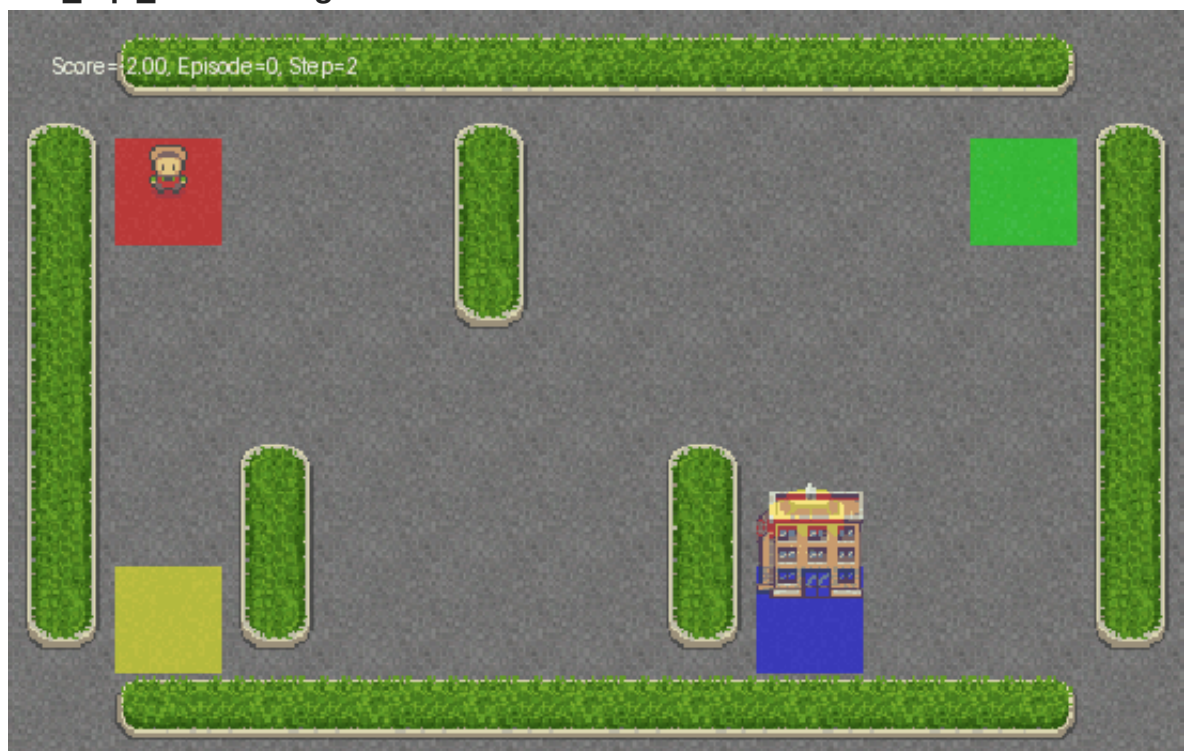
Compare the prediction of controlled models

In [59]:
```python
#my_q_agent_controlled_avg_reward_lst
#my_dqn_agent_controlled_avg_reward_lst
```

In [60]:
```python
x = range(1,1000 + 1)
```

```
plt.figure(figsize=(10, 6))
plt.plot(x, my_dqn_agent_controlled_avg_reward_lst, label='DQN', marker='
plt.plot(x, my_q_agent_controlled_avg_reward_lst, label='Q-learning', mar

# titles and labels
plt.title('Training Comparison (Controlled Parameters)')
plt.xlabel('Time Episodes')
plt.ylabel('Average Reward')
plt.legend()
# display
plt.grid(True)
plt.show()
```
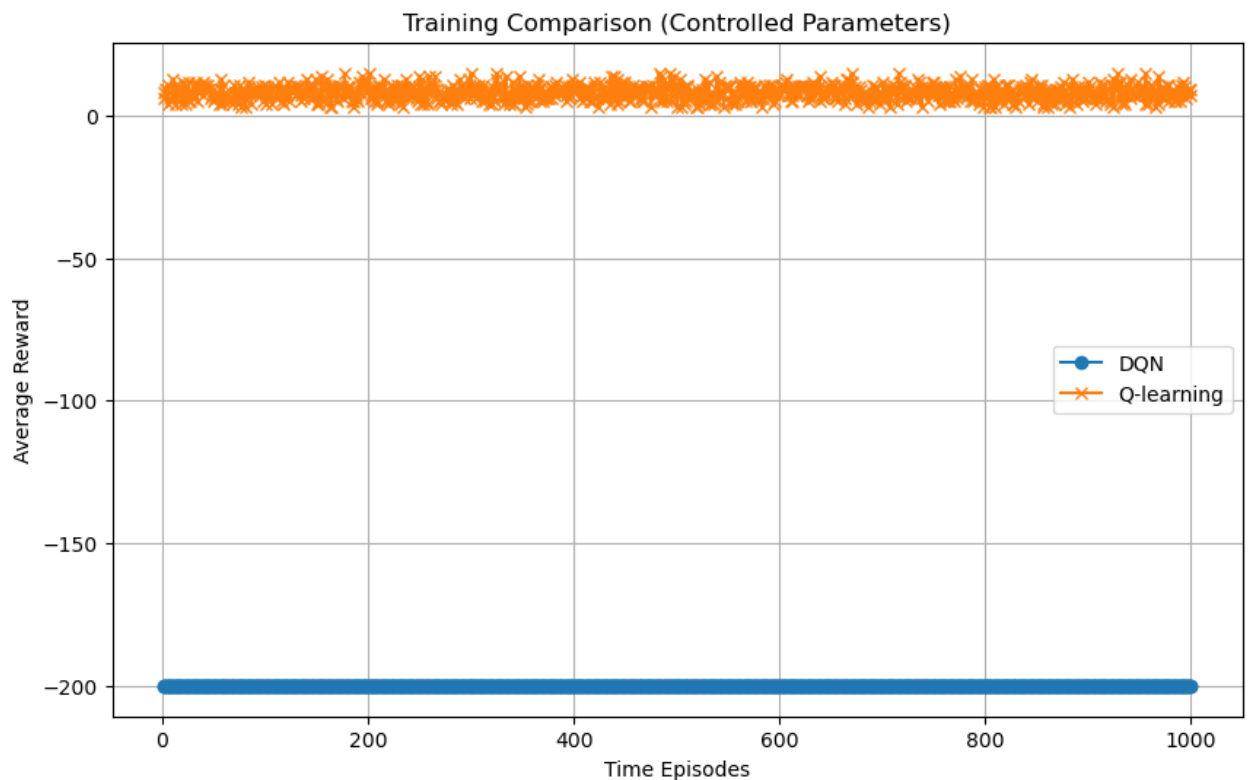

Fine-tunred Parameter Training

Q Learning Agent with Fine-Tuned Parameters

```
In [63]: my_q_agent_fine_tuned = QLearningAgent(game_id="Taxi-v3", render_mode="rg
                                    gamma=0.90, alpha=0.5, epsilon=1.0
                                    epsilon_decay=1000, epsilon_end=0.
                                    result_model_name="taxi_q_fine_tun
```

```
In [64]: %%time
         my_q_agent_fine_tuned_training_rewards = my_q_agent_fine_tuned.train(1000
```

```
Episode 100/100000, Avg Reward: -465.18, Epsilon: 0.32
Episode 200/100000, Avg Reward: -142.22, Epsilon: 0.17
Episode 300/100000, Avg Reward: -36.70, Epsilon: 0.13
Episode 400/100000, Avg Reward: -7.76, Epsilon: 0.11
Episode 500/100000, Avg Reward: -0.62, Epsilon: 0.10
Episode 600/100000, Avg Reward: -0.22, Epsilon: 0.09
Episode 700/100000, Avg Reward: 0.70, Epsilon: 0.08
Episode 800/100000, Avg Reward: 2.45, Epsilon: 0.07
```

```
Episode 900/100000, Avg Reward: 3.85, Epsilon: 0.06
Episode 1000/100000, Avg Reward: 4.94, Epsilon: 0.06
Episode 1100/100000, Avg Reward: 4.78, Epsilon: 0.05
Episode 1200/100000, Avg Reward: 4.92, Epsilon: 0.05
Episode 1300/100000, Avg Reward: 5.68, Epsilon: 0.04
Episode 1400/100000, Avg Reward: 5.50, Epsilon: 0.04
Episode 1500/100000, Avg Reward: 5.83, Epsilon: 0.04
Episode 1600/100000, Avg Reward: 5.92, Epsilon: 0.03
Episode 1700/100000, Avg Reward: 6.80, Epsilon: 0.03
Episode 1800/100000, Avg Reward: 6.25, Epsilon: 0.03
Episode 1900/100000, Avg Reward: 6.95, Epsilon: 0.02
Episode 2000/100000, Avg Reward: 7.17, Epsilon: 0.02
Episode 2100/100000, Avg Reward: 6.68, Epsilon: 0.02
Episode 2200/100000, Avg Reward: 7.42, Epsilon: 0.02
Episode 2300/100000, Avg Reward: 7.25, Epsilon: 0.02
Episode 2400/100000, Avg Reward: 6.69, Epsilon: 0.02
Episode 2500/100000, Avg Reward: 6.71, Epsilon: 0.01
Episode 2600/100000, Avg Reward: 6.80, Epsilon: 0.01
Episode 2700/100000, Avg Reward: 6.56, Epsilon: 0.01
Episode 2800/100000, Avg Reward: 7.59, Epsilon: 0.01
Episode 2900/100000, Avg Reward: 8.22, Epsilon: 0.01
Episode 3000/100000, Avg Reward: 7.73, Epsilon: 0.01
Episode 3100/100000, Avg Reward: 7.78, Epsilon: 0.01
Episode 3200/100000, Avg Reward: 7.06, Epsilon: 0.01
Episode 3300/100000, Avg Reward: 7.73, Epsilon: 0.01
Episode 3400/100000, Avg Reward: 7.24, Epsilon: 0.01
Episode 3500/100000, Avg Reward: 7.33, Epsilon: 0.01
Episode 3600/100000, Avg Reward: 7.61, Epsilon: 0.01
Episode 3700/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 3800/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 3900/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 4000/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 4100/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 4200/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 4300/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 4400/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 4500/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 4600/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 4700/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 4800/100000, Avg Reward: 7.28, Epsilon: 0.00
Episode 4900/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 5000/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 5100/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 5200/100000, Avg Reward: 7.20, Epsilon: 0.00
Episode 5300/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 5400/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 5500/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 5600/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 5700/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 5800/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 5900/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 6000/100000, Avg Reward: 8.33, Epsilon: 0.00
Episode 6100/100000, Avg Reward: 7.27, Epsilon: 0.00
Episode 6200/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 6300/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 6400/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 6500/100000, Avg Reward: 8.23, Epsilon: 0.00
```

```
Episode 6600/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 6700/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 6800/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 6900/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 7000/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 7100/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 7200/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 7300/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 7400/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 7500/100000, Avg Reward: 8.58, Epsilon: 0.00
Episode 7600/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 7700/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 7800/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 7900/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 8000/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 8100/100000, Avg Reward: 8.36, Epsilon: 0.00
Episode 8200/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 8300/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 8400/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 8500/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 8600/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 8700/100000, Avg Reward: 7.27, Epsilon: 0.00
Episode 8800/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 8900/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 9000/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 9100/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 9200/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 9300/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 9400/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 9500/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 9600/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 9700/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 9800/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 9900/100000, Avg Reward: 8.40, Epsilon: 0.00
Episode 10000/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 10100/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 10200/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 10300/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 10400/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 10500/100000, Avg Reward: 8.57, Epsilon: 0.00
Episode 10600/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 10700/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 10800/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 10900/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 11000/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 11100/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 11200/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 11300/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 11400/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 11500/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 11600/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 11700/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 11800/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 11900/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 12000/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 12100/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 12200/100000, Avg Reward: 8.23, Epsilon: 0.00
```

```
Episode 12300/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 12400/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 12500/100000, Avg Reward: 8.41, Epsilon: 0.00
Episode 12600/100000, Avg Reward: 7.45, Epsilon: 0.00
Episode 12700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 12800/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 12900/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 13000/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 13100/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 13200/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 13300/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 13400/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 13500/100000, Avg Reward: 7.39, Epsilon: 0.00
Episode 13600/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 13700/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 13800/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 13900/100000, Avg Reward: 7.44, Epsilon: 0.00
Episode 14000/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 14100/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 14200/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 14300/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 14400/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 14500/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 14600/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 14700/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 14800/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 14900/100000, Avg Reward: 8.48, Epsilon: 0.00
Episode 15000/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 15100/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 15200/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 15300/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 15400/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 15500/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 15600/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 15700/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 15800/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 15900/100000, Avg Reward: 8.47, Epsilon: 0.00
Episode 16000/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 16100/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 16200/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 16300/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 16400/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 16500/100000, Avg Reward: 7.33, Epsilon: 0.00
Episode 16600/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 16700/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 16800/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 16900/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 17000/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 17100/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 17200/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 17300/100000, Avg Reward: 7.47, Epsilon: 0.00
Episode 17400/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 17500/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 17600/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 17700/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 17800/100000, Avg Reward: 7.47, Epsilon: 0.00
Episode 17900/100000, Avg Reward: 7.87, Epsilon: 0.00
```

```
Episode 18000/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 18100/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 18200/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 18300/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 18400/100000, Avg Reward: 8.43, Epsilon: 0.00
Episode 18500/100000, Avg Reward: 8.42, Epsilon: 0.00
Episode 18600/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 18700/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 18800/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 18900/100000, Avg Reward: 7.42, Epsilon: 0.00
Episode 19000/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 19100/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 19200/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 19300/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 19400/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 19500/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 19600/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 19700/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 19800/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 19900/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 20000/100000, Avg Reward: 8.43, Epsilon: 0.00
Episode 20100/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 20200/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 20300/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 20400/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 20500/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 20600/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 20700/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 20800/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 20900/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 21000/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 21100/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 21200/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 21300/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 21400/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 21500/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 21600/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 21700/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 21800/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 21900/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 22000/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 22100/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 22200/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 22300/100000, Avg Reward: 7.33, Epsilon: 0.00
Episode 22400/100000, Avg Reward: 7.47, Epsilon: 0.00
Episode 22500/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 22600/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 22700/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 22800/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 22900/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 23000/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 23100/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 23200/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 23300/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 23400/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 23500/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 23600/100000, Avg Reward: 8.36, Epsilon: 0.00
```

```
Episode 23700/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 23800/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 23900/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 24000/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 24100/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 24200/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 24300/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 24400/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 24500/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 24600/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 24700/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 24800/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 24900/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 25000/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 25100/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 25200/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 25300/100000, Avg Reward: 7.32, Epsilon: 0.00
Episode 25400/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 25500/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 25600/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 25700/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 25800/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 25900/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 26000/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 26100/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 26200/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 26300/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 26400/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 26500/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 26600/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 26700/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 26800/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 26900/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 27000/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 27100/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 27200/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 27300/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 27400/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 27500/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 27600/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 27700/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 27800/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 27900/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 28000/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 28100/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 28200/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 28300/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 28400/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 28500/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 28600/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 28700/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 28800/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 28900/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 29000/100000, Avg Reward: 7.35, Epsilon: 0.00
Episode 29100/100000, Avg Reward: 7.32, Epsilon: 0.00
Episode 29200/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 29300/100000, Avg Reward: 7.83, Epsilon: 0.00
```

```
Episode 29400/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 29500/100000, Avg Reward: 7.32, Epsilon: 0.00
Episode 29600/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 29700/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 29800/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 29900/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 30000/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 30100/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 30200/100000, Avg Reward: 8.40, Epsilon: 0.00
Episode 30300/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 30400/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 30500/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 30600/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 30700/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 30800/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 30900/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 31000/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 31100/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 31200/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 31300/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 31400/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 31500/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 31600/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 31700/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 31800/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 31900/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 32000/100000, Avg Reward: 6.98, Epsilon: 0.00
Episode 32100/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 32200/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 32300/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 32400/100000, Avg Reward: 8.41, Epsilon: 0.00
Episode 32500/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 32600/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 32700/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 32800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 32900/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 33000/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 33100/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 33200/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 33300/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 33400/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 33500/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 33600/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 33700/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 33800/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 33900/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 34000/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 34100/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 34200/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 34300/100000, Avg Reward: 8.51, Epsilon: 0.00
Episode 34400/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 34500/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 34600/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 34700/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 34800/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 34900/100000, Avg Reward: 7.14, Epsilon: 0.00
Episode 35000/100000, Avg Reward: 7.70, Epsilon: 0.00
```

```
Episode 35100/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 35200/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 35300/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 35400/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 35500/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 35600/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 35700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 35800/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 35900/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 36000/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 36100/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 36200/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 36300/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 36400/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 36500/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 36600/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 36700/100000, Avg Reward: 8.41, Epsilon: 0.00
Episode 36800/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 36900/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 37000/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 37100/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 37200/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 37300/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 37400/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 37500/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 37600/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 37700/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 37800/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 37900/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 38000/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 38100/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 38200/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 38300/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 38400/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 38500/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 38600/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 38700/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 38800/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 38900/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 39000/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 39100/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 39200/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 39300/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 39400/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 39500/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 39600/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 39700/100000, Avg Reward: 8.52, Epsilon: 0.00
Episode 39800/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 39900/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 40000/100000, Avg Reward: 8.53, Epsilon: 0.00
Episode 40100/100000, Avg Reward: 8.47, Epsilon: 0.00
Episode 40200/100000, Avg Reward: 7.23, Epsilon: 0.00
Episode 40300/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 40400/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 40500/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 40600/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 40700/100000, Avg Reward: 7.64, Epsilon: 0.00
```

```
Episode 40800/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 40900/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 41000/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 41100/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 41200/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 41300/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 41400/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 41500/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 41600/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 41700/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 41800/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 41900/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 42000/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 42100/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 42200/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 42300/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 42400/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 42500/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 42600/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 42700/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 42800/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 42900/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 43000/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 43100/100000, Avg Reward: 7.36, Epsilon: 0.00
Episode 43200/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 43300/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 43400/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 43500/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 43600/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 43700/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 43800/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 43900/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 44000/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 44100/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 44200/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 44300/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 44400/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 44500/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 44600/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 44700/100000, Avg Reward: 7.18, Epsilon: 0.00
Episode 44800/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 44900/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 45000/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 45100/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 45200/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 45300/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 45400/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 45500/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 45600/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 45700/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 45800/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 45900/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 46000/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 46100/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 46200/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 46300/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 46400/100000, Avg Reward: 8.19, Epsilon: 0.00
```

```
Episode 46500/100000, Avg Reward: 8.33, Epsilon: 0.00
Episode 46600/100000, Avg Reward: 8.85, Epsilon: 0.00
Episode 46700/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 46800/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 46900/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 47000/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 47100/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 47200/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 47300/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 47400/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 47500/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 47600/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 47700/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 47800/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 47900/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 48000/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 48100/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 48200/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 48300/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 48400/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 48500/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 48600/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 48700/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 48800/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 48900/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 49000/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 49100/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 49200/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 49300/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 49400/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 49500/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 49600/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 49700/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 49800/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 49900/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 50000/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 50100/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 50200/100000, Avg Reward: 7.30, Epsilon: 0.00
Episode 50300/100000, Avg Reward: 8.36, Epsilon: 0.00
Episode 50400/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 50500/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 50600/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 50700/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 50800/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 50900/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 51000/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 51100/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 51200/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 51300/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 51400/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 51500/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 51600/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 51700/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 51800/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 51900/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 52000/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 52100/100000, Avg Reward: 7.98, Epsilon: 0.00
```

```
Episode 52200/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 52300/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 52400/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 52500/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 52600/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 52700/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 52800/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 52900/100000, Avg Reward: 7.50, Epsilon: 0.00
Episode 53000/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 53100/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 53200/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 53300/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 53400/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 53500/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 53600/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 53700/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 53800/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 53900/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 54000/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 54100/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 54200/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 54300/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 54400/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 54500/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 54600/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 54700/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 54800/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 54900/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 55000/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 55100/100000, Avg Reward: 7.19, Epsilon: 0.00
Episode 55200/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 55300/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 55400/100000, Avg Reward: 7.20, Epsilon: 0.00
Episode 55500/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 55600/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 55700/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 55800/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 55900/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 56000/100000, Avg Reward: 8.70, Epsilon: 0.00
Episode 56100/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 56200/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 56300/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 56400/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 56500/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 56600/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 56700/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 56800/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 56900/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 57000/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 57100/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 57200/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 57300/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 57400/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 57500/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 57600/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 57700/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 57800/100000, Avg Reward: 8.07, Epsilon: 0.00
```

```
Episode 57900/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 58000/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 58100/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 58200/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 58300/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 58400/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 58500/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 58600/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 58700/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 58800/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 58900/100000, Avg Reward: 8.34, Epsilon: 0.00
Episode 59000/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 59100/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 59200/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 59300/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 59400/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 59500/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 59600/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 59700/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 59800/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 59900/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 60000/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 60100/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 60200/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 60300/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 60400/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 60500/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 60600/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 60700/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 60800/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 60900/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 61000/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 61100/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 61200/100000, Avg Reward: 7.30, Epsilon: 0.00
Episode 61300/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 61400/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 61500/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 61600/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 61700/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 61800/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 61900/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 62000/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 62100/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 62200/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 62300/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 62400/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 62500/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 62600/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 62700/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 62800/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 62900/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 63000/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 63100/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 63200/100000, Avg Reward: 7.39, Epsilon: 0.00
Episode 63300/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 63400/100000, Avg Reward: 8.37, Epsilon: 0.00
Episode 63500/100000, Avg Reward: 8.00, Epsilon: 0.00
```

```
Episode 63600/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 63700/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 63800/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 63900/100000, Avg Reward: 8.41, Epsilon: 0.00
Episode 64000/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 64100/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 64200/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 64300/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 64400/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 64500/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 64600/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 64700/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 64800/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 64900/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 65000/100000, Avg Reward: 8.49, Epsilon: 0.00
Episode 65100/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 65200/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 65300/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 65400/100000, Avg Reward: 7.45, Epsilon: 0.00
Episode 65500/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 65600/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 65700/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 65800/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 65900/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 66000/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 66100/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 66200/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 66300/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 66400/100000, Avg Reward: 7.47, Epsilon: 0.00
Episode 66500/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 66600/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 66700/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 66800/100000, Avg Reward: 8.33, Epsilon: 0.00
Episode 66900/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 67000/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 67100/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 67200/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 67300/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 67400/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 67500/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 67600/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 67700/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 67800/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 67900/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 68000/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 68100/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 68200/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 68300/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 68400/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 68500/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 68600/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 68700/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 68800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 68900/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 69000/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 69100/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 69200/100000, Avg Reward: 8.15, Epsilon: 0.00
```

```
Episode 69300/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 69400/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 69500/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 69600/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 69700/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 69800/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 69900/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 70000/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 70100/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 70200/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 70300/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 70400/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 70500/100000, Avg Reward: 7.25, Epsilon: 0.00
Episode 70600/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 70700/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 70800/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 70900/100000, Avg Reward: 8.21, Epsilon: 0.00
Episode 71000/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 71100/100000, Avg Reward: 8.24, Epsilon: 0.00
Episode 71200/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 71300/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 71400/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 71500/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 71600/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 71700/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 71800/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 71900/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 72000/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 72100/100000, Avg Reward: 7.20, Epsilon: 0.00
Episode 72200/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 72300/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 72400/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 72500/100000, Avg Reward: 8.47, Epsilon: 0.00
Episode 72600/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 72700/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 72800/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 72900/100000, Avg Reward: 7.33, Epsilon: 0.00
Episode 73000/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 73100/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 73200/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 73300/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 73400/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 73500/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 73600/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 73700/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 73800/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 73900/100000, Avg Reward: 7.32, Epsilon: 0.00
Episode 74000/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 74100/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 74200/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 74300/100000, Avg Reward: 7.37, Epsilon: 0.00
Episode 74400/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 74500/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 74600/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 74700/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 74800/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 74900/100000, Avg Reward: 8.03, Epsilon: 0.00
```

```
Episode 75000/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 75100/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 75200/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 75300/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 75400/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 75500/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 75600/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 75700/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 75800/100000, Avg Reward: 7.31, Epsilon: 0.00
Episode 75900/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 76000/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 76100/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 76200/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 76300/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 76400/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 76500/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 76600/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 76700/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 76800/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 76900/100000, Avg Reward: 8.59, Epsilon: 0.00
Episode 77000/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 77100/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 77200/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 77300/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 77400/100000, Avg Reward: 8.33, Epsilon: 0.00
Episode 77500/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 77600/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 77700/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 77800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 77900/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 78000/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 78100/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 78200/100000, Avg Reward: 7.42, Epsilon: 0.00
Episode 78300/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 78400/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 78500/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 78600/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 78700/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 78800/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 78900/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 79000/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 79100/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 79200/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 79300/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 79400/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 79500/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 79600/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 79700/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 79800/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 79900/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 80000/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 80100/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 80200/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 80300/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 80400/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 80500/100000, Avg Reward: 8.70, Epsilon: 0.00
Episode 80600/100000, Avg Reward: 7.67, Epsilon: 0.00
```

```
Episode 80700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 80800/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 80900/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 81000/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 81100/100000, Avg Reward: 8.62, Epsilon: 0.00
Episode 81200/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 81300/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 81400/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 81500/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 81600/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 81700/100000, Avg Reward: 8.25, Epsilon: 0.00
Episode 81800/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 81900/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 82000/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 82100/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 82200/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 82300/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 82400/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 82500/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 82600/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 82700/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 82800/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 82900/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 83000/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 83100/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 83200/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 83300/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 83400/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 83500/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 83600/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 83700/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 83800/100000, Avg Reward: 8.47, Epsilon: 0.00
Episode 83900/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 84000/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 84100/100000, Avg Reward: 7.98, Epsilon: 0.00
Episode 84200/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 84300/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 84400/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 84500/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 84600/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 84700/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 84800/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 84900/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 85000/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 85100/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 85200/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 85300/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 85400/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 85500/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 85600/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 85700/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 85800/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 85900/100000, Avg Reward: 7.52, Epsilon: 0.00
Episode 86000/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 86100/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 86200/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 86300/100000, Avg Reward: 7.90, Epsilon: 0.00
```

```
Episode 86400/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 86500/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 86600/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 86700/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 86800/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 86900/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 87000/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 87100/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 87200/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 87300/100000, Avg Reward: 8.35, Epsilon: 0.00
Episode 87400/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 87500/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 87600/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 87700/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 87800/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 87900/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 88000/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 88100/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 88200/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 88300/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 88400/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 88500/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 88600/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 88700/100000, Avg Reward: 8.13, Epsilon: 0.00
Episode 88800/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 88900/100000, Avg Reward: 7.27, Epsilon: 0.00
Episode 89000/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 89100/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 89200/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 89300/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 89400/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 89500/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 89600/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 89700/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 89800/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 89900/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 90000/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 90100/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 90200/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 90300/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 90400/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 90500/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 90600/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 90700/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 90800/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 90900/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 91000/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 91100/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 91200/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 91300/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 91400/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 91500/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 91600/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 91700/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 91800/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 91900/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 92000/100000, Avg Reward: 8.04, Epsilon: 0.00
```

```
Episode 92100/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 92200/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 92300/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 92400/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 92500/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 92600/100000, Avg Reward: 8.16, Epsilon: 0.00
Episode 92700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 92800/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 92900/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 93000/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 93100/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 93200/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 93300/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 93400/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 93500/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 93600/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 93700/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 93800/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 93900/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 94000/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 94100/100000, Avg Reward: 8.45, Epsilon: 0.00
Episode 94200/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 94300/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 94400/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 94500/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 94600/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 94700/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 94800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 94900/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 95000/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 95100/100000, Avg Reward: 7.56, Epsilon: 0.00
Episode 95200/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 95300/100000, Avg Reward: 7.50, Epsilon: 0.00
Episode 95400/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 95500/100000, Avg Reward: 8.31, Epsilon: 0.00
Episode 95600/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 95700/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 95800/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 95900/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 96000/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 96100/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 96200/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 96300/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 96400/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 96500/100000, Avg Reward: 8.38, Epsilon: 0.00
Episode 96600/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 96700/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 96800/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 96900/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 97000/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 97100/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 97200/100000, Avg Reward: 8.62, Epsilon: 0.00
Episode 97300/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 97400/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 97500/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 97600/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 97700/100000, Avg Reward: 8.29, Epsilon: 0.00
```

```
Episode 97800/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 97900/100000, Avg Reward: 8.03, Epsilon: 0.00
Episode 98000/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 98100/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 98200/100000, Avg Reward: 8.19, Epsilon: 0.00
Episode 98300/100000, Avg Reward: 8.32, Epsilon: 0.00
Episode 98400/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 98500/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 98600/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 98700/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 98800/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 98900/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 99000/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 99100/100000, Avg Reward: 8.22, Epsilon: 0.00
Episode 99200/100000, Avg Reward: 8.14, Epsilon: 0.00
Episode 99300/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 99400/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 99500/100000, Avg Reward: 8.09, Epsilon: 0.00
Episode 99600/100000, Avg Reward: 8.23, Epsilon: 0.00
Episode 99700/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 99800/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 99900/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 100000/100000, Avg Reward: 7.56, Epsilon: 0.00
CPU times: user 11.9 s, sys: 461 ms, total: 12.4 s
Wall time: 11.5 s
```
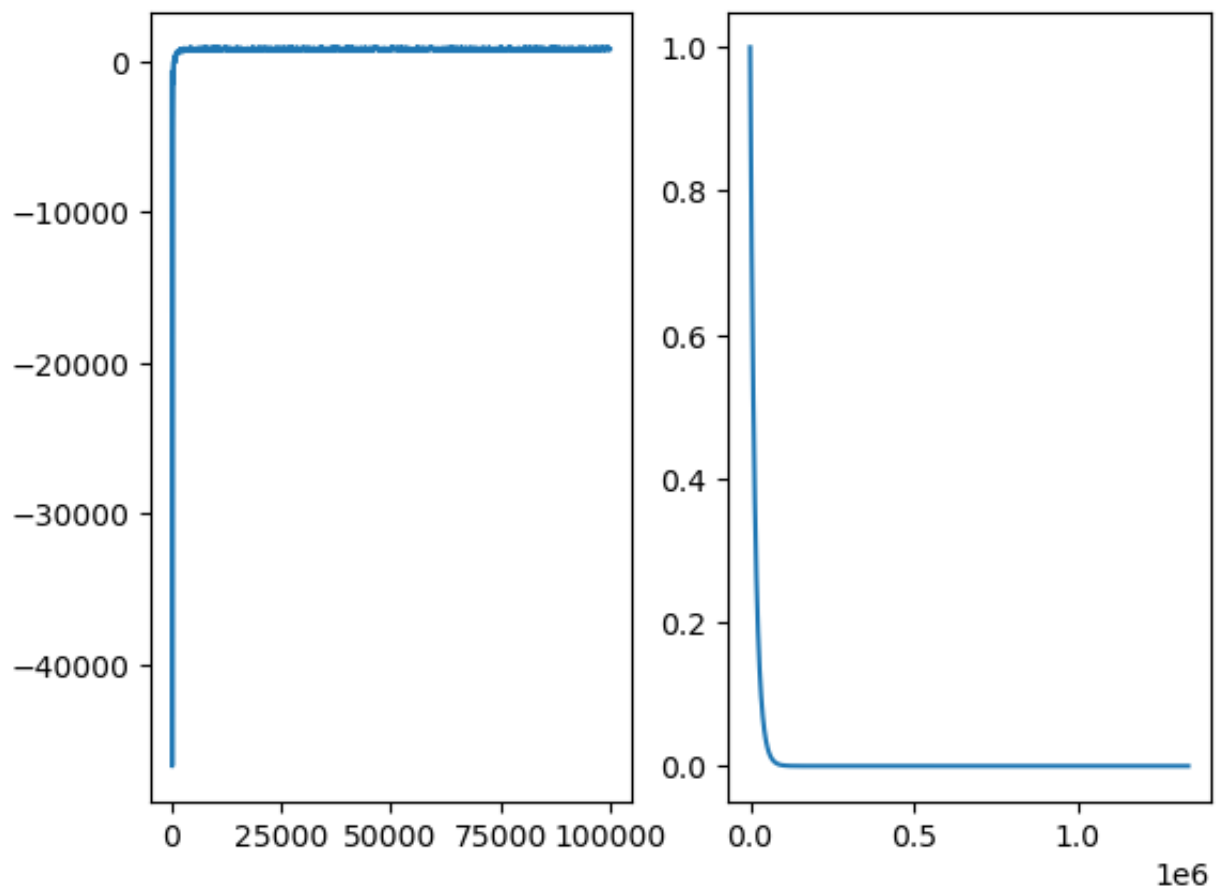

In [65]: `print(my_q_agent_fine_tuned_training_rewards)`

```
[-794. -839. -704. ...    4.    9.   11.]
```

DQN Agent with Fine-Tuned Parameters

```python
In [67]: my_dqn_agent_fine_tuned = DQNAgent(game_id="Taxi-v3", gamma=0.90, alpha=0
                                            epsilon_start=1.0, epsilon_decay=1000,
                                            max_steps=200, render_mode="rgb_array"
```

```python
In [68]: %%time
         my_dqn_agent_fine_tuned_training_rewards = my_dqn_agent_fine_tuned.train(
```

```
Non-trained Policy Network Model with randomly initialized values:
State 000, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 001, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 002, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 003, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 004, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 005, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 006, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 007, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 008, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 009, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.06]
State 010, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 011, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 012, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 013, Action N, Q-values: [+0.03 +0.09 -0.07 +0.09 +0.07 -0.05]
State 014, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.09 -0.05]
State 015, Action N, Q-values: [+0.03 +0.09 -0.07 +0.09 +0.08 -0.06]
State 016, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 017, Action N, Q-values: [+0.04 +0.10 -0.07 +0.08 +0.07 -0.06]
State 018, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.07 -0.06]
State 019, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 020, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 021, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.06]
State 022, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 023, Action P, Q-values: [+0.04 +0.08 -0.09 +0.08 +0.09 -0.05]
State 024, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 025, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 026, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 027, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 028, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 029, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 030, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 031, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 032, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.05]
State 033, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 034, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 035, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.07 -0.06]
State 036, Action P, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.06]
State 037, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 038, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 039, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 040, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.09 -0.05]
State 041, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 042, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 043, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 044, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 045, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.04]
State 046, Action W, Q-values: [+0.04 +0.09 -0.09 +0.09 +0.09 -0.05]
State 047, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 048, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
```

```
State 049, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.04]
State 050, Action P, Q-values: [+0.03 +0.08 -0.09 +0.08 +0.09 -0.05]
State 051, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 052, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 053, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 054, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 055, Action N, Q-values: [+0.04 +0.09 -0.08 +0.07 +0.08 -0.06]
State 056, Action N, Q-values: [+0.03 +0.09 -0.07 +0.09 +0.08 -0.05]
State 057, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 058, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 059, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.07 -0.05]
State 060, Action P, Q-values: [+0.04 +0.08 -0.08 +0.09 +0.09 -0.06]
State 061, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 062, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.06]
State 063, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.07 -0.05]
State 064, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 065, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 066, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 067, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 068, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 069, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 070, Action P, Q-values: [+0.04 +0.09 -0.09 +0.09 +0.09 -0.05]
State 071, Action N, Q-values: [+0.04 +0.08 -0.07 +0.08 +0.08 -0.06]
State 072, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 073, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 074, Action N, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.08 -0.05]
State 075, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.06]
State 076, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 077, Action N, Q-values: [+0.04 +0.08 -0.07 +0.08 +0.08 -0.06]
State 078, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 079, Action W, Q-values: [+0.04 +0.08 -0.07 +0.09 +0.08 -0.06]
State 080, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 081, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 082, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 083, Action P, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.05]
State 084, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 085, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 086, Action W, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 087, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 088, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 089, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 090, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 091, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.04]
State 092, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 093, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 094, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 095, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.06]
State 096, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 097, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.05]
State 098, Action N, Q-values: [+0.05 +0.09 -0.08 +0.08 +0.09 -0.05]
State 099, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 100, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.05]
State 101, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 102, Action P, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.05]
State 103, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 104, Action N, Q-values: [+0.04 +0.10 -0.08 +0.07 +0.07 -0.06]
State 105, Action N, Q-values: [+0.03 +0.09 -0.07 +0.09 +0.08 -0.05]
```

```
State 106, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 107, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 108, Action W, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 109, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 110, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 111, Action N, Q-values: [+0.04 +0.09 -0.08 +0.07 +0.08 -0.06]
State 112, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 113, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 114, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 115, Action P, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 116, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 117, Action W, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.06]
State 118, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.06]
State 119, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 120, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.06]
State 121, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 122, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 123, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 124, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 125, Action P, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.05]
State 126, Action N, Q-values: [+0.03 +0.10 -0.07 +0.09 +0.07 -0.06]
State 127, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 128, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 129, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 130, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.05]
State 131, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.08 -0.05]
State 132, Action P, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 133, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 134, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.05]
State 135, Action N, Q-values: [+0.03 +0.09 -0.07 +0.09 +0.08 -0.05]
State 136, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 137, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 138, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 139, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 140, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 141, Action P, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.05]
State 142, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 143, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 144, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 145, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 146, Action N, Q-values: [+0.03 +0.10 -0.08 +0.09 +0.09 -0.06]
State 147, Action P, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 148, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 149, Action W, Q-values: [+0.03 +0.08 -0.07 +0.09 +0.08 -0.05]
State 150, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 151, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 152, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 153, Action N, Q-values: [+0.04 +0.10 -0.07 +0.08 +0.08 -0.05]
State 154, Action P, Q-values: [+0.04 +0.08 -0.09 +0.08 +0.09 -0.05]
State 155, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 156, Action P, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.09 -0.04]
State 157, Action W, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.06]
State 158, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 159, Action W, Q-values: [+0.03 +0.08 -0.08 +0.09 +0.09 -0.05]
State 160, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 161, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 162, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
```

```
State 163, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 164, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.04]
State 165, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 166, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.06]
State 167, Action N, Q-values: [+0.04 +0.08 -0.07 +0.08 +0.08 -0.06]
State 168, Action P, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.05]
State 169, Action N, Q-values: [+0.04 +0.10 -0.06 +0.09 +0.08 -0.06]
State 170, Action N, Q-values: [+0.03 +0.09 -0.07 +0.09 +0.08 -0.05]
State 171, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 172, Action N, Q-values: [+0.04 +0.10 -0.07 +0.09 +0.08 -0.06]
State 173, Action N, Q-values: [+0.05 +0.09 -0.08 +0.08 +0.08 -0.05]
State 174, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 175, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 176, Action N, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.08 -0.06]
State 177, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 178, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 179, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.04]
State 180, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 181, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 182, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.09 -0.05]
State 183, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 184, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 185, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.09 -0.05]
State 186, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.06]
State 187, Action P, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.09 -0.05]
State 188, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 189, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 190, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 191, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 192, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 193, Action P, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.05]
State 194, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 195, Action N, Q-values: [+0.04 +0.10 -0.09 +0.09 +0.08 -0.05]
State 196, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 197, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 198, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 199, Action N, Q-values: [+0.05 +0.09 -0.08 +0.08 +0.08 -0.06]
State 200, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 201, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 202, Action N, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.08 -0.05]
State 203, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 204, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 205, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 206, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 207, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 208, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.06]
State 209, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 210, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 211, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 212, Action W, Q-values: [+0.03 +0.08 -0.08 +0.09 +0.09 -0.05]
State 213, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.05]
State 214, Action P, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.06]
State 215, Action W, Q-values: [+0.04 +0.08 -0.07 +0.08 +0.08 -0.06]
State 216, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 217, Action P, Q-values: [+0.03 +0.08 -0.09 +0.08 +0.09 -0.05]
State 218, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 219, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
```

```
State 220, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 221, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.06]
State 222, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 223, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 224, Action W, Q-values: [+0.03 +0.08 -0.08 +0.09 +0.08 -0.05]
State 225, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 226, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 227, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 228, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.06]
State 229, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 230, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 231, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 232, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 233, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.06]
State 234, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 235, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 236, Action N, Q-values: [+0.03 +0.10 -0.08 +0.09 +0.08 -0.05]
State 237, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 238, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 239, Action P, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.09 -0.05]
State 240, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.09 -0.05]
State 241, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 242, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 243, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.06]
State 244, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 245, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.06]
State 246, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 247, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 248, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 249, Action N, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.08 -0.05]
State 250, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 251, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.08 -0.05]
State 252, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 253, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 254, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 255, Action P, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.05]
State 256, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 257, Action N, Q-values: [+0.05 +0.09 -0.07 +0.08 +0.08 -0.06]
State 258, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 259, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 260, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 261, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 262, Action N, Q-values: [+0.05 +0.09 -0.09 +0.08 +0.08 -0.06]
State 263, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 264, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 265, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.06]
State 266, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.04]
State 267, Action P, Q-values: [+0.04 +0.08 -0.08 +0.07 +0.09 -0.05]
State 268, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 269, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 270, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 271, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.05]
State 272, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.07 -0.06]
State 273, Action W, Q-values: [+0.03 +0.08 -0.08 +0.09 +0.08 -0.05]
State 274, Action P, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.09 -0.04]
State 275, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.06]
State 276, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
```

```
State 277, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 278, Action N, Q-values: [+0.04 +0.10 -0.08 +0.09 +0.08 -0.06]
State 279, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.05]
State 280, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.04]
State 281, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 282, Action P, Q-values: [+0.05 +0.09 -0.08 +0.08 +0.09 -0.05]
State 283, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 284, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 285, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 286, Action P, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.06]
State 287, Action W, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 288, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.06]
State 289, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 290, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.07 -0.06]
State 291, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 292, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 293, Action W, Q-values: [+0.04 +0.08 -0.08 +0.09 +0.09 -0.05]
State 294, Action P, Q-values: [+0.03 +0.08 -0.09 +0.08 +0.08 -0.06]
State 295, Action W, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 296, Action N, Q-values: [+0.04 +0.08 -0.07 +0.08 +0.08 -0.05]
State 297, Action P, Q-values: [+0.03 +0.08 -0.09 +0.08 +0.09 -0.05]
State 298, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 299, Action W, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 300, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 301, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.05]
State 302, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 303, Action P, Q-values: [+0.04 +0.08 -0.09 +0.08 +0.08 -0.05]
State 304, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 305, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 306, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.06]
State 307, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 308, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 309, Action P, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.05]
State 310, Action P, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.09 -0.05]
State 311, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.08 -0.06]
State 312, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.06]
State 313, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.09 -0.05]
State 314, Action N, Q-values: [+0.04 +0.09 -0.08 +0.07 +0.08 -0.06]
State 315, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 316, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 317, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 318, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 319, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 320, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 321, Action P, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.04]
State 322, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 323, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.09 -0.05]
State 324, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 325, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 326, Action P, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.05]
State 327, Action N, Q-values: [+0.04 +0.10 -0.08 +0.09 +0.08 -0.06]
State 328, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.05]
State 329, Action P, Q-values: [+0.04 +0.08 -0.08 +0.09 +0.09 -0.05]
State 330, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 331, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.06]
State 332, Action P, Q-values: [+0.04 +0.08 -0.07 +0.09 +0.09 -0.05]
State 333, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.06]
```

```
State 334, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 335, Action W, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.09 -0.05]
State 336, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.05]
State 337, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.06]
State 338, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.06]
State 339, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 340, Action W, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.06]
State 341, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.09 -0.05]
State 342, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 343, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 344, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.07 -0.06]
State 345, Action N, Q-values: [+0.03 +0.09 -0.07 +0.09 +0.09 -0.05]
State 346, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 347, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.06]
State 348, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 349, Action P, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 350, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.08 -0.06]
State 351, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 352, Action P, Q-values: [+0.04 +0.08 -0.09 +0.08 +0.09 -0.05]
State 353, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.05]
State 354, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 355, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 356, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 357, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.05]
State 358, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.05]
State 359, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 360, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 361, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 362, Action P, Q-values: [+0.03 +0.08 -0.09 +0.08 +0.09 -0.05]
State 363, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 364, Action N, Q-values: [+0.04 +0.10 -0.07 +0.09 +0.08 -0.06]
State 365, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 366, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 367, Action W, Q-values: [+0.03 +0.08 -0.08 +0.09 +0.08 -0.06]
State 368, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.05]
State 369, Action N, Q-values: [+0.04 +0.09 -0.08 +0.07 +0.09 -0.05]
State 370, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.06]
State 371, Action P, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.04]
State 372, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 373, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 374, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 375, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 376, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 377, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 378, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.09 -0.05]
State 379, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.06]
State 380, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 381, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 382, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 383, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 384, Action N, Q-values: [+0.04 +0.08 -0.07 +0.08 +0.08 -0.06]
State 385, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.09 -0.05]
State 386, Action P, Q-values: [+0.04 +0.08 -0.07 +0.08 +0.09 -0.06]
State 387, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 388, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 389, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.06]
State 390, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.08 -0.06]
```

```
State 391, Action P, Q-values: [+0.04 +0.08 -0.09 +0.08 +0.09 -0.05]
State 392, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.07 -0.06]
State 393, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 394, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 395, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 396, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.06]
State 397, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 398, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 399, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 400, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 401, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 402, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 403, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 404, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 405, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 406, Action P, Q-values: [+0.04 +0.08 -0.08 +0.09 +0.09 -0.05]
State 407, Action W, Q-values: [+0.04 +0.08 -0.07 +0.09 +0.08 -0.06]
State 408, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.05]
State 409, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 410, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.06]
State 411, Action N, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.06]
State 412, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 413, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 414, Action N, Q-values: [+0.03 +0.10 -0.08 +0.09 +0.08 -0.06]
State 415, Action N, Q-values: [+0.03 +0.10 -0.09 +0.08 +0.08 -0.05]
State 416, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 417, Action N, Q-values: [+0.05 +0.09 -0.07 +0.08 +0.08 -0.06]
State 418, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.04]
State 419, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 420, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 421, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 422, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.05]
State 423, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 424, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.10 -0.05]
State 425, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 426, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 427, Action W, Q-values: [+0.03 +0.08 -0.07 +0.08 +0.08 -0.06]
State 428, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 429, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.06]
State 430, Action P, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 431, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.06]
State 432, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.09 -0.05]
State 433, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 434, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
State 435, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 436, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.05]
State 437, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.06]
State 438, Action W, Q-values: [+0.03 +0.08 -0.09 +0.09 +0.08 -0.05]
State 439, Action N, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 440, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 441, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.06]
State 442, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 443, Action P, Q-values: [+0.04 +0.08 -0.08 +0.09 +0.09 -0.05]
State 444, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 445, Action P, Q-values: [+0.04 +0.08 -0.09 +0.08 +0.09 -0.06]
State 446, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 447, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.05]
```

```
State 448, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.08 -0.05]
State 449, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 450, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 451, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 452, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.06]
State 453, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 454, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.07 -0.06]
State 455, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 456, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.08 -0.05]
State 457, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.09 -0.05]
State 458, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.06]
State 459, Action N, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.08 -0.05]
State 460, Action N, Q-values: [+0.04 +0.10 -0.07 +0.08 +0.08 -0.05]
State 461, Action P, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.06]
State 462, Action N, Q-values: [+0.04 +0.08 -0.08 +0.08 +0.08 -0.05]
State 463, Action N, Q-values: [+0.05 +0.09 -0.08 +0.08 +0.07 -0.05]
State 464, Action N, Q-values: [+0.03 +0.09 -0.07 +0.08 +0.09 -0.06]
State 465, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 466, Action N, Q-values: [+0.04 +0.09 -0.07 +0.09 +0.08 -0.06]
State 467, Action N, Q-values: [+0.03 +0.09 -0.09 +0.08 +0.09 -0.05]
State 468, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 469, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.09 -0.06]
State 470, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 471, Action N, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.08 -0.05]
State 472, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.05]
State 473, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 474, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 475, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 476, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 477, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.07 -0.06]
State 478, Action N, Q-values: [+0.03 +0.09 -0.08 +0.08 +0.09 -0.05]
State 479, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 480, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 481, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.05]
State 482, Action N, Q-values: [+0.03 +0.09 -0.08 +0.09 +0.08 -0.06]
State 483, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 484, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 485, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 486, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.09 -0.06]
State 487, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 488, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.06]
State 489, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 490, Action P, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.08 -0.06]
State 491, Action N, Q-values: [+0.04 +0.10 -0.08 +0.08 +0.07 -0.06]
State 492, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
State 493, Action P, Q-values: [+0.04 +0.09 -0.09 +0.08 +0.09 -0.05]
State 494, Action P, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 495, Action N, Q-values: [+0.04 +0.09 -0.07 +0.08 +0.08 -0.06]
State 496, Action N, Q-values: [+0.04 +0.09 -0.08 +0.09 +0.08 -0.05]
State 497, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.09 -0.05]
State 498, Action P, Q-values: [+0.03 +0.08 -0.08 +0.08 +0.09 -0.05]
State 499, Action N, Q-values: [+0.04 +0.09 -0.08 +0.08 +0.08 -0.05]
Episode 100/100000, Avg Reward: -783.15, Epsilon: 0.99
Episode 200/100000, Avg Reward: -769.30, Epsilon: 0.99
Episode 300/100000, Avg Reward: -772.94, Epsilon: 0.98
Episode 400/100000, Avg Reward: -739.93, Epsilon: 0.97
Episode 500/100000, Avg Reward: -727.96, Epsilon: 0.97
```

```
Episode 600/100000, Avg Reward: -749.61, Epsilon: 0.96
Episode 700/100000, Avg Reward: -733.56, Epsilon: 0.95
Episode 800/100000, Avg Reward: -740.32, Epsilon: 0.95
Episode 900/100000, Avg Reward: -739.75, Epsilon: 0.94
Episode 1000/100000, Avg Reward: -734.60, Epsilon: 0.93
Episode 1100/100000, Avg Reward: -713.57, Epsilon: 0.93
Episode 1200/100000, Avg Reward: -724.29, Epsilon: 0.92
Episode 1300/100000, Avg Reward: -700.74, Epsilon: 0.91
Episode 1400/100000, Avg Reward: -699.57, Epsilon: 0.91
Episode 1500/100000, Avg Reward: -714.60, Epsilon: 0.90
Episode 1600/100000, Avg Reward: -705.89, Epsilon: 0.90
Episode 1700/100000, Avg Reward: -685.57, Epsilon: 0.89
Episode 1800/100000, Avg Reward: -712.21, Epsilon: 0.88
Episode 1900/100000, Avg Reward: -686.91, Epsilon: 0.88
Episode 2000/100000, Avg Reward: -684.94, Epsilon: 0.87
Episode 2100/100000, Avg Reward: -681.78, Epsilon: 0.86
Episode 2200/100000, Avg Reward: -678.93, Epsilon: 0.86
Episode 2300/100000, Avg Reward: -675.60, Epsilon: 0.85
Episode 2400/100000, Avg Reward: -672.28, Epsilon: 0.85
Episode 2500/100000, Avg Reward: -677.92, Epsilon: 0.84
Episode 2600/100000, Avg Reward: -663.48, Epsilon: 0.84
Episode 2700/100000, Avg Reward: -664.63, Epsilon: 0.83
Episode 2800/100000, Avg Reward: -655.69, Epsilon: 0.82
Episode 2900/100000, Avg Reward: -654.29, Epsilon: 0.82
Episode 3000/100000, Avg Reward: -654.22, Epsilon: 0.81
Episode 3100/100000, Avg Reward: -644.39, Epsilon: 0.81
Episode 3200/100000, Avg Reward: -629.19, Epsilon: 0.80
Episode 3300/100000, Avg Reward: -655.14, Epsilon: 0.80
Episode 3400/100000, Avg Reward: -635.72, Epsilon: 0.79
Episode 3500/100000, Avg Reward: -640.50, Epsilon: 0.79
Episode 3600/100000, Avg Reward: -634.83, Epsilon: 0.78
Episode 3700/100000, Avg Reward: -626.74, Epsilon: 0.77
Episode 3800/100000, Avg Reward: -615.85, Epsilon: 0.77
Episode 3900/100000, Avg Reward: -617.86, Epsilon: 0.76
Episode 4000/100000, Avg Reward: -618.65, Epsilon: 0.76
Episode 4100/100000, Avg Reward: -608.12, Epsilon: 0.75
Episode 4200/100000, Avg Reward: -608.77, Epsilon: 0.75
Episode 4300/100000, Avg Reward: -605.91, Epsilon: 0.74
Episode 4400/100000, Avg Reward: -618.29, Epsilon: 0.74
Episode 4500/100000, Avg Reward: -607.49, Epsilon: 0.73
Episode 4600/100000, Avg Reward: -597.27, Epsilon: 0.73
Episode 4700/100000, Avg Reward: -599.52, Epsilon: 0.72
Episode 4800/100000, Avg Reward: -566.45, Epsilon: 0.72
Episode 4900/100000, Avg Reward: -581.66, Epsilon: 0.71
Episode 5000/100000, Avg Reward: -594.40, Epsilon: 0.71
Episode 5100/100000, Avg Reward: -582.87, Epsilon: 0.70
Episode 5200/100000, Avg Reward: -563.19, Epsilon: 0.70
Episode 5300/100000, Avg Reward: -578.49, Epsilon: 0.69
Episode 5400/100000, Avg Reward: -550.46, Epsilon: 0.69
Episode 5500/100000, Avg Reward: -558.83, Epsilon: 0.68
Episode 5600/100000, Avg Reward: -560.26, Epsilon: 0.68
Episode 5700/100000, Avg Reward: -545.78, Epsilon: 0.67
Episode 5800/100000, Avg Reward: -566.16, Epsilon: 0.67
Episode 5900/100000, Avg Reward: -547.38, Epsilon: 0.67
Episode 6000/100000, Avg Reward: -537.22, Epsilon: 0.66
Episode 6100/100000, Avg Reward: -526.08, Epsilon: 0.66
Episode 6200/100000, Avg Reward: -530.38, Epsilon: 0.65
```

```
Episode 6300/100000, Avg Reward: -504.97, Epsilon: 0.65
Episode 6400/100000, Avg Reward: -503.64, Epsilon: 0.64
Episode 6500/100000, Avg Reward: -470.47, Epsilon: 0.64
Episode 6600/100000, Avg Reward: -476.69, Epsilon: 0.63
Episode 6700/100000, Avg Reward: -489.39, Epsilon: 0.63
Episode 6800/100000, Avg Reward: -451.69, Epsilon: 0.63
Episode 6900/100000, Avg Reward: -450.12, Epsilon: 0.62
Episode 7000/100000, Avg Reward: -413.00, Epsilon: 0.62
Episode 7100/100000, Avg Reward: -387.06, Epsilon: 0.61
Episode 7200/100000, Avg Reward: -352.17, Epsilon: 0.61
Episode 7300/100000, Avg Reward: -342.69, Epsilon: 0.60
Episode 7400/100000, Avg Reward: -365.44, Epsilon: 0.60
Episode 7500/100000, Avg Reward: -328.69, Epsilon: 0.60
Episode 7600/100000, Avg Reward: -243.58, Epsilon: 0.59
Episode 7700/100000, Avg Reward: -217.15, Epsilon: 0.59
Episode 7800/100000, Avg Reward: -164.37, Epsilon: 0.58
Episode 7900/100000, Avg Reward: -152.63, Epsilon: 0.58
Episode 8000/100000, Avg Reward: -102.79, Epsilon: 0.58
Episode 8100/100000, Avg Reward: -70.08, Epsilon: 0.57
Episode 8200/100000, Avg Reward: -67.15, Epsilon: 0.57
Episode 8300/100000, Avg Reward: -80.93, Epsilon: 0.56
Episode 8400/100000, Avg Reward: -67.87, Epsilon: 0.56
Episode 8500/100000, Avg Reward: -61.28, Epsilon: 0.56
Episode 8600/100000, Avg Reward: -60.72, Epsilon: 0.55
Episode 8700/100000, Avg Reward: -56.27, Epsilon: 0.55
Episode 8800/100000, Avg Reward: -59.26, Epsilon: 0.54
Episode 8900/100000, Avg Reward: -62.11, Epsilon: 0.54
Episode 9000/100000, Avg Reward: -53.24, Epsilon: 0.54
Episode 9100/100000, Avg Reward: -52.20, Epsilon: 0.53
Episode 9200/100000, Avg Reward: -51.56, Epsilon: 0.53
Episode 9300/100000, Avg Reward: -49.89, Epsilon: 0.53
Episode 9400/100000, Avg Reward: -54.05, Epsilon: 0.52
Episode 9500/100000, Avg Reward: -54.13, Epsilon: 0.52
Episode 9600/100000, Avg Reward: -48.34, Epsilon: 0.52
Episode 9700/100000, Avg Reward: -53.00, Epsilon: 0.51
Episode 9800/100000, Avg Reward: -44.30, Epsilon: 0.51
Episode 9900/100000, Avg Reward: -45.73, Epsilon: 0.50
Episode 10000/100000, Avg Reward: -54.97, Epsilon: 0.50
Episode 10100/100000, Avg Reward: -50.38, Epsilon: 0.50
Episode 10200/100000, Avg Reward: -47.84, Epsilon: 0.49
Episode 10300/100000, Avg Reward: -48.60, Epsilon: 0.49
Episode 10400/100000, Avg Reward: -47.80, Epsilon: 0.49
Episode 10500/100000, Avg Reward: -42.56, Epsilon: 0.48
Episode 10600/100000, Avg Reward: -47.44, Epsilon: 0.48
Episode 10700/100000, Avg Reward: -47.77, Epsilon: 0.48
Episode 10800/100000, Avg Reward: -43.18, Epsilon: 0.47
Episode 10900/100000, Avg Reward: -41.63, Epsilon: 0.47
Episode 11000/100000, Avg Reward: -36.26, Epsilon: 0.47
Episode 11100/100000, Avg Reward: -36.70, Epsilon: 0.46
Episode 11200/100000, Avg Reward: -38.70, Epsilon: 0.46
Episode 11300/100000, Avg Reward: -35.24, Epsilon: 0.46
Episode 11400/100000, Avg Reward: -42.79, Epsilon: 0.45
Episode 11500/100000, Avg Reward: -40.24, Epsilon: 0.45
Episode 11600/100000, Avg Reward: -32.52, Epsilon: 0.45
Episode 11700/100000, Avg Reward: -35.77, Epsilon: 0.45
Episode 11800/100000, Avg Reward: -32.42, Epsilon: 0.44
Episode 11900/100000, Avg Reward: -29.23, Epsilon: 0.44
```

```
Episode 12000/100000, Avg Reward: -38.42, Epsilon: 0.44
Episode 12100/100000, Avg Reward: -30.10, Epsilon: 0.43
Episode 12200/100000, Avg Reward: -36.11, Epsilon: 0.43
Episode 12300/100000, Avg Reward: -28.22, Epsilon: 0.43
Episode 12400/100000, Avg Reward: -33.12, Epsilon: 0.42
Episode 12500/100000, Avg Reward: -35.76, Epsilon: 0.42
Episode 12600/100000, Avg Reward: -32.56, Epsilon: 0.42
Episode 12700/100000, Avg Reward: -32.33, Epsilon: 0.42
Episode 12800/100000, Avg Reward: -29.71, Epsilon: 0.41
Episode 12900/100000, Avg Reward: -29.51, Epsilon: 0.41
Episode 13000/100000, Avg Reward: -30.23, Epsilon: 0.41
Episode 13100/100000, Avg Reward: -30.64, Epsilon: 0.40
Episode 13200/100000, Avg Reward: -29.19, Epsilon: 0.40
Episode 13300/100000, Avg Reward: -28.37, Epsilon: 0.40
Episode 13400/100000, Avg Reward: -26.84, Epsilon: 0.40
Episode 13500/100000, Avg Reward: -29.16, Epsilon: 0.39
Episode 13600/100000, Avg Reward: -28.56, Epsilon: 0.39
Episode 13700/100000, Avg Reward: -23.38, Epsilon: 0.39
Episode 13800/100000, Avg Reward: -23.47, Epsilon: 0.39
Episode 13900/100000, Avg Reward: -24.27, Epsilon: 0.38
Episode 14000/100000, Avg Reward: -23.81, Epsilon: 0.38
Episode 14100/100000, Avg Reward: -21.59, Epsilon: 0.38
Episode 14200/100000, Avg Reward: -24.69, Epsilon: 0.37
Episode 14300/100000, Avg Reward: -24.67, Epsilon: 0.37
Episode 14400/100000, Avg Reward: -22.41, Epsilon: 0.37
Episode 14500/100000, Avg Reward: -23.96, Epsilon: 0.37
Episode 14600/100000, Avg Reward: -21.18, Epsilon: 0.36
Episode 14700/100000, Avg Reward: -22.14, Epsilon: 0.36
Episode 14800/100000, Avg Reward: -22.19, Epsilon: 0.36
Episode 14900/100000, Avg Reward: -20.64, Epsilon: 0.36
Episode 15000/100000, Avg Reward: -24.98, Epsilon: 0.35
Episode 15100/100000, Avg Reward: -19.25, Epsilon: 0.35
Episode 15200/100000, Avg Reward: -20.85, Epsilon: 0.35
Episode 15300/100000, Avg Reward: -19.24, Epsilon: 0.35
Episode 15400/100000, Avg Reward: -18.49, Epsilon: 0.35
Episode 15500/100000, Avg Reward: -22.54, Epsilon: 0.34
Episode 15600/100000, Avg Reward: -23.71, Epsilon: 0.34
Episode 15700/100000, Avg Reward: -17.66, Epsilon: 0.34
Episode 15800/100000, Avg Reward: -22.03, Epsilon: 0.34
Episode 15900/100000, Avg Reward: -20.39, Epsilon: 0.33
Episode 16000/100000, Avg Reward: -19.94, Epsilon: 0.33
Episode 16100/100000, Avg Reward: -20.74, Epsilon: 0.33
Episode 16200/100000, Avg Reward: -16.68, Epsilon: 0.33
Episode 16300/100000, Avg Reward: -16.20, Epsilon: 0.32
Episode 16400/100000, Avg Reward: -18.21, Epsilon: 0.32
Episode 16500/100000, Avg Reward: -15.92, Epsilon: 0.32
Episode 16600/100000, Avg Reward: -17.70, Epsilon: 0.32
Episode 16700/100000, Avg Reward: -18.93, Epsilon: 0.32
Episode 16800/100000, Avg Reward: -16.16, Epsilon: 0.31
Episode 16900/100000, Avg Reward: -15.38, Epsilon: 0.31
Episode 17000/100000, Avg Reward: -15.37, Epsilon: 0.31
Episode 17100/100000, Avg Reward: -14.19, Epsilon: 0.31
Episode 17200/100000, Avg Reward: -15.57, Epsilon: 0.30
Episode 17300/100000, Avg Reward: -14.88, Epsilon: 0.30
Episode 17400/100000, Avg Reward: -15.94, Epsilon: 0.30
Episode 17500/100000, Avg Reward: -13.69, Epsilon: 0.30
Episode 17600/100000, Avg Reward: -17.24, Epsilon: 0.30
```

```
Episode 17700/100000, Avg Reward: -14.87, Epsilon: 0.29
Episode 17800/100000, Avg Reward: -12.43, Epsilon: 0.29
Episode 17900/100000, Avg Reward: -10.16, Epsilon: 0.29
Episode 18000/100000, Avg Reward: -13.96, Epsilon: 0.29
Episode 18100/100000, Avg Reward: -15.19, Epsilon: 0.29
Episode 18200/100000, Avg Reward: -10.62, Epsilon: 0.28
Episode 18300/100000, Avg Reward: -14.13, Epsilon: 0.28
Episode 18400/100000, Avg Reward: -10.01, Epsilon: 0.28
Episode 18500/100000, Avg Reward: -12.07, Epsilon: 0.28
Episode 18600/100000, Avg Reward: -11.35, Epsilon: 0.28
Episode 18700/100000, Avg Reward: -11.53, Epsilon: 0.27
Episode 18800/100000, Avg Reward: -12.67, Epsilon: 0.27
Episode 18900/100000, Avg Reward: -12.15, Epsilon: 0.27
Episode 19000/100000, Avg Reward: -9.78, Epsilon: 0.27
Episode 19100/100000, Avg Reward: -11.89, Epsilon: 0.27
Episode 19200/100000, Avg Reward: -13.21, Epsilon: 0.27
Episode 19300/100000, Avg Reward: -9.54, Epsilon: 0.26
Episode 19400/100000, Avg Reward: -10.10, Epsilon: 0.26
Episode 19500/100000, Avg Reward: -8.49, Epsilon: 0.26
Episode 19600/100000, Avg Reward: -11.13, Epsilon: 0.26
Episode 19700/100000, Avg Reward: -10.32, Epsilon: 0.26
Episode 19800/100000, Avg Reward: -9.88, Epsilon: 0.25
Episode 19900/100000, Avg Reward: -8.95, Epsilon: 0.25
Episode 20000/100000, Avg Reward: -11.46, Epsilon: 0.25
Episode 20100/100000, Avg Reward: -9.39, Epsilon: 0.25
Episode 20200/100000, Avg Reward: -11.08, Epsilon: 0.25
Episode 20300/100000, Avg Reward: -8.80, Epsilon: 0.25
Episode 20400/100000, Avg Reward: -6.54, Epsilon: 0.24
Episode 20500/100000, Avg Reward: -10.54, Epsilon: 0.24
Episode 20600/100000, Avg Reward: -10.07, Epsilon: 0.24
Episode 20700/100000, Avg Reward: -8.06, Epsilon: 0.24
Episode 20800/100000, Avg Reward: -9.20, Epsilon: 0.24
Episode 20900/100000, Avg Reward: -7.74, Epsilon: 0.24
Episode 21000/100000, Avg Reward: -7.20, Epsilon: 0.23
Episode 21100/100000, Avg Reward: -8.08, Epsilon: 0.23
Episode 21200/100000, Avg Reward: -7.74, Epsilon: 0.23
Episode 21300/100000, Avg Reward: -8.12, Epsilon: 0.23
Episode 21400/100000, Avg Reward: -9.81, Epsilon: 0.23
Episode 21500/100000, Avg Reward: -6.67, Epsilon: 0.23
Episode 21600/100000, Avg Reward: -8.14, Epsilon: 0.22
Episode 21700/100000, Avg Reward: -7.58, Epsilon: 0.22
Episode 21800/100000, Avg Reward: -5.76, Epsilon: 0.22
Episode 21900/100000, Avg Reward: -8.69, Epsilon: 0.22
Episode 22000/100000, Avg Reward: -6.08, Epsilon: 0.22
Episode 22100/100000, Avg Reward: -6.85, Epsilon: 0.22
Episode 22200/100000, Avg Reward: -5.01, Epsilon: 0.22
Episode 22300/100000, Avg Reward: -5.99, Epsilon: 0.21
Episode 22400/100000, Avg Reward: -5.68, Epsilon: 0.21
Episode 22500/100000, Avg Reward: -5.23, Epsilon: 0.21
Episode 22600/100000, Avg Reward: -5.89, Epsilon: 0.21
Episode 22700/100000, Avg Reward: -4.80, Epsilon: 0.21
Episode 22800/100000, Avg Reward: -5.00, Epsilon: 0.21
Episode 22900/100000, Avg Reward: -7.24, Epsilon: 0.21
Episode 23000/100000, Avg Reward: -5.36, Epsilon: 0.20
Episode 23100/100000, Avg Reward: -6.93, Epsilon: 0.20
Episode 23200/100000, Avg Reward: -2.89, Epsilon: 0.20
Episode 23300/100000, Avg Reward: -3.54, Epsilon: 0.20
```

```
Episode 23400/100000, Avg Reward: -6.23, Epsilon: 0.20
Episode 23500/100000, Avg Reward: -3.98, Epsilon: 0.20
Episode 23600/100000, Avg Reward: -4.35, Epsilon: 0.20
Episode 23700/100000, Avg Reward: -3.90, Epsilon: 0.19
Episode 23800/100000, Avg Reward: -3.63, Epsilon: 0.19
Episode 23900/100000, Avg Reward: -4.72, Epsilon: 0.19
Episode 24000/100000, Avg Reward: -4.26, Epsilon: 0.19
Episode 24100/100000, Avg Reward: -3.96, Epsilon: 0.19
Episode 24200/100000, Avg Reward: -3.23, Epsilon: 0.19
Episode 24300/100000, Avg Reward: -4.96, Epsilon: 0.19
Episode 24400/100000, Avg Reward: -2.62, Epsilon: 0.19
Episode 24500/100000, Avg Reward: -6.74, Epsilon: 0.18
Episode 24600/100000, Avg Reward: -2.74, Epsilon: 0.18
Episode 24700/100000, Avg Reward: -4.34, Epsilon: 0.18
Episode 24800/100000, Avg Reward: -3.74, Epsilon: 0.18
Episode 24900/100000, Avg Reward: -2.53, Epsilon: 0.18
Episode 25000/100000, Avg Reward: -1.70, Epsilon: 0.18
Episode 25100/100000, Avg Reward: -2.32, Epsilon: 0.18
Episode 25200/100000, Avg Reward: -3.11, Epsilon: 0.18
Episode 25300/100000, Avg Reward: -1.28, Epsilon: 0.17
Episode 25400/100000, Avg Reward: -1.86, Epsilon: 0.17
Episode 25500/100000, Avg Reward: -3.51, Epsilon: 0.17
Episode 25600/100000, Avg Reward: -4.86, Epsilon: 0.17
Episode 25700/100000, Avg Reward: -5.31, Epsilon: 0.17
Episode 25800/100000, Avg Reward: -1.55, Epsilon: 0.17
Episode 25900/100000, Avg Reward: -2.74, Epsilon: 0.17
Episode 26000/100000, Avg Reward: -3.49, Epsilon: 0.17
Episode 26100/100000, Avg Reward: 0.06, Epsilon: 0.16
Episode 26200/100000, Avg Reward: -3.92, Epsilon: 0.16
Episode 26300/100000, Avg Reward: -3.33, Epsilon: 0.16
Episode 26400/100000, Avg Reward: -3.69, Epsilon: 0.16
Episode 26500/100000, Avg Reward: -5.53, Epsilon: 0.16
Episode 26600/100000, Avg Reward: -2.56, Epsilon: 0.16
Episode 26700/100000, Avg Reward: -1.12, Epsilon: 0.16
Episode 26800/100000, Avg Reward: -2.81, Epsilon: 0.16
Episode 26900/100000, Avg Reward: -0.59, Epsilon: 0.16
Episode 27000/100000, Avg Reward: -0.05, Epsilon: 0.15
Episode 27100/100000, Avg Reward: -0.27, Epsilon: 0.15
Episode 27200/100000, Avg Reward: 0.44, Epsilon: 0.15
Episode 27300/100000, Avg Reward: -2.10, Epsilon: 0.15
Episode 27400/100000, Avg Reward: -0.42, Epsilon: 0.15
Episode 27500/100000, Avg Reward: -0.50, Epsilon: 0.15
Episode 27600/100000, Avg Reward: -1.89, Epsilon: 0.15
Episode 27700/100000, Avg Reward: 0.09, Epsilon: 0.15
Episode 27800/100000, Avg Reward: -1.65, Epsilon: 0.15
Episode 27900/100000, Avg Reward: -4.99, Epsilon: 0.15
Episode 28000/100000, Avg Reward: -1.97, Epsilon: 0.14
Episode 28100/100000, Avg Reward: -2.22, Epsilon: 0.14
Episode 28200/100000, Avg Reward: 1.12, Epsilon: 0.14
Episode 28300/100000, Avg Reward: -3.02, Epsilon: 0.14
Episode 28400/100000, Avg Reward: -4.00, Epsilon: 0.14
Episode 28500/100000, Avg Reward: 0.81, Epsilon: 0.14
Episode 28600/100000, Avg Reward: -2.16, Epsilon: 0.14
Episode 28700/100000, Avg Reward: -2.33, Epsilon: 0.14
Episode 28800/100000, Avg Reward: -0.42, Epsilon: 0.14
Episode 28900/100000, Avg Reward: 0.28, Epsilon: 0.14
Episode 29000/100000, Avg Reward: -0.13, Epsilon: 0.13
```

```
Episode 29100/100000, Avg Reward: -0.19, Epsilon: 0.13
Episode 29200/100000, Avg Reward: 0.89, Epsilon: 0.13
Episode 29300/100000, Avg Reward: -0.91, Epsilon: 0.13
Episode 29400/100000, Avg Reward: -2.30, Epsilon: 0.13
Episode 29500/100000, Avg Reward: 0.58, Epsilon: 0.13
Episode 29600/100000, Avg Reward: 0.44, Epsilon: 0.13
Episode 29700/100000, Avg Reward: 1.62, Epsilon: 0.13
Episode 29800/100000, Avg Reward: 1.16, Epsilon: 0.13
Episode 29900/100000, Avg Reward: 0.96, Epsilon: 0.13
Episode 30000/100000, Avg Reward: 2.41, Epsilon: 0.13
Episode 30100/100000, Avg Reward: -0.29, Epsilon: 0.13
Episode 30200/100000, Avg Reward: 0.81, Epsilon: 0.12
Episode 30300/100000, Avg Reward: -3.30, Epsilon: 0.12
Episode 30400/100000, Avg Reward: 0.80, Epsilon: 0.12
Episode 30500/100000, Avg Reward: 0.36, Epsilon: 0.12
Episode 30600/100000, Avg Reward: 0.49, Epsilon: 0.12
Episode 30700/100000, Avg Reward: -0.41, Epsilon: 0.12
Episode 30800/100000, Avg Reward: 1.48, Epsilon: 0.12
Episode 30900/100000, Avg Reward: 0.38, Epsilon: 0.12
Episode 31000/100000, Avg Reward: 1.71, Epsilon: 0.12
Episode 31100/100000, Avg Reward: 0.21, Epsilon: 0.12
Episode 31200/100000, Avg Reward: 2.92, Epsilon: 0.12
Episode 31300/100000, Avg Reward: 1.92, Epsilon: 0.12
Episode 31400/100000, Avg Reward: 0.53, Epsilon: 0.11
Episode 31500/100000, Avg Reward: 1.94, Epsilon: 0.11
Episode 31600/100000, Avg Reward: 1.28, Epsilon: 0.11
Episode 31700/100000, Avg Reward: 2.43, Epsilon: 0.11
Episode 31800/100000, Avg Reward: 0.82, Epsilon: 0.11
Episode 31900/100000, Avg Reward: -0.23, Epsilon: 0.11
Episode 32000/100000, Avg Reward: 0.24, Epsilon: 0.11
Episode 32100/100000, Avg Reward: 2.22, Epsilon: 0.11
Episode 32200/100000, Avg Reward: 2.54, Epsilon: 0.11
Episode 32300/100000, Avg Reward: 0.81, Epsilon: 0.11
Episode 32400/100000, Avg Reward: 2.05, Epsilon: 0.11
Episode 32500/100000, Avg Reward: 1.74, Epsilon: 0.11
Episode 32600/100000, Avg Reward: 1.40, Epsilon: 0.11
Episode 32700/100000, Avg Reward: 1.54, Epsilon: 0.10
Episode 32800/100000, Avg Reward: 2.21, Epsilon: 0.10
Episode 32900/100000, Avg Reward: 0.75, Epsilon: 0.10
Episode 33000/100000, Avg Reward: 2.22, Epsilon: 0.10
Episode 33100/100000, Avg Reward: 1.80, Epsilon: 0.10
Episode 33200/100000, Avg Reward: 2.11, Epsilon: 0.10
Episode 33300/100000, Avg Reward: 2.09, Epsilon: 0.10
Episode 33400/100000, Avg Reward: 3.57, Epsilon: 0.10
Episode 33500/100000, Avg Reward: 1.20, Epsilon: 0.10
Episode 33600/100000, Avg Reward: 2.36, Epsilon: 0.10
Episode 33700/100000, Avg Reward: 2.32, Epsilon: 0.10
Episode 33800/100000, Avg Reward: 2.97, Epsilon: 0.10
Episode 33900/100000, Avg Reward: 1.79, Epsilon: 0.10
Episode 34000/100000, Avg Reward: 3.12, Epsilon: 0.10
Episode 34100/100000, Avg Reward: 3.05, Epsilon: 0.09
Episode 34200/100000, Avg Reward: 2.74, Epsilon: 0.09
Episode 34300/100000, Avg Reward: 1.59, Epsilon: 0.09
Episode 34400/100000, Avg Reward: 1.98, Epsilon: 0.09
Episode 34500/100000, Avg Reward: 2.72, Epsilon: 0.09
Episode 34600/100000, Avg Reward: 0.52, Epsilon: 0.09
Episode 34700/100000, Avg Reward: 4.31, Epsilon: 0.09
```

```
Episode 34800/100000, Avg Reward: 1.91, Epsilon: 0.09
Episode 34900/100000, Avg Reward: 3.14, Epsilon: 0.09
Episode 35000/100000, Avg Reward: 2.67, Epsilon: 0.09
Episode 35100/100000, Avg Reward: 3.34, Epsilon: 0.09
Episode 35200/100000, Avg Reward: 3.35, Epsilon: 0.09
Episode 35300/100000, Avg Reward: 3.75, Epsilon: 0.09
Episode 35400/100000, Avg Reward: 3.91, Epsilon: 0.09
Episode 35500/100000, Avg Reward: 3.41, Epsilon: 0.09
Episode 35600/100000, Avg Reward: 3.51, Epsilon: 0.09
Episode 35700/100000, Avg Reward: 2.11, Epsilon: 0.08
Episode 35800/100000, Avg Reward: 0.42, Epsilon: 0.08
Episode 35900/100000, Avg Reward: 3.55, Epsilon: 0.08
Episode 36000/100000, Avg Reward: 2.07, Epsilon: 0.08
Episode 36100/100000, Avg Reward: 2.47, Epsilon: 0.08
Episode 36200/100000, Avg Reward: 2.38, Epsilon: 0.08
Episode 36300/100000, Avg Reward: 3.22, Epsilon: 0.08
Episode 36400/100000, Avg Reward: 4.66, Epsilon: 0.08
Episode 36500/100000, Avg Reward: 3.92, Epsilon: 0.08
Episode 36600/100000, Avg Reward: 3.06, Epsilon: 0.08
Episode 36700/100000, Avg Reward: 3.91, Epsilon: 0.08
Episode 36800/100000, Avg Reward: 4.37, Epsilon: 0.08
Episode 36900/100000, Avg Reward: 4.23, Epsilon: 0.08
Episode 37000/100000, Avg Reward: 0.23, Epsilon: 0.08
Episode 37100/100000, Avg Reward: 2.70, Epsilon: 0.08
Episode 37200/100000, Avg Reward: 2.47, Epsilon: 0.08
Episode 37300/100000, Avg Reward: 2.71, Epsilon: 0.08
Episode 37400/100000, Avg Reward: 4.25, Epsilon: 0.08
Episode 37500/100000, Avg Reward: 3.79, Epsilon: 0.07
Episode 37600/100000, Avg Reward: 4.39, Epsilon: 0.07
Episode 37700/100000, Avg Reward: 2.85, Epsilon: 0.07
Episode 37800/100000, Avg Reward: 3.91, Epsilon: 0.07
Episode 37900/100000, Avg Reward: 4.06, Epsilon: 0.07
Episode 38000/100000, Avg Reward: 3.53, Epsilon: 0.07
Episode 38100/100000, Avg Reward: 3.24, Epsilon: 0.07
Episode 38200/100000, Avg Reward: 5.48, Epsilon: 0.07
Episode 38300/100000, Avg Reward: 4.25, Epsilon: 0.07
Episode 38400/100000, Avg Reward: 4.42, Epsilon: 0.07
Episode 38500/100000, Avg Reward: 3.18, Epsilon: 0.07
Episode 38600/100000, Avg Reward: 4.07, Epsilon: 0.07
Episode 38700/100000, Avg Reward: 3.99, Epsilon: 0.07
Episode 38800/100000, Avg Reward: 3.67, Epsilon: 0.07
Episode 38900/100000, Avg Reward: 4.71, Epsilon: 0.07
Episode 39000/100000, Avg Reward: 3.82, Epsilon: 0.07
Episode 39100/100000, Avg Reward: 4.35, Epsilon: 0.07
Episode 39200/100000, Avg Reward: 3.68, Epsilon: 0.07
Episode 39300/100000, Avg Reward: 5.35, Epsilon: 0.07
Episode 39400/100000, Avg Reward: 3.82, Epsilon: 0.07
Episode 39500/100000, Avg Reward: 2.29, Epsilon: 0.07
Episode 39600/100000, Avg Reward: 5.61, Epsilon: 0.06
Episode 39700/100000, Avg Reward: 3.79, Epsilon: 0.06
Episode 39800/100000, Avg Reward: 5.15, Epsilon: 0.06
Episode 39900/100000, Avg Reward: 4.30, Epsilon: 0.06
Episode 40000/100000, Avg Reward: 3.25, Epsilon: 0.06
Episode 40100/100000, Avg Reward: 0.81, Epsilon: 0.06
Episode 40200/100000, Avg Reward: 4.49, Epsilon: 0.06
Episode 40300/100000, Avg Reward: 2.71, Epsilon: 0.06
Episode 40400/100000, Avg Reward: 5.09, Epsilon: 0.06
```

```
Episode 40500/100000, Avg Reward: 4.24, Epsilon: 0.06
Episode 40600/100000, Avg Reward: 5.08, Epsilon: 0.06
Episode 40700/100000, Avg Reward: 5.61, Epsilon: 0.06
Episode 40800/100000, Avg Reward: 4.60, Epsilon: 0.06
Episode 40900/100000, Avg Reward: 5.71, Epsilon: 0.06
Episode 41000/100000, Avg Reward: 4.57, Epsilon: 0.06
Episode 41100/100000, Avg Reward: 5.02, Epsilon: 0.06
Episode 41200/100000, Avg Reward: 5.07, Epsilon: 0.06
Episode 41300/100000, Avg Reward: 5.00, Epsilon: 0.06
Episode 41400/100000, Avg Reward: 5.21, Epsilon: 0.06
Episode 41500/100000, Avg Reward: 5.40, Epsilon: 0.06
Episode 41600/100000, Avg Reward: 5.57, Epsilon: 0.06
Episode 41700/100000, Avg Reward: 4.46, Epsilon: 0.06
Episode 41800/100000, Avg Reward: 4.66, Epsilon: 0.06
Episode 41900/100000, Avg Reward: 4.78, Epsilon: 0.06
Episode 42000/100000, Avg Reward: 4.66, Epsilon: 0.05
Episode 42100/100000, Avg Reward: 2.93, Epsilon: 0.05
Episode 42200/100000, Avg Reward: 6.30, Epsilon: 0.05
Episode 42300/100000, Avg Reward: 5.93, Epsilon: 0.05
Episode 42400/100000, Avg Reward: 4.57, Epsilon: 0.05
Episode 42500/100000, Avg Reward: 4.53, Epsilon: 0.05
Episode 42600/100000, Avg Reward: 5.48, Epsilon: 0.05
Episode 42700/100000, Avg Reward: 1.82, Epsilon: 0.05
Episode 42800/100000, Avg Reward: 2.76, Epsilon: 0.05
Episode 42900/100000, Avg Reward: 5.22, Epsilon: 0.05
Episode 43000/100000, Avg Reward: 5.50, Epsilon: 0.05
Episode 43100/100000, Avg Reward: 5.38, Epsilon: 0.05
Episode 43200/100000, Avg Reward: 5.13, Epsilon: 0.05
Episode 43300/100000, Avg Reward: 5.08, Epsilon: 0.05
Episode 43400/100000, Avg Reward: 5.72, Epsilon: 0.05
Episode 43500/100000, Avg Reward: 5.32, Epsilon: 0.05
Episode 43600/100000, Avg Reward: 4.31, Epsilon: 0.05
Episode 43700/100000, Avg Reward: 5.03, Epsilon: 0.05
Episode 43800/100000, Avg Reward: 3.72, Epsilon: 0.05
Episode 43900/100000, Avg Reward: 4.49, Epsilon: 0.05
Episode 44000/100000, Avg Reward: 6.13, Epsilon: 0.05
Episode 44100/100000, Avg Reward: 5.51, Epsilon: 0.05
Episode 44200/100000, Avg Reward: 5.08, Epsilon: 0.05
Episode 44300/100000, Avg Reward: 5.97, Epsilon: 0.05
Episode 44400/100000, Avg Reward: 5.82, Epsilon: 0.05
Episode 44500/100000, Avg Reward: 4.22, Epsilon: 0.05
Episode 44600/100000, Avg Reward: 5.60, Epsilon: 0.05
Episode 44700/100000, Avg Reward: 5.19, Epsilon: 0.05
Episode 44800/100000, Avg Reward: 5.71, Epsilon: 0.05
Episode 44900/100000, Avg Reward: 3.62, Epsilon: 0.04
Episode 45000/100000, Avg Reward: 5.86, Epsilon: 0.04
Episode 45100/100000, Avg Reward: 4.90, Epsilon: 0.04
Episode 45200/100000, Avg Reward: 4.67, Epsilon: 0.04
Episode 45300/100000, Avg Reward: 6.13, Epsilon: 0.04
Episode 45400/100000, Avg Reward: 5.69, Epsilon: 0.04
Episode 45500/100000, Avg Reward: 5.19, Epsilon: 0.04
Episode 45600/100000, Avg Reward: 5.80, Epsilon: 0.04
Episode 45700/100000, Avg Reward: 5.94, Epsilon: 0.04
Episode 45800/100000, Avg Reward: 4.47, Epsilon: 0.04
Episode 45900/100000, Avg Reward: 1.62, Epsilon: 0.04
Episode 46000/100000, Avg Reward: 6.18, Epsilon: 0.04
Episode 46100/100000, Avg Reward: 6.82, Epsilon: 0.04
```

```
Episode 46200/100000, Avg Reward: 6.20, Epsilon: 0.04
Episode 46300/100000, Avg Reward: 6.21, Epsilon: 0.04
Episode 46400/100000, Avg Reward: 5.81, Epsilon: 0.04
Episode 46500/100000, Avg Reward: 6.02, Epsilon: 0.04
Episode 46600/100000, Avg Reward: 5.73, Epsilon: 0.04
Episode 46700/100000, Avg Reward: 6.83, Epsilon: 0.04
Episode 46800/100000, Avg Reward: 5.44, Epsilon: 0.04
Episode 46900/100000, Avg Reward: 6.48, Epsilon: 0.04
Episode 47000/100000, Avg Reward: 6.00, Epsilon: 0.04
Episode 47100/100000, Avg Reward: 5.66, Epsilon: 0.04
Episode 47200/100000, Avg Reward: 6.19, Epsilon: 0.04
Episode 47300/100000, Avg Reward: 5.86, Epsilon: 0.04
Episode 47400/100000, Avg Reward: 4.92, Epsilon: 0.04
Episode 47500/100000, Avg Reward: 6.15, Epsilon: 0.04
Episode 47600/100000, Avg Reward: 6.18, Epsilon: 0.04
Episode 47700/100000, Avg Reward: 6.47, Epsilon: 0.04
Episode 47800/100000, Avg Reward: 5.15, Epsilon: 0.04
Episode 47900/100000, Avg Reward: 4.62, Epsilon: 0.04
Episode 48000/100000, Avg Reward: 3.02, Epsilon: 0.04
Episode 48100/100000, Avg Reward: 4.25, Epsilon: 0.04
Episode 48200/100000, Avg Reward: 6.55, Epsilon: 0.04
Episode 48300/100000, Avg Reward: 6.48, Epsilon: 0.04
Episode 48400/100000, Avg Reward: 5.85, Epsilon: 0.04
Episode 48500/100000, Avg Reward: 5.97, Epsilon: 0.04
Episode 48600/100000, Avg Reward: 6.15, Epsilon: 0.03
Episode 48700/100000, Avg Reward: 6.63, Epsilon: 0.03
Episode 48800/100000, Avg Reward: 3.71, Epsilon: 0.03
Episode 48900/100000, Avg Reward: 5.92, Epsilon: 0.03
Episode 49000/100000, Avg Reward: 5.69, Epsilon: 0.03
Episode 49100/100000, Avg Reward: 4.17, Epsilon: 0.03
Episode 49200/100000, Avg Reward: 6.15, Epsilon: 0.03
Episode 49300/100000, Avg Reward: 6.76, Epsilon: 0.03
Episode 49400/100000, Avg Reward: 6.65, Epsilon: 0.03
Episode 49500/100000, Avg Reward: 6.52, Epsilon: 0.03
Episode 49600/100000, Avg Reward: 4.25, Epsilon: 0.03
Episode 49700/100000, Avg Reward: 4.50, Epsilon: 0.03
Episode 49800/100000, Avg Reward: 6.67, Epsilon: 0.03
Episode 49900/100000, Avg Reward: 6.00, Epsilon: 0.03
Episode 50000/100000, Avg Reward: 7.13, Epsilon: 0.03
Episode 50100/100000, Avg Reward: 5.53, Epsilon: 0.03
Episode 50200/100000, Avg Reward: 3.64, Epsilon: 0.03
Episode 50300/100000, Avg Reward: 6.01, Epsilon: 0.03
Episode 50400/100000, Avg Reward: 5.38, Epsilon: 0.03
Episode 50500/100000, Avg Reward: 2.74, Epsilon: 0.03
Episode 50600/100000, Avg Reward: 5.69, Epsilon: 0.03
Episode 50700/100000, Avg Reward: 6.43, Epsilon: 0.03
Episode 50800/100000, Avg Reward: 7.36, Epsilon: 0.03
Episode 50900/100000, Avg Reward: 6.13, Epsilon: 0.03
Episode 51000/100000, Avg Reward: 4.48, Epsilon: 0.03
Episode 51100/100000, Avg Reward: 4.79, Epsilon: 0.03
Episode 51200/100000, Avg Reward: 7.30, Epsilon: 0.03
Episode 51300/100000, Avg Reward: 6.81, Epsilon: 0.03
Episode 51400/100000, Avg Reward: 6.17, Epsilon: 0.03
Episode 51500/100000, Avg Reward: 4.46, Epsilon: 0.03
Episode 51600/100000, Avg Reward: 6.73, Epsilon: 0.03
Episode 51700/100000, Avg Reward: 6.14, Epsilon: 0.03
Episode 51800/100000, Avg Reward: 6.72, Epsilon: 0.03
```

```
Episode 51900/100000, Avg Reward: 7.01, Epsilon: 0.03
Episode 52000/100000, Avg Reward: 6.28, Epsilon: 0.03
Episode 52100/100000, Avg Reward: 6.83, Epsilon: 0.03
Episode 52200/100000, Avg Reward: 6.70, Epsilon: 0.03
Episode 52300/100000, Avg Reward: 6.14, Epsilon: 0.03
Episode 52400/100000, Avg Reward: 5.93, Epsilon: 0.03
Episode 52500/100000, Avg Reward: 6.40, Epsilon: 0.03
Episode 52600/100000, Avg Reward: 6.60, Epsilon: 0.03
Episode 52700/100000, Avg Reward: 4.41, Epsilon: 0.03
Episode 52800/100000, Avg Reward: 6.31, Epsilon: 0.03
Episode 52900/100000, Avg Reward: 1.68, Epsilon: 0.03
Episode 53000/100000, Avg Reward: 5.05, Epsilon: 0.03
Episode 53100/100000, Avg Reward: 5.76, Epsilon: 0.03
Episode 53200/100000, Avg Reward: 6.37, Epsilon: 0.03
Episode 53300/100000, Avg Reward: 7.05, Epsilon: 0.03
Episode 53400/100000, Avg Reward: 6.84, Epsilon: 0.03
Episode 53500/100000, Avg Reward: 6.28, Epsilon: 0.02
Episode 53600/100000, Avg Reward: 6.79, Epsilon: 0.02
Episode 53700/100000, Avg Reward: 6.93, Epsilon: 0.02
Episode 53800/100000, Avg Reward: 6.68, Epsilon: 0.02
Episode 53900/100000, Avg Reward: 6.41, Epsilon: 0.02
Episode 54000/100000, Avg Reward: 6.31, Epsilon: 0.02
Episode 54100/100000, Avg Reward: 4.41, Epsilon: 0.02
Episode 54200/100000, Avg Reward: 5.94, Epsilon: 0.02
Episode 54300/100000, Avg Reward: 6.64, Epsilon: 0.02
Episode 54400/100000, Avg Reward: 7.06, Epsilon: 0.02
Episode 54500/100000, Avg Reward: 7.14, Epsilon: 0.02
Episode 54600/100000, Avg Reward: 7.50, Epsilon: 0.02
Episode 54700/100000, Avg Reward: 7.18, Epsilon: 0.02
Episode 54800/100000, Avg Reward: 4.65, Epsilon: 0.02
Episode 54900/100000, Avg Reward: 6.73, Epsilon: 0.02
Episode 55000/100000, Avg Reward: 6.29, Epsilon: 0.02
Episode 55100/100000, Avg Reward: -1.17, Epsilon: 0.02
Episode 55200/100000, Avg Reward: 5.83, Epsilon: 0.02
Episode 55300/100000, Avg Reward: 4.21, Epsilon: 0.02
Episode 55400/100000, Avg Reward: 6.75, Epsilon: 0.02
Episode 55500/100000, Avg Reward: 7.43, Epsilon: 0.02
Episode 55600/100000, Avg Reward: 6.51, Epsilon: 0.02
Episode 55700/100000, Avg Reward: 6.47, Epsilon: 0.02
Episode 55800/100000, Avg Reward: 6.54, Epsilon: 0.02
Episode 55900/100000, Avg Reward: 7.15, Epsilon: 0.02
Episode 56000/100000, Avg Reward: 6.56, Epsilon: 0.02
Episode 56100/100000, Avg Reward: 6.53, Epsilon: 0.02
Episode 56200/100000, Avg Reward: 7.12, Epsilon: 0.02
Episode 56300/100000, Avg Reward: 6.31, Epsilon: 0.02
Episode 56400/100000, Avg Reward: 6.99, Epsilon: 0.02
Episode 56500/100000, Avg Reward: 5.70, Epsilon: 0.02
Episode 56600/100000, Avg Reward: 7.22, Epsilon: 0.02
Episode 56700/100000, Avg Reward: 4.29, Epsilon: 0.02
Episode 56800/100000, Avg Reward: 7.26, Epsilon: 0.02
Episode 56900/100000, Avg Reward: 6.73, Epsilon: 0.02
Episode 57000/100000, Avg Reward: 7.01, Epsilon: 0.02
Episode 57100/100000, Avg Reward: 7.01, Epsilon: 0.02
Episode 57200/100000, Avg Reward: 7.43, Epsilon: 0.02
Episode 57300/100000, Avg Reward: 7.38, Epsilon: 0.02
Episode 57400/100000, Avg Reward: 6.96, Epsilon: 0.02
Episode 57500/100000, Avg Reward: 6.16, Epsilon: 0.02
```

```
Episode 57600/100000, Avg Reward: 6.92, Epsilon: 0.02
Episode 57700/100000, Avg Reward: 6.86, Epsilon: 0.02
Episode 57800/100000, Avg Reward: 3.97, Epsilon: 0.02
Episode 57900/100000, Avg Reward: -2.20, Epsilon: 0.02
Episode 58000/100000, Avg Reward: 4.99, Epsilon: 0.02
Episode 58100/100000, Avg Reward: 6.17, Epsilon: 0.02
Episode 58200/100000, Avg Reward: 6.59, Epsilon: 0.02
Episode 58300/100000, Avg Reward: 7.06, Epsilon: 0.02
Episode 58400/100000, Avg Reward: 4.87, Epsilon: 0.02
Episode 58500/100000, Avg Reward: 6.71, Epsilon: 0.02
Episode 58600/100000, Avg Reward: 6.94, Epsilon: 0.02
Episode 58700/100000, Avg Reward: 6.39, Epsilon: 0.02
Episode 58800/100000, Avg Reward: 7.09, Epsilon: 0.02
Episode 58900/100000, Avg Reward: 4.80, Epsilon: 0.02
Episode 59000/100000, Avg Reward: 7.05, Epsilon: 0.02
Episode 59100/100000, Avg Reward: 7.55, Epsilon: 0.02
Episode 59200/100000, Avg Reward: 4.58, Epsilon: 0.02
Episode 59300/100000, Avg Reward: 7.15, Epsilon: 0.02
Episode 59400/100000, Avg Reward: 7.26, Epsilon: 0.02
Episode 59500/100000, Avg Reward: 7.04, Epsilon: 0.02
Episode 59600/100000, Avg Reward: 3.06, Epsilon: 0.02
Episode 59700/100000, Avg Reward: 5.16, Epsilon: 0.02
Episode 59800/100000, Avg Reward: 7.08, Epsilon: 0.02
Episode 59900/100000, Avg Reward: 7.09, Epsilon: 0.02
Episode 60000/100000, Avg Reward: 7.61, Epsilon: 0.02
Episode 60100/100000, Avg Reward: 7.44, Epsilon: 0.02
Episode 60200/100000, Avg Reward: 7.47, Epsilon: 0.02
Episode 60300/100000, Avg Reward: 6.77, Epsilon: 0.02
Episode 60400/100000, Avg Reward: 6.92, Epsilon: 0.02
Episode 60500/100000, Avg Reward: 4.88, Epsilon: 0.02
Episode 60600/100000, Avg Reward: 3.80, Epsilon: 0.02
Episode 60700/100000, Avg Reward: -0.18, Epsilon: 0.02
Episode 60800/100000, Avg Reward: 3.32, Epsilon: 0.01
Episode 60900/100000, Avg Reward: 5.61, Epsilon: 0.01
Episode 61000/100000, Avg Reward: 7.53, Epsilon: 0.01
Episode 61100/100000, Avg Reward: 7.22, Epsilon: 0.01
Episode 61200/100000, Avg Reward: 4.51, Epsilon: 0.01
Episode 61300/100000, Avg Reward: 7.34, Epsilon: 0.01
Episode 61400/100000, Avg Reward: 7.00, Epsilon: 0.01
Episode 61500/100000, Avg Reward: 6.97, Epsilon: 0.01
Episode 61600/100000, Avg Reward: 7.10, Epsilon: 0.01
Episode 61700/100000, Avg Reward: 6.40, Epsilon: 0.01
Episode 61800/100000, Avg Reward: 5.30, Epsilon: 0.01
Episode 61900/100000, Avg Reward: 7.81, Epsilon: 0.01
Episode 62000/100000, Avg Reward: 7.05, Epsilon: 0.01
Episode 62100/100000, Avg Reward: 4.94, Epsilon: 0.01
Episode 62200/100000, Avg Reward: 6.85, Epsilon: 0.01
Episode 62300/100000, Avg Reward: 7.34, Epsilon: 0.01
Episode 62400/100000, Avg Reward: 7.75, Epsilon: 0.01
Episode 62500/100000, Avg Reward: 7.12, Epsilon: 0.01
Episode 62600/100000, Avg Reward: 7.39, Epsilon: 0.01
Episode 62700/100000, Avg Reward: 7.82, Epsilon: 0.01
Episode 62800/100000, Avg Reward: 7.36, Epsilon: 0.01
Episode 62900/100000, Avg Reward: 7.75, Epsilon: 0.01
Episode 63000/100000, Avg Reward: 6.97, Epsilon: 0.01
Episode 63100/100000, Avg Reward: 7.17, Epsilon: 0.01
Episode 63200/100000, Avg Reward: 6.98, Epsilon: 0.01
```

```
Episode 63300/100000, Avg Reward: 7.47, Epsilon: 0.01
Episode 63400/100000, Avg Reward: 4.87, Epsilon: 0.01
Episode 63500/100000, Avg Reward: 6.93, Epsilon: 0.01
Episode 63600/100000, Avg Reward: 7.11, Epsilon: 0.01
Episode 63700/100000, Avg Reward: 7.14, Epsilon: 0.01
Episode 63800/100000, Avg Reward: 7.05, Epsilon: 0.01
Episode 63900/100000, Avg Reward: 5.49, Epsilon: 0.01
Episode 64000/100000, Avg Reward: 7.19, Epsilon: 0.01
Episode 64100/100000, Avg Reward: 7.67, Epsilon: 0.01
Episode 64200/100000, Avg Reward: 5.13, Epsilon: 0.01
Episode 64300/100000, Avg Reward: 7.57, Epsilon: 0.01
Episode 64400/100000, Avg Reward: 6.72, Epsilon: 0.01
Episode 64500/100000, Avg Reward: 7.20, Epsilon: 0.01
Episode 64600/100000, Avg Reward: 7.86, Epsilon: 0.01
Episode 64700/100000, Avg Reward: 7.75, Epsilon: 0.01
Episode 64800/100000, Avg Reward: 7.50, Epsilon: 0.01
Episode 64900/100000, Avg Reward: 7.18, Epsilon: 0.01
Episode 65000/100000, Avg Reward: 6.73, Epsilon: 0.01
Episode 65100/100000, Avg Reward: 7.15, Epsilon: 0.01
Episode 65200/100000, Avg Reward: 3.23, Epsilon: 0.01
Episode 65300/100000, Avg Reward: 4.04, Epsilon: 0.01
Episode 65400/100000, Avg Reward: 7.44, Epsilon: 0.01
Episode 65500/100000, Avg Reward: 7.98, Epsilon: 0.01
Episode 65600/100000, Avg Reward: 7.71, Epsilon: 0.01
Episode 65700/100000, Avg Reward: 7.26, Epsilon: 0.01
Episode 65800/100000, Avg Reward: 7.58, Epsilon: 0.01
Episode 65900/100000, Avg Reward: 7.62, Epsilon: 0.01
Episode 66000/100000, Avg Reward: 5.53, Epsilon: 0.01
Episode 66100/100000, Avg Reward: 5.77, Epsilon: 0.01
Episode 66200/100000, Avg Reward: 6.82, Epsilon: 0.01
Episode 66300/100000, Avg Reward: 7.50, Epsilon: 0.01
Episode 66400/100000, Avg Reward: 7.61, Epsilon: 0.01
Episode 66500/100000, Avg Reward: 7.33, Epsilon: 0.01
Episode 66600/100000, Avg Reward: 7.28, Epsilon: 0.01
Episode 66700/100000, Avg Reward: 7.73, Epsilon: 0.01
Episode 66800/100000, Avg Reward: 7.77, Epsilon: 0.01
Episode 66900/100000, Avg Reward: 7.40, Epsilon: 0.01
Episode 67000/100000, Avg Reward: 7.78, Epsilon: 0.01
Episode 67100/100000, Avg Reward: 7.33, Epsilon: 0.01
Episode 67200/100000, Avg Reward: 7.28, Epsilon: 0.01
Episode 67300/100000, Avg Reward: 5.09, Epsilon: 0.01
Episode 67400/100000, Avg Reward: 7.88, Epsilon: 0.01
Episode 67500/100000, Avg Reward: 7.33, Epsilon: 0.01
Episode 67600/100000, Avg Reward: 5.92, Epsilon: 0.01
Episode 67700/100000, Avg Reward: 4.93, Epsilon: 0.01
Episode 67800/100000, Avg Reward: 7.41, Epsilon: 0.01
Episode 67900/100000, Avg Reward: 7.29, Epsilon: 0.01
Episode 68000/100000, Avg Reward: 7.92, Epsilon: 0.01
Episode 68100/100000, Avg Reward: 7.81, Epsilon: 0.01
Episode 68200/100000, Avg Reward: 8.13, Epsilon: 0.01
Episode 68300/100000, Avg Reward: 4.88, Epsilon: 0.01
Episode 68400/100000, Avg Reward: 1.24, Epsilon: 0.01
Episode 68500/100000, Avg Reward: 7.37, Epsilon: 0.01
Episode 68600/100000, Avg Reward: 7.21, Epsilon: 0.01
Episode 68700/100000, Avg Reward: 7.86, Epsilon: 0.01
Episode 68800/100000, Avg Reward: 7.36, Epsilon: 0.01
Episode 68900/100000, Avg Reward: 7.13, Epsilon: 0.01
```

```
Episode 69000/100000, Avg Reward: 7.90, Epsilon: 0.01
Episode 69100/100000, Avg Reward: 7.99, Epsilon: 0.01
Episode 69200/100000, Avg Reward: 5.47, Epsilon: 0.01
Episode 69300/100000, Avg Reward: 8.20, Epsilon: 0.01
Episode 69400/100000, Avg Reward: 7.18, Epsilon: 0.01
Episode 69500/100000, Avg Reward: 7.01, Epsilon: 0.01
Episode 69600/100000, Avg Reward: 7.69, Epsilon: 0.01
Episode 69700/100000, Avg Reward: 5.59, Epsilon: 0.01
Episode 69800/100000, Avg Reward: 7.65, Epsilon: 0.01
Episode 69900/100000, Avg Reward: 7.24, Epsilon: 0.01
Episode 70000/100000, Avg Reward: 8.00, Epsilon: 0.01
Episode 70100/100000, Avg Reward: 7.62, Epsilon: 0.01
Episode 70200/100000, Avg Reward: 2.28, Epsilon: 0.01
Episode 70300/100000, Avg Reward: 6.24, Epsilon: 0.01
Episode 70400/100000, Avg Reward: 6.46, Epsilon: 0.01
Episode 70500/100000, Avg Reward: 8.07, Epsilon: 0.01
Episode 70600/100000, Avg Reward: 3.48, Epsilon: 0.01
Episode 70700/100000, Avg Reward: 7.39, Epsilon: 0.01
Episode 70800/100000, Avg Reward: 7.09, Epsilon: 0.01
Episode 70900/100000, Avg Reward: 7.87, Epsilon: 0.01
Episode 71000/100000, Avg Reward: 7.24, Epsilon: 0.01
Episode 71100/100000, Avg Reward: 7.54, Epsilon: 0.01
Episode 71200/100000, Avg Reward: 7.96, Epsilon: 0.01
Episode 71300/100000, Avg Reward: 5.47, Epsilon: 0.01
Episode 71400/100000, Avg Reward: 8.01, Epsilon: 0.01
Episode 71500/100000, Avg Reward: 7.54, Epsilon: 0.01
Episode 71600/100000, Avg Reward: 7.69, Epsilon: 0.01
Episode 71700/100000, Avg Reward: 7.71, Epsilon: 0.01
Episode 71800/100000, Avg Reward: 6.70, Epsilon: 0.01
Episode 71900/100000, Avg Reward: 7.70, Epsilon: 0.01
Episode 72000/100000, Avg Reward: 7.74, Epsilon: 0.01
Episode 72100/100000, Avg Reward: 3.76, Epsilon: 0.01
Episode 72200/100000, Avg Reward: 7.12, Epsilon: 0.01
Episode 72300/100000, Avg Reward: 7.32, Epsilon: 0.01
Episode 72400/100000, Avg Reward: 6.53, Epsilon: 0.01
Episode 72500/100000, Avg Reward: 7.70, Epsilon: 0.01
Episode 72600/100000, Avg Reward: 7.43, Epsilon: 0.01
Episode 72700/100000, Avg Reward: 7.56, Epsilon: 0.01
Episode 72800/100000, Avg Reward: 5.83, Epsilon: 0.01
Episode 72900/100000, Avg Reward: 2.95, Epsilon: 0.01
Episode 73000/100000, Avg Reward: 7.71, Epsilon: 0.01
Episode 73100/100000, Avg Reward: 7.39, Epsilon: 0.01
Episode 73200/100000, Avg Reward: 7.61, Epsilon: 0.01
Episode 73300/100000, Avg Reward: 7.24, Epsilon: 0.01
Episode 73400/100000, Avg Reward: 7.71, Epsilon: 0.01
Episode 73500/100000, Avg Reward: 7.43, Epsilon: 0.01
Episode 73600/100000, Avg Reward: 1.58, Epsilon: 0.01
Episode 73700/100000, Avg Reward: 8.21, Epsilon: 0.01
Episode 73800/100000, Avg Reward: 5.62, Epsilon: 0.01
Episode 73900/100000, Avg Reward: 5.72, Epsilon: 0.01
Episode 74000/100000, Avg Reward: 7.80, Epsilon: 0.01
Episode 74100/100000, Avg Reward: 7.53, Epsilon: 0.01
Episode 74200/100000, Avg Reward: 7.66, Epsilon: 0.01
Episode 74300/100000, Avg Reward: 7.56, Epsilon: 0.01
Episode 74400/100000, Avg Reward: 6.70, Epsilon: 0.01
Episode 74500/100000, Avg Reward: 7.53, Epsilon: 0.01
Episode 74600/100000, Avg Reward: 1.45, Epsilon: 0.01
```

```
Episode 74700/100000, Avg Reward: 7.68, Epsilon: 0.01
Episode 74800/100000, Avg Reward: 7.76, Epsilon: 0.01
Episode 74900/100000, Avg Reward: 7.50, Epsilon: 0.01
Episode 75000/100000, Avg Reward: 7.14, Epsilon: 0.01
Episode 75100/100000, Avg Reward: 7.60, Epsilon: 0.01
Episode 75200/100000, Avg Reward: 7.75, Epsilon: 0.01
Episode 75300/100000, Avg Reward: 8.21, Epsilon: 0.01
Episode 75400/100000, Avg Reward: 7.76, Epsilon: 0.01
Episode 75500/100000, Avg Reward: 7.73, Epsilon: 0.01
Episode 75600/100000, Avg Reward: 7.92, Epsilon: 0.01
Episode 75700/100000, Avg Reward: 7.90, Epsilon: 0.01
Episode 75800/100000, Avg Reward: 7.49, Epsilon: 0.01
Episode 75900/100000, Avg Reward: 7.77, Epsilon: 0.01
Episode 76000/100000, Avg Reward: 1.60, Epsilon: 0.01
Episode 76100/100000, Avg Reward: 7.75, Epsilon: 0.01
Episode 76200/100000, Avg Reward: 7.52, Epsilon: 0.01
Episode 76300/100000, Avg Reward: 8.27, Epsilon: 0.01
Episode 76400/100000, Avg Reward: 7.87, Epsilon: 0.01
Episode 76500/100000, Avg Reward: 7.35, Epsilon: 0.01
Episode 76600/100000, Avg Reward: 5.36, Epsilon: 0.01
Episode 76700/100000, Avg Reward: 7.62, Epsilon: 0.01
Episode 76800/100000, Avg Reward: 5.39, Epsilon: 0.00
Episode 76900/100000, Avg Reward: 7.26, Epsilon: 0.00
Episode 77000/100000, Avg Reward: 5.51, Epsilon: 0.00
Episode 77100/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 77200/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 77300/100000, Avg Reward: 7.20, Epsilon: 0.00
Episode 77400/100000, Avg Reward: 3.36, Epsilon: 0.00
Episode 77500/100000, Avg Reward: 5.50, Epsilon: 0.00
Episode 77600/100000, Avg Reward: 5.43, Epsilon: 0.00
Episode 77700/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 77800/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 77900/100000, Avg Reward: 7.29, Epsilon: 0.00
Episode 78000/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 78100/100000, Avg Reward: 7.39, Epsilon: 0.00
Episode 78200/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 78300/100000, Avg Reward: 7.19, Epsilon: 0.00
Episode 78400/100000, Avg Reward: 7.87, Epsilon: 0.00
Episode 78500/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 78600/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 78700/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 78800/100000, Avg Reward: 6.21, Epsilon: 0.00
Episode 78900/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 79000/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 79100/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 79200/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 79300/100000, Avg Reward: 7.54, Epsilon: 0.00
Episode 79400/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 79500/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 79600/100000, Avg Reward: 5.85, Epsilon: 0.00
Episode 79700/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 79800/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 79900/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 80000/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 80100/100000, Avg Reward: 6.87, Epsilon: 0.00
Episode 80200/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 80300/100000, Avg Reward: 8.01, Epsilon: 0.00
```

```
Episode 80400/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 80500/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 80600/100000, Avg Reward: 7.27, Epsilon: 0.00
Episode 80700/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 80800/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 80900/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 81000/100000, Avg Reward: 5.25, Epsilon: 0.00
Episode 81100/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 81200/100000, Avg Reward: 7.22, Epsilon: 0.00
Episode 81300/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 81400/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 81500/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 81600/100000, Avg Reward: 7.03, Epsilon: 0.00
Episode 81700/100000, Avg Reward: 6.93, Epsilon: 0.00
Episode 81800/100000, Avg Reward: 7.49, Epsilon: 0.00
Episode 81900/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 82000/100000, Avg Reward: 7.40, Epsilon: 0.00
Episode 82100/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 82200/100000, Avg Reward: 5.45, Epsilon: 0.00
Episode 82300/100000, Avg Reward: 8.30, Epsilon: 0.00
Episode 82400/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 82500/100000, Avg Reward: 5.11, Epsilon: 0.00
Episode 82600/100000, Avg Reward: 7.89, Epsilon: 0.00
Episode 82700/100000, Avg Reward: 8.15, Epsilon: 0.00
Episode 82800/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 82900/100000, Avg Reward: 7.65, Epsilon: 0.00
Episode 83000/100000, Avg Reward: 5.62, Epsilon: 0.00
Episode 83100/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 83200/100000, Avg Reward: -2.09, Epsilon: 0.00
Episode 83300/100000, Avg Reward: 6.85, Epsilon: 0.00
Episode 83400/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 83500/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 83600/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 83700/100000, Avg Reward: 7.47, Epsilon: 0.00
Episode 83800/100000, Avg Reward: 8.41, Epsilon: 0.00
Episode 83900/100000, Avg Reward: 8.18, Epsilon: 0.00
Episode 84000/100000, Avg Reward: 7.46, Epsilon: 0.00
Episode 84100/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 84200/100000, Avg Reward: 7.81, Epsilon: 0.00
Episode 84300/100000, Avg Reward: 7.58, Epsilon: 0.00
Episode 84400/100000, Avg Reward: 7.38, Epsilon: 0.00
Episode 84500/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 84600/100000, Avg Reward: 5.29, Epsilon: 0.00
Episode 84700/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 84800/100000, Avg Reward: 5.66, Epsilon: 0.00
Episode 84900/100000, Avg Reward: 7.42, Epsilon: 0.00
Episode 85000/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 85100/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 85200/100000, Avg Reward: 7.44, Epsilon: 0.00
Episode 85300/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 85400/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 85500/100000, Avg Reward: 5.20, Epsilon: 0.00
Episode 85600/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 85700/100000, Avg Reward: 5.45, Epsilon: 0.00
Episode 85800/100000, Avg Reward: 7.69, Epsilon: 0.00
Episode 85900/100000, Avg Reward: 7.44, Epsilon: 0.00
Episode 86000/100000, Avg Reward: 7.77, Epsilon: 0.00
```

```
Episode 86100/100000, Avg Reward: 7.41, Epsilon: 0.00
Episode 86200/100000, Avg Reward: 7.22, Epsilon: 0.00
Episode 86300/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 86400/100000, Avg Reward: 7.30, Epsilon: 0.00
Episode 86500/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 86600/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 86700/100000, Avg Reward: 8.05, Epsilon: 0.00
Episode 86800/100000, Avg Reward: 7.36, Epsilon: 0.00
Episode 86900/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 87000/100000, Avg Reward: 3.49, Epsilon: 0.00
Episode 87100/100000, Avg Reward: 7.33, Epsilon: 0.00
Episode 87200/100000, Avg Reward: 3.59, Epsilon: 0.00
Episode 87300/100000, Avg Reward: 5.43, Epsilon: 0.00
Episode 87400/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 87500/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 87600/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 87700/100000, Avg Reward: 7.91, Epsilon: 0.00
Episode 87800/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 87900/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 88000/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 88100/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 88200/100000, Avg Reward: 5.79, Epsilon: 0.00
Episode 88300/100000, Avg Reward: 1.57, Epsilon: 0.00
Episode 88400/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 88500/100000, Avg Reward: 8.33, Epsilon: 0.00
Episode 88600/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 88700/100000, Avg Reward: 1.02, Epsilon: 0.00
Episode 88800/100000, Avg Reward: 7.86, Epsilon: 0.00
Episode 88900/100000, Avg Reward: 7.60, Epsilon: 0.00
Episode 89000/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 89100/100000, Avg Reward: 7.61, Epsilon: 0.00
Episode 89200/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 89300/100000, Avg Reward: 7.00, Epsilon: 0.00
Episode 89400/100000, Avg Reward: 7.75, Epsilon: 0.00
Episode 89500/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 89600/100000, Avg Reward: 8.10, Epsilon: 0.00
Episode 89700/100000, Avg Reward: 1.84, Epsilon: 0.00
Episode 89800/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 89900/100000, Avg Reward: 7.92, Epsilon: 0.00
Episode 90000/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 90100/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 90200/100000, Avg Reward: 7.95, Epsilon: 0.00
Episode 90300/100000, Avg Reward: 7.48, Epsilon: 0.00
Episode 90400/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 90500/100000, Avg Reward: 7.18, Epsilon: 0.00
Episode 90600/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 90700/100000, Avg Reward: 3.54, Epsilon: 0.00
Episode 90800/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 90900/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 91000/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 91100/100000, Avg Reward: 7.78, Epsilon: 0.00
Episode 91200/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 91300/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 91400/100000, Avg Reward: 8.42, Epsilon: 0.00
Episode 91500/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 91600/100000, Avg Reward: 7.79, Epsilon: 0.00
Episode 91700/100000, Avg Reward: 8.10, Epsilon: 0.00
```

```
Episode 91800/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 91900/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 92000/100000, Avg Reward: 5.76, Epsilon: 0.00
Episode 92100/100000, Avg Reward: 7.97, Epsilon: 0.00
Episode 92200/100000, Avg Reward: 8.11, Epsilon: 0.00
Episode 92300/100000, Avg Reward: 7.59, Epsilon: 0.00
Episode 92400/100000, Avg Reward: 7.80, Epsilon: 0.00
Episode 92500/100000, Avg Reward: 5.81, Epsilon: 0.00
Episode 92600/100000, Avg Reward: 5.49, Epsilon: 0.00
Episode 92700/100000, Avg Reward: 7.85, Epsilon: 0.00
Episode 92800/100000, Avg Reward: 8.04, Epsilon: 0.00
Episode 92900/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 93000/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 93100/100000, Avg Reward: 7.66, Epsilon: 0.00
Episode 93200/100000, Avg Reward: 7.73, Epsilon: 0.00
Episode 93300/100000, Avg Reward: 7.53, Epsilon: 0.00
Episode 93400/100000, Avg Reward: 7.68, Epsilon: 0.00
Episode 93500/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 93600/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 93700/100000, Avg Reward: 7.72, Epsilon: 0.00
Episode 93800/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 93900/100000, Avg Reward: 7.67, Epsilon: 0.00
Episode 94000/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 94100/100000, Avg Reward: 3.33, Epsilon: 0.00
Episode 94200/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 94300/100000, Avg Reward: 5.98, Epsilon: 0.00
Episode 94400/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 94500/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 94600/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 94700/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 94800/100000, Avg Reward: 7.88, Epsilon: 0.00
Episode 94900/100000, Avg Reward: 8.06, Epsilon: 0.00
Episode 95000/100000, Avg Reward: 8.07, Epsilon: 0.00
Episode 95100/100000, Avg Reward: 7.71, Epsilon: 0.00
Episode 95200/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 95300/100000, Avg Reward: 7.96, Epsilon: 0.00
Episode 95400/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 95500/100000, Avg Reward: 8.29, Epsilon: 0.00
Episode 95600/100000, Avg Reward: 5.73, Epsilon: 0.00
Episode 95700/100000, Avg Reward: 6.26, Epsilon: 0.00
Episode 95800/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 95900/100000, Avg Reward: 8.39, Epsilon: 0.00
Episode 96000/100000, Avg Reward: 7.63, Epsilon: 0.00
Episode 96100/100000, Avg Reward: 8.28, Epsilon: 0.00
Episode 96200/100000, Avg Reward: 7.64, Epsilon: 0.00
Episode 96300/100000, Avg Reward: 7.55, Epsilon: 0.00
Episode 96400/100000, Avg Reward: 7.57, Epsilon: 0.00
Episode 96500/100000, Avg Reward: 8.26, Epsilon: 0.00
Episode 96600/100000, Avg Reward: 7.51, Epsilon: 0.00
Episode 96700/100000, Avg Reward: 8.58, Epsilon: 0.00
Episode 96800/100000, Avg Reward: 7.93, Epsilon: 0.00
Episode 96900/100000, Avg Reward: 7.90, Epsilon: 0.00
Episode 97000/100000, Avg Reward: 3.63, Epsilon: 0.00
Episode 97100/100000, Avg Reward: 6.16, Epsilon: 0.00
Episode 97200/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 97300/100000, Avg Reward: 8.67, Epsilon: 0.00
Episode 97400/100000, Avg Reward: 8.04, Epsilon: 0.00
```

```
Episode 97500/100000, Avg Reward: 7.77, Epsilon: 0.00
Episode 97600/100000, Avg Reward: 8.00, Epsilon: 0.00
Episode 97700/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 97800/100000, Avg Reward: 7.83, Epsilon: 0.00
Episode 97900/100000, Avg Reward: 8.20, Epsilon: 0.00
Episode 98000/100000, Avg Reward: 7.94, Epsilon: 0.00
Episode 98100/100000, Avg Reward: 5.74, Epsilon: 0.00
Episode 98200/100000, Avg Reward: 3.14, Epsilon: 0.00
Episode 98300/100000, Avg Reward: 7.84, Epsilon: 0.00
Episode 98400/100000, Avg Reward: 8.01, Epsilon: 0.00
Episode 98500/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 98600/100000, Avg Reward: 6.57, Epsilon: 0.00
Episode 98700/100000, Avg Reward: 8.12, Epsilon: 0.00
Episode 98800/100000, Avg Reward: 7.82, Epsilon: 0.00
Episode 98900/100000, Avg Reward: 8.02, Epsilon: 0.00
Episode 99000/100000, Avg Reward: 8.17, Epsilon: 0.00
Episode 99100/100000, Avg Reward: 7.43, Epsilon: 0.00
Episode 99200/100000, Avg Reward: 7.99, Epsilon: 0.00
Episode 99300/100000, Avg Reward: 4.92, Epsilon: 0.00
Episode 99400/100000, Avg Reward: 8.27, Epsilon: 0.00
Episode 99500/100000, Avg Reward: 8.08, Epsilon: 0.00
Episode 99600/100000, Avg Reward: 7.62, Epsilon: 0.00
Episode 99700/100000, Avg Reward: 7.76, Epsilon: 0.00
Episode 99800/100000, Avg Reward: 7.70, Epsilon: 0.00
Episode 99900/100000, Avg Reward: 7.74, Epsilon: 0.00
Episode 100000/100000, Avg Reward: 8.01, Epsilon: 0.00
!!!!!!!!!!!!!!!!!Trained and Stored the model successfully!!!!!!!!!!!!!!!!!
CPU times: user 1h 22min 47s, sys: 3min 25s, total: 1h 26min 13s
Wall time: 21min 39s
```
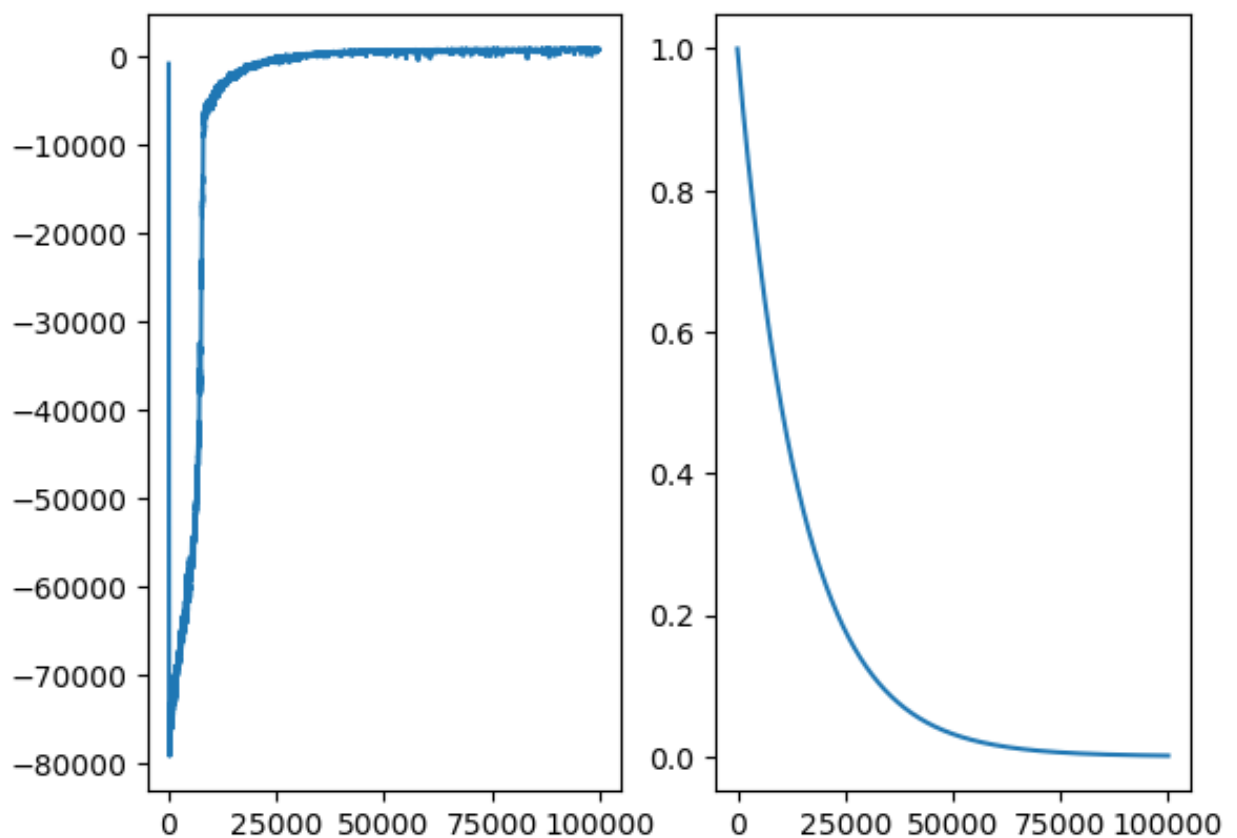


```
In [69]:  print(my_dqn_agent_fine_tuned_training_rewards)

[-812. -884. -821. ...    3.    6.   14.]
```

```
In [70]:  fine_tuned_training = {
              "my_q_agent_fine_tuned_training_rewards": list(my_q_agent_fine_tuned_tr
              "my_dqn_agent_fine_tuned_training_rewards": list(my_dqn_agent_fine_tune
          }
```

```
In [71]:  print(type(my_q_agent_fine_tuned_training_rewards))
          print(type(my_dqn_agent_fine_tuned_training_rewards))
```

```
<class 'numpy.ndarray'>
<class 'numpy.ndarray'>
```

```
In [72]:  with open("fine_tuned_training.json", "w") as f:
              json.dump(fine_tuned_training , f)
```

```
In [73]:  with open("fine_tuned_training.json", "r") as f:
              fine_tuned_training_dict = json.load(f)
```

Compare the training reward

```
In [75]:  x = range(1,100000 + 1)

          plt.figure(figsize=(10, 6))
          plt.plot(x, fine_tuned_training_dict["my_dqn_agent_fine_tuned_training_re
          plt.plot(x, fine_tuned_training_dict["my_q_agent_fine_tuned_training_rewa

          # titles and labels
          plt.title('Prediction Comparison (Fine Tuned Unique Parameters)')
          plt.xlabel('Time Episodes')
          plt.ylabel('Average Reward')
          plt.legend()
          # display
          plt.grid(True)
          plt.show()
```
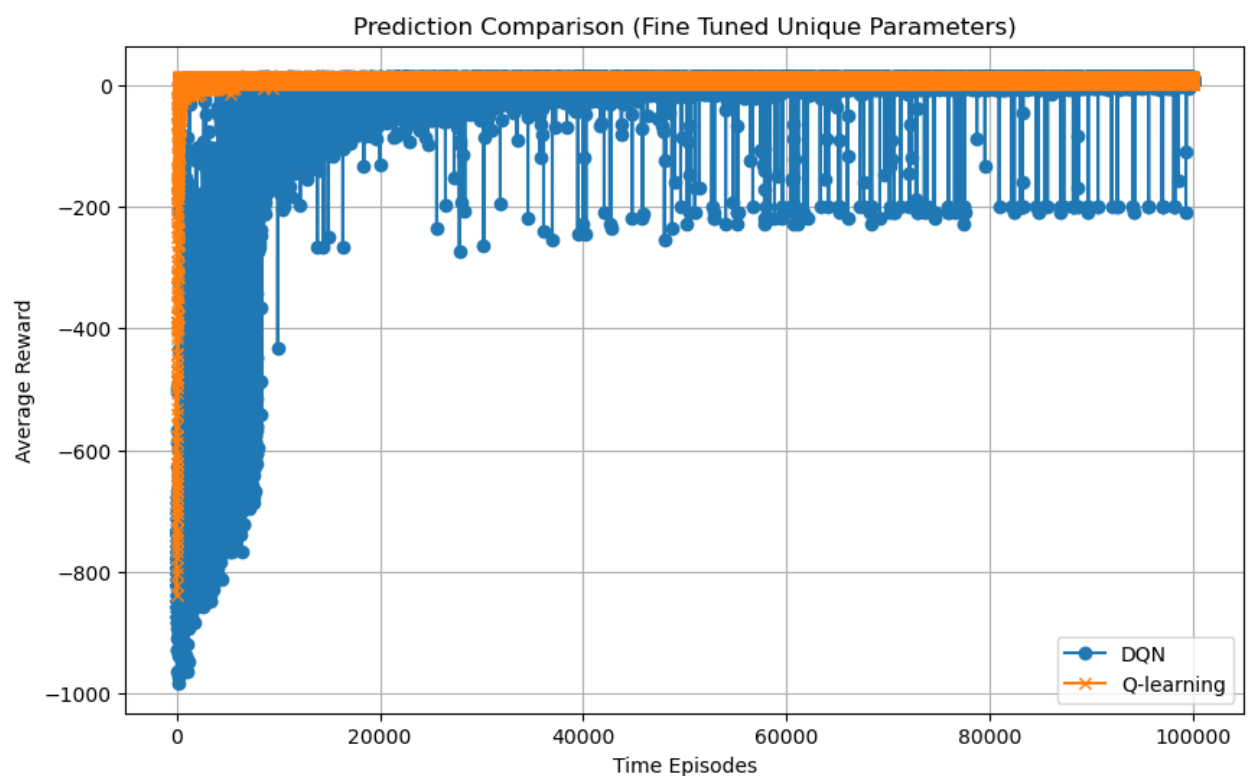
Prediction with the Fine-tunred models

Q Learning Agent with Fine Tuned Parameters

```
In [78]: my_q_agent_fine_tuned_avg_reward_lst = my_q_agent_fine_tuned.test(test_ep
```

```
============================================================
EPISODE  0
Score of total rewards: 7
============================================================
EPISODE  1
Score of total rewards: 10
============================================================
EPISODE  2
Score of total rewards: 8
============================================================
EPISODE  3
Score of total rewards: 14
============================================================
EPISODE  4
Score of total rewards: 4
============================================================
EPISODE  5
Score of total rewards: 8
============================================================
EPISODE  6
Score of total rewards: 5
============================================================
EPISODE  7
Score of total rewards: 10
============================================================
EPISODE  8
Score of total rewards: 13
============================================================
EPISODE  9
Score of total rewards: 8
============================================================
EPISODE  10
Score of total rewards: 10
============================================================
EPISODE  11
Score of total rewards: 4
============================================================
EPISODE  12
Score of total rewards: 9
============================================================
EPISODE  13
Score of total rewards: 6
============================================================
EPISODE  14
Score of total rewards: 5
============================================================
EPISODE  15
Score of total rewards: 5
============================================================
EPISODE  16
Score of total rewards: 3
```

```
=========================================================
EPISODE  17
Score of total rewards: 6
=========================================================
EPISODE  18
Score of total rewards: 8
=========================================================
EPISODE  19
Score of total rewards: 9
=========================================================
EPISODE  20
Score of total rewards: 5
=========================================================
EPISODE  21
Score of total rewards: 7
=========================================================
EPISODE  22
Score of total rewards: 6
=========================================================
EPISODE  23
Score of total rewards: 12
=========================================================
EPISODE  24
Score of total rewards: 8
=========================================================
EPISODE  25
Score of total rewards: 7
=========================================================
EPISODE  26
Score of total rewards: 6
=========================================================
EPISODE  27
Score of total rewards: 8
=========================================================
EPISODE  28
Score of total rewards: 6
=========================================================
EPISODE  29
Score of total rewards: 8
=========================================================
EPISODE  30
Score of total rewards: 10
=========================================================
EPISODE  31
Score of total rewards: 9
=========================================================
EPISODE  32
Score of total rewards: 13
=========================================================
EPISODE  33
Score of total rewards: 12
=========================================================
EPISODE  34
Score of total rewards: 11
=========================================================
EPISODE  35
Score of total rewards: 12
```

```
========================================================
EPISODE  36
Score of total rewards: 8
========================================================
EPISODE  37
Score of total rewards: 11
========================================================
EPISODE  38
Score of total rewards: 7
========================================================
EPISODE  39
Score of total rewards: 12
========================================================
EPISODE  40
Score of total rewards: 5
========================================================
EPISODE  41
Score of total rewards: 8
========================================================
EPISODE  42
Score of total rewards: 10
========================================================
EPISODE  43
Score of total rewards: 10
========================================================
EPISODE  44
Score of total rewards: 3
========================================================
EPISODE  45
Score of total rewards: 7
========================================================
EPISODE  46
Score of total rewards: 5
========================================================
EPISODE  47
Score of total rewards: 10
========================================================
EPISODE  48
Score of total rewards: 10
========================================================
EPISODE  49
Score of total rewards: 10
========================================================
EPISODE  50
Score of total rewards: 5
========================================================
EPISODE  51
Score of total rewards: 13
========================================================
EPISODE  52
Score of total rewards: 11
========================================================
EPISODE  53
Score of total rewards: 9
========================================================
EPISODE  54
Score of total rewards: 12
```

```
========================================================
EPISODE  55
Score of total rewards: 6
========================================================
EPISODE  56
Score of total rewards: 7
========================================================
EPISODE  57
Score of total rewards: 6
========================================================
EPISODE  58
Score of total rewards: 9
========================================================
EPISODE  59
Score of total rewards: 15
========================================================
EPISODE  60
Score of total rewards: 12
========================================================
EPISODE  61
Score of total rewards: 3
========================================================
EPISODE  62
Score of total rewards: 7
========================================================
EPISODE  63
Score of total rewards: 5
========================================================
EPISODE  64
Score of total rewards: 8
========================================================
EPISODE  65
Score of total rewards: 8
========================================================
EPISODE  66
Score of total rewards: 10
========================================================
EPISODE  67
Score of total rewards: 3
========================================================
EPISODE  68
Score of total rewards: 5
========================================================
EPISODE  69
Score of total rewards: 13
========================================================
EPISODE  70
Score of total rewards: 10
========================================================
EPISODE  71
Score of total rewards: 5
========================================================
EPISODE  72
Score of total rewards: 4
========================================================
EPISODE  73
Score of total rewards: 11
```

```
========================================================
EPISODE  74
Score of total rewards: 11
========================================================
EPISODE  75
Score of total rewards: 8
========================================================
EPISODE  76
Score of total rewards: 13
========================================================
EPISODE  77
Score of total rewards: 4
========================================================
EPISODE  78
Score of total rewards: 7
========================================================
EPISODE  79
Score of total rewards: 9
========================================================
EPISODE  80
Score of total rewards: 5
========================================================
EPISODE  81
Score of total rewards: 6
========================================================
EPISODE  82
Score of total rewards: 7
========================================================
EPISODE  83
Score of total rewards: 6
========================================================
EPISODE  84
Score of total rewards: 7
========================================================
EPISODE  85
Score of total rewards: 7
========================================================
EPISODE  86
Score of total rewards: 9
========================================================
EPISODE  87
Score of total rewards: 7
========================================================
EPISODE  88
Score of total rewards: 8
========================================================
EPISODE  89
Score of total rewards: 14
========================================================
EPISODE  90
Score of total rewards: 8
========================================================
EPISODE  91
Score of total rewards: 12
========================================================
EPISODE  92
Score of total rewards: 5
```

```
========================================================
EPISODE  93
Score of total rewards: 8
========================================================
EPISODE  94
Score of total rewards: 9
========================================================
EPISODE  95
Score of total rewards: 8
========================================================
EPISODE  96
Score of total rewards: 11
========================================================
EPISODE  97
Score of total rewards: 8
========================================================
EPISODE  98
Score of total rewards: 9
========================================================
EPISODE  99
Score of total rewards: 4
========================================================
EPISODE  100
Score of total rewards: 5
========================================================
EPISODE  101
Score of total rewards: 7
========================================================
EPISODE  102
Score of total rewards: 10
========================================================
EPISODE  103
Score of total rewards: 13
========================================================
EPISODE  104
Score of total rewards: 3
========================================================
EPISODE  105
Score of total rewards: 7
========================================================
EPISODE  106
Score of total rewards: 11
========================================================
EPISODE  107
Score of total rewards: 7
========================================================
EPISODE  108
Score of total rewards: 12
========================================================
EPISODE  109
Score of total rewards: 8
========================================================
EPISODE  110
Score of total rewards: 5
========================================================
EPISODE  111
Score of total rewards: 4
```

```
========================================================
EPISODE  112
Score of total rewards: 9
========================================================
EPISODE  113
Score of total rewards: 7
========================================================
EPISODE  114
Score of total rewards: 7
========================================================
EPISODE  115
Score of total rewards: 8
========================================================
EPISODE  116
Score of total rewards: 7
========================================================
EPISODE  117
Score of total rewards: 11
========================================================
EPISODE  118
Score of total rewards: 7
========================================================
EPISODE  119
Score of total rewards: 11
========================================================
EPISODE  120
Score of total rewards: 11
========================================================
EPISODE  121
Score of total rewards: 8
========================================================
EPISODE  122
Score of total rewards: 10
========================================================
EPISODE  123
Score of total rewards: 7
========================================================
EPISODE  124
Score of total rewards: 11
========================================================
EPISODE  125
Score of total rewards: 6
========================================================
EPISODE  126
Score of total rewards: 7
========================================================
EPISODE  127
Score of total rewards: 4
========================================================
EPISODE  128
Score of total rewards: 12
========================================================
EPISODE  129
Score of total rewards: 4
========================================================
EPISODE  130
Score of total rewards: 7
```

```
========================================================
EPISODE  131
Score of total rewards: 8
========================================================
EPISODE  132
Score of total rewards: 4
========================================================
EPISODE  133
Score of total rewards: 5
========================================================
EPISODE  134
Score of total rewards: 8
========================================================
EPISODE  135
Score of total rewards: 8
========================================================
EPISODE  136
Score of total rewards: 5
========================================================
EPISODE  137
Score of total rewards: 8
========================================================
EPISODE  138
Score of total rewards: 8
========================================================
EPISODE  139
Score of total rewards: 14
========================================================
EPISODE  140
Score of total rewards: 4
========================================================
EPISODE  141
Score of total rewards: 10
========================================================
EPISODE  142
Score of total rewards: 4
========================================================
EPISODE  143
Score of total rewards: 8
========================================================
EPISODE  144
Score of total rewards: 11
========================================================
EPISODE  145
Score of total rewards: 10
========================================================
EPISODE  146
Score of total rewards: 9
========================================================
EPISODE  147
Score of total rewards: 11
========================================================
EPISODE  148
Score of total rewards: 12
========================================================
EPISODE  149
Score of total rewards: 7
```

```
========================================================
EPISODE  150
Score of total rewards: 9
========================================================
EPISODE  151
Score of total rewards: 4
========================================================
EPISODE  152
Score of total rewards: 10
========================================================
EPISODE  153
Score of total rewards: 10
========================================================
EPISODE  154
Score of total rewards: 13
========================================================
EPISODE  155
Score of total rewards: 4
========================================================
EPISODE  156
Score of total rewards: 3
========================================================
EPISODE  157
Score of total rewards: 9
========================================================
EPISODE  158
Score of total rewards: 7
========================================================
EPISODE  159
Score of total rewards: 4
========================================================
EPISODE  160
Score of total rewards: 5
========================================================
EPISODE  161
Score of total rewards: 10
========================================================
EPISODE  162
Score of total rewards: 3
========================================================
EPISODE  163
Score of total rewards: 5
========================================================
EPISODE  164
Score of total rewards: 4
========================================================
EPISODE  165
Score of total rewards: 5
========================================================
EPISODE  166
Score of total rewards: 7
========================================================
EPISODE  167
Score of total rewards: 11
========================================================
EPISODE  168
Score of total rewards: 14
```

```
========================================================
EPISODE  169
Score of total rewards: 3
========================================================
EPISODE  170
Score of total rewards: 10
========================================================
EPISODE  171
Score of total rewards: 9
========================================================
EPISODE  172
Score of total rewards: 6
========================================================
EPISODE  173
Score of total rewards: 7
========================================================
EPISODE  174
Score of total rewards: 10
========================================================
EPISODE  175
Score of total rewards: 6
========================================================
EPISODE  176
Score of total rewards: 7
========================================================
EPISODE  177
Score of total rewards: 8
========================================================
EPISODE  178
Score of total rewards: 4
========================================================
EPISODE  179
Score of total rewards: 7
========================================================
EPISODE  180
Score of total rewards: 4
========================================================
EPISODE  181
Score of total rewards: 8
========================================================
EPISODE  182
Score of total rewards: 12
========================================================
EPISODE  183
Score of total rewards: 12
========================================================
EPISODE  184
Score of total rewards: 7
========================================================
EPISODE  185
Score of total rewards: 8
========================================================
EPISODE  186
Score of total rewards: 4
========================================================
EPISODE  187
Score of total rewards: 6
```

```
========================================================
EPISODE  188
Score of total rewards: 4
========================================================
EPISODE  189
Score of total rewards: 13
========================================================
EPISODE  190
Score of total rewards: 5
========================================================
EPISODE  191
Score of total rewards: 6
========================================================
EPISODE  192
Score of total rewards: 9
========================================================
EPISODE  193
Score of total rewards: 3
========================================================
EPISODE  194
Score of total rewards: 7
========================================================
EPISODE  195
Score of total rewards: 7
========================================================
EPISODE  196
Score of total rewards: 4
========================================================
EPISODE  197
Score of total rewards: 7
========================================================
EPISODE  198
Score of total rewards: 8
========================================================
EPISODE  199
Score of total rewards: 10
========================================================
EPISODE  200
Score of total rewards: 12
========================================================
EPISODE  201
Score of total rewards: 11
========================================================
EPISODE  202
Score of total rewards: 9
========================================================
EPISODE  203
Score of total rewards: 14
========================================================
EPISODE  204
Score of total rewards: 7
========================================================
EPISODE  205
Score of total rewards: 8
========================================================
EPISODE  206
Score of total rewards: 5
```

```
========================================================
EPISODE  207
Score of total rewards: 7
========================================================
EPISODE  208
Score of total rewards: 9
========================================================
EPISODE  209
Score of total rewards: 6
========================================================
EPISODE  210
Score of total rewards: 11
========================================================
EPISODE  211
Score of total rewards: 14
========================================================
EPISODE  212
Score of total rewards: 6
========================================================
EPISODE  213
Score of total rewards: 8
========================================================
EPISODE  214
Score of total rewards: 9
========================================================
EPISODE  215
Score of total rewards: 11
========================================================
EPISODE  216
Score of total rewards: 6
========================================================
EPISODE  217
Score of total rewards: 9
========================================================
EPISODE  218
Score of total rewards: 8
========================================================
EPISODE  219
Score of total rewards: 5
========================================================
EPISODE  220
Score of total rewards: 9
========================================================
EPISODE  221
Score of total rewards: 11
========================================================
EPISODE  222
Score of total rewards: 5
========================================================
EPISODE  223
Score of total rewards: 4
========================================================
EPISODE  224
Score of total rewards: 10
========================================================
EPISODE  225
Score of total rewards: 4
```

```
========================================================
EPISODE  226
Score of total rewards: 11
========================================================
EPISODE  227
Score of total rewards: 8
========================================================
EPISODE  228
Score of total rewards: 3
========================================================
EPISODE  229
Score of total rewards: 11
========================================================
EPISODE  230
Score of total rewards: 7
========================================================
EPISODE  231
Score of total rewards: 9
========================================================
EPISODE  232
Score of total rewards: 10
========================================================
EPISODE  233
Score of total rewards: 12
========================================================
EPISODE  234
Score of total rewards: 10
========================================================
EPISODE  235
Score of total rewards: 6
========================================================
EPISODE  236
Score of total rewards: 8
========================================================
EPISODE  237
Score of total rewards: 14
========================================================
EPISODE  238
Score of total rewards: 6
========================================================
EPISODE  239
Score of total rewards: 6
========================================================
EPISODE  240
Score of total rewards: 9
========================================================
EPISODE  241
Score of total rewards: 7
========================================================
EPISODE  242
Score of total rewards: 10
========================================================
EPISODE  243
Score of total rewards: 7
========================================================
EPISODE  244
Score of total rewards: 7
```

```
========================================================
EPISODE  245
Score of total rewards: 8
========================================================
EPISODE  246
Score of total rewards: 7
========================================================
EPISODE  247
Score of total rewards: 8
========================================================
EPISODE  248
Score of total rewards: 7
========================================================
EPISODE  249
Score of total rewards: 4
========================================================
EPISODE  250
Score of total rewards: 7
========================================================
EPISODE  251
Score of total rewards: 7
========================================================
EPISODE  252
Score of total rewards: 4
========================================================
EPISODE  253
Score of total rewards: 7
========================================================
EPISODE  254
Score of total rewards: 7
========================================================
EPISODE  255
Score of total rewards: 7
========================================================
EPISODE  256
Score of total rewards: 9
========================================================
EPISODE  257
Score of total rewards: 15
========================================================
EPISODE  258
Score of total rewards: 7
========================================================
EPISODE  259
Score of total rewards: 9
========================================================
EPISODE  260
Score of total rewards: 8
========================================================
EPISODE  261
Score of total rewards: 6
========================================================
EPISODE  262
Score of total rewards: 5
========================================================
EPISODE  263
Score of total rewards: 8
```

```
========================================================
EPISODE  264
Score of total rewards: 8
========================================================
EPISODE  265
Score of total rewards: 10
========================================================
EPISODE  266
Score of total rewards: 7
========================================================
EPISODE  267
Score of total rewards: 10
========================================================
EPISODE  268
Score of total rewards: 8
========================================================
EPISODE  269
Score of total rewards: 7
========================================================
EPISODE  270
Score of total rewards: 5
========================================================
EPISODE  271
Score of total rewards: 7
========================================================
EPISODE  272
Score of total rewards: 7
========================================================
EPISODE  273
Score of total rewards: 8
========================================================
EPISODE  274
Score of total rewards: 8
========================================================
EPISODE  275
Score of total rewards: 7
========================================================
EPISODE  276
Score of total rewards: 10
========================================================
EPISODE  277
Score of total rewards: 12
========================================================
EPISODE  278
Score of total rewards: 7
========================================================
EPISODE  279
Score of total rewards: 7
========================================================
EPISODE  280
Score of total rewards: 8
========================================================
EPISODE  281
Score of total rewards: 8
========================================================
EPISODE  282
Score of total rewards: 5
```

```
========================================================
EPISODE  283
Score of total rewards: 15
========================================================
EPISODE  284
Score of total rewards: 7
========================================================
EPISODE  285
Score of total rewards: 8
========================================================
EPISODE  286
Score of total rewards: 8
========================================================
EPISODE  287
Score of total rewards: 8
========================================================
EPISODE  288
Score of total rewards: 8
========================================================
EPISODE  289
Score of total rewards: 3
========================================================
EPISODE  290
Score of total rewards: 4
========================================================
EPISODE  291
Score of total rewards: 12
========================================================
EPISODE  292
Score of total rewards: 8
========================================================
EPISODE  293
Score of total rewards: 8
========================================================
EPISODE  294
Score of total rewards: 9
========================================================
EPISODE  295
Score of total rewards: 7
========================================================
EPISODE  296
Score of total rewards: 12
========================================================
EPISODE  297
Score of total rewards: 12
========================================================
EPISODE  298
Score of total rewards: 4
========================================================
EPISODE  299
Score of total rewards: 5
========================================================
EPISODE  300
Score of total rewards: 6
========================================================
EPISODE  301
Score of total rewards: 7
```

```
========================================================
EPISODE  302
Score of total rewards: 11
========================================================
EPISODE  303
Score of total rewards: 6
========================================================
EPISODE  304
Score of total rewards: 11
========================================================
EPISODE  305
Score of total rewards: 13
========================================================
EPISODE  306
Score of total rewards: 4
========================================================
EPISODE  307
Score of total rewards: 8
========================================================
EPISODE  308
Score of total rewards: 9
========================================================
EPISODE  309
Score of total rewards: 11
========================================================
EPISODE  310
Score of total rewards: 11
========================================================
EPISODE  311
Score of total rewards: 6
========================================================
EPISODE  312
Score of total rewards: 12
========================================================
EPISODE  313
Score of total rewards: 10
========================================================
EPISODE  314
Score of total rewards: 10
========================================================
EPISODE  315
Score of total rewards: 7
========================================================
EPISODE  316
Score of total rewards: 6
========================================================
EPISODE  317
Score of total rewards: 8
========================================================
EPISODE  318
Score of total rewards: 4
========================================================
EPISODE  319
Score of total rewards: 5
========================================================
EPISODE  320
Score of total rewards: 7
```

```
========================================================
EPISODE  321
Score of total rewards: 10
========================================================
EPISODE  322
Score of total rewards: 9
========================================================
EPISODE  323
Score of total rewards: 6
========================================================
EPISODE  324
Score of total rewards: 11
========================================================
EPISODE  325
Score of total rewards: 6
========================================================
EPISODE  326
Score of total rewards: 11
========================================================
EPISODE  327
Score of total rewards: 8
========================================================
EPISODE  328
Score of total rewards: 9
========================================================
EPISODE  329
Score of total rewards: 9
========================================================
EPISODE  330
Score of total rewards: 8
========================================================
EPISODE  331
Score of total rewards: 6
========================================================
EPISODE  332
Score of total rewards: 6
========================================================
EPISODE  333
Score of total rewards: 9
========================================================
EPISODE  334
Score of total rewards: 8
========================================================
EPISODE  335
Score of total rewards: 7
========================================================
EPISODE  336
Score of total rewards: 7
========================================================
EPISODE  337
Score of total rewards: 7
========================================================
EPISODE  338
Score of total rewards: 12
========================================================
EPISODE  339
Score of total rewards: 8
```

```
========================================================
EPISODE  340
Score of total rewards: 7
========================================================
EPISODE  341
Score of total rewards: 4
========================================================
EPISODE  342
Score of total rewards: 11
========================================================
EPISODE  343
Score of total rewards: 7
========================================================
EPISODE  344
Score of total rewards: 6
========================================================
EPISODE  345
Score of total rewards: 9
========================================================
EPISODE  346
Score of total rewards: 6
========================================================
EPISODE  347
Score of total rewards: 8
========================================================
EPISODE  348
Score of total rewards: 9
========================================================
EPISODE  349
Score of total rewards: 6
========================================================
EPISODE  350
Score of total rewards: 8
========================================================
EPISODE  351
Score of total rewards: 10
========================================================
EPISODE  352
Score of total rewards: 11
========================================================
EPISODE  353
Score of total rewards: 9
========================================================
EPISODE  354
Score of total rewards: 6
========================================================
EPISODE  355
Score of total rewards: 14
========================================================
EPISODE  356
Score of total rewards: 14
========================================================
EPISODE  357
Score of total rewards: 6
========================================================
EPISODE  358
Score of total rewards: 7
```

```
========================================================
EPISODE  359
Score of total rewards: 6
========================================================
EPISODE  360
Score of total rewards: 7
========================================================
EPISODE  361
Score of total rewards: 9
========================================================
EPISODE  362
Score of total rewards: 7
========================================================
EPISODE  363
Score of total rewards: 8
========================================================
EPISODE  364
Score of total rewards: 7
========================================================
EPISODE  365
Score of total rewards: 9
========================================================
EPISODE  366
Score of total rewards: 6
========================================================
EPISODE  367
Score of total rewards: 10
========================================================
EPISODE  368
Score of total rewards: 10
========================================================
EPISODE  369
Score of total rewards: 8
========================================================
EPISODE  370
Score of total rewards: 8
========================================================
EPISODE  371
Score of total rewards: 4
========================================================
EPISODE  372
Score of total rewards: 13
========================================================
EPISODE  373
Score of total rewards: 8
========================================================
EPISODE  374
Score of total rewards: 5
========================================================
EPISODE  375
Score of total rewards: 10
========================================================
EPISODE  376
Score of total rewards: 6
========================================================
EPISODE  377
Score of total rewards: 9
```

```
============================================================
EPISODE  378
Score of total rewards: 8
============================================================
EPISODE  379
Score of total rewards: 3
============================================================
EPISODE  380
Score of total rewards: 11
============================================================
EPISODE  381
Score of total rewards: 7
============================================================
EPISODE  382
Score of total rewards: 5
============================================================
EPISODE  383
Score of total rewards: 11
============================================================
EPISODE  384
Score of total rewards: 9
============================================================
EPISODE  385
Score of total rewards: 10
============================================================
EPISODE  386
Score of total rewards: 6
============================================================
EPISODE  387
Score of total rewards: 5
============================================================
EPISODE  388
Score of total rewards: 8
============================================================
EPISODE  389
Score of total rewards: 5
============================================================
EPISODE  390
Score of total rewards: 10
============================================================
EPISODE  391
Score of total rewards: 5
============================================================
EPISODE  392
Score of total rewards: 6
============================================================
EPISODE  393
Score of total rewards: 8
============================================================
EPISODE  394
Score of total rewards: 10
============================================================
EPISODE  395
Score of total rewards: 8
============================================================
EPISODE  396
Score of total rewards: 5
```

```
==========================================================
EPISODE  397
Score of total rewards: 9
==========================================================
EPISODE  398
Score of total rewards: 10
==========================================================
EPISODE  399
Score of total rewards: 7
==========================================================
EPISODE  400
Score of total rewards: 9
==========================================================
EPISODE  401
Score of total rewards: 8
==========================================================
EPISODE  402
Score of total rewards: 6
==========================================================
EPISODE  403
Score of total rewards: 9
==========================================================
EPISODE  404
Score of total rewards: 10
==========================================================
EPISODE  405
Score of total rewards: 7
==========================================================
EPISODE  406
Score of total rewards: 5
==========================================================
EPISODE  407
Score of total rewards: 13
==========================================================
EPISODE  408
Score of total rewards: 7
==========================================================
EPISODE  409
Score of total rewards: 11
==========================================================
EPISODE  410
Score of total rewards: 7
==========================================================
EPISODE  411
Score of total rewards: 8
==========================================================
EPISODE  412
Score of total rewards: 7
==========================================================
EPISODE  413
Score of total rewards: 10
==========================================================
EPISODE  414
Score of total rewards: 8
==========================================================
EPISODE  415
Score of total rewards: 6
```

```
========================================================
EPISODE  416
Score of total rewards: 5
========================================================
EPISODE  417
Score of total rewards: 10
========================================================
EPISODE  418
Score of total rewards: 4
========================================================
EPISODE  419
Score of total rewards: 10
========================================================
EPISODE  420
Score of total rewards: 8
========================================================
EPISODE  421
Score of total rewards: 5
========================================================
EPISODE  422
Score of total rewards: 5
========================================================
EPISODE  423
Score of total rewards: 8
========================================================
EPISODE  424
Score of total rewards: 14
========================================================
EPISODE  425
Score of total rewards: 8
========================================================
EPISODE  426
Score of total rewards: 6
========================================================
EPISODE  427
Score of total rewards: 9
========================================================
EPISODE  428
Score of total rewards: 8
========================================================
EPISODE  429
Score of total rewards: 9
========================================================
EPISODE  430
Score of total rewards: 7
========================================================
EPISODE  431
Score of total rewards: 7
========================================================
EPISODE  432
Score of total rewards: 7
========================================================
EPISODE  433
Score of total rewards: 9
========================================================
EPISODE  434
Score of total rewards: 5
```

```
========================================================
EPISODE  435
Score of total rewards: 13
========================================================
EPISODE  436
Score of total rewards: 7
========================================================
EPISODE  437
Score of total rewards: 12
========================================================
EPISODE  438
Score of total rewards: 9
========================================================
EPISODE  439
Score of total rewards: 9
========================================================
EPISODE  440
Score of total rewards: 6
========================================================
EPISODE  441
Score of total rewards: 6
========================================================
EPISODE  442
Score of total rewards: 5
========================================================
EPISODE  443
Score of total rewards: 8
========================================================
EPISODE  444
Score of total rewards: 14
========================================================
EPISODE  445
Score of total rewards: 5
========================================================
EPISODE  446
Score of total rewards: 5
========================================================
EPISODE  447
Score of total rewards: 9
========================================================
EPISODE  448
Score of total rewards: 12
========================================================
EPISODE  449
Score of total rewards: 11
========================================================
EPISODE  450
Score of total rewards: 9
========================================================
EPISODE  451
Score of total rewards: 6
========================================================
EPISODE  452
Score of total rewards: 7
========================================================
EPISODE  453
Score of total rewards: 6
```

```
===========================================================
EPISODE  454
Score of total rewards: 3
===========================================================
EPISODE  455
Score of total rewards: 12
===========================================================
EPISODE  456
Score of total rewards: 12
===========================================================
EPISODE  457
Score of total rewards: 5
===========================================================
EPISODE  458
Score of total rewards: 13
===========================================================
EPISODE  459
Score of total rewards: 9
===========================================================
EPISODE  460
Score of total rewards: 5
===========================================================
EPISODE  461
Score of total rewards: 3
===========================================================
EPISODE  462
Score of total rewards: 9
===========================================================
EPISODE  463
Score of total rewards: 4
===========================================================
EPISODE  464
Score of total rewards: 12
===========================================================
EPISODE  465
Score of total rewards: 9
===========================================================
EPISODE  466
Score of total rewards: 9
===========================================================
EPISODE  467
Score of total rewards: 4
===========================================================
EPISODE  468
Score of total rewards: 11
===========================================================
EPISODE  469
Score of total rewards: 7
===========================================================
EPISODE  470
Score of total rewards: 9
===========================================================
EPISODE  471
Score of total rewards: 5
===========================================================
EPISODE  472
Score of total rewards: 7
```

```
========================================================
EPISODE  473
Score of total rewards: 7
========================================================
EPISODE  474
Score of total rewards: 6
========================================================
EPISODE  475
Score of total rewards: 8
========================================================
EPISODE  476
Score of total rewards: 8
========================================================
EPISODE  477
Score of total rewards: 6
========================================================
EPISODE  478
Score of total rewards: 10
========================================================
EPISODE  479
Score of total rewards: 6
========================================================
EPISODE  480
Score of total rewards: 6
========================================================
EPISODE  481
Score of total rewards: 8
========================================================
EPISODE  482
Score of total rewards: 4
========================================================
EPISODE  483
Score of total rewards: 13
========================================================
EPISODE  484
Score of total rewards: 4
========================================================
EPISODE  485
Score of total rewards: 10
========================================================
EPISODE  486
Score of total rewards: 9
========================================================
EPISODE  487
Score of total rewards: 10
========================================================
EPISODE  488
Score of total rewards: 9
========================================================
EPISODE  489
Score of total rewards: 6
========================================================
EPISODE  490
Score of total rewards: 8
========================================================
EPISODE  491
Score of total rewards: 3
```

```
========================================================
EPISODE  492
Score of total rewards: 8
========================================================
EPISODE  493
Score of total rewards: 8
========================================================
EPISODE  494
Score of total rewards: 9
========================================================
EPISODE  495
Score of total rewards: 7
========================================================
EPISODE  496
Score of total rewards: 7
========================================================
EPISODE  497
Score of total rewards: 13
========================================================
EPISODE  498
Score of total rewards: 7
========================================================
EPISODE  499
Score of total rewards: 8
========================================================
EPISODE  500
Score of total rewards: 8
========================================================
EPISODE  501
Score of total rewards: 4
========================================================
EPISODE  502
Score of total rewards: 8
========================================================
EPISODE  503
Score of total rewards: 8
========================================================
EPISODE  504
Score of total rewards: 11
========================================================
EPISODE  505
Score of total rewards: 5
========================================================
EPISODE  506
Score of total rewards: 4
========================================================
EPISODE  507
Score of total rewards: 14
========================================================
EPISODE  508
Score of total rewards: 7
========================================================
EPISODE  509
Score of total rewards: 7
========================================================
EPISODE  510
Score of total rewards: 8
```

```
========================================================
EPISODE  511
Score of total rewards: 8
========================================================
EPISODE  512
Score of total rewards: 8
========================================================
EPISODE  513
Score of total rewards: 8
========================================================
EPISODE  514
Score of total rewards: 6
========================================================
EPISODE  515
Score of total rewards: 8
========================================================
EPISODE  516
Score of total rewards: 7
========================================================
EPISODE  517
Score of total rewards: 9
========================================================
EPISODE  518
Score of total rewards: 8
========================================================
EPISODE  519
Score of total rewards: 6
========================================================
EPISODE  520
Score of total rewards: 8
========================================================
EPISODE  521
Score of total rewards: 13
========================================================
EPISODE  522
Score of total rewards: 10
========================================================
EPISODE  523
Score of total rewards: 9
========================================================
EPISODE  524
Score of total rewards: 8
========================================================
EPISODE  525
Score of total rewards: 4
========================================================
EPISODE  526
Score of total rewards: 7
========================================================
EPISODE  527
Score of total rewards: 7
========================================================
EPISODE  528
Score of total rewards: 10
========================================================
EPISODE  529
Score of total rewards: 11
```

```
========================================================
EPISODE  530
Score of total rewards: 7
========================================================
EPISODE  531
Score of total rewards: 5
========================================================
EPISODE  532
Score of total rewards: 5
========================================================
EPISODE  533
Score of total rewards: 5
========================================================
EPISODE  534
Score of total rewards: 14
========================================================
EPISODE  535
Score of total rewards: 7
========================================================
EPISODE  536
Score of total rewards: 8
========================================================
EPISODE  537
Score of total rewards: 6
========================================================
EPISODE  538
Score of total rewards: 11
========================================================
EPISODE  539
Score of total rewards: 7
========================================================
EPISODE  540
Score of total rewards: 6
========================================================
EPISODE  541
Score of total rewards: 9
========================================================
EPISODE  542
Score of total rewards: 8
========================================================
EPISODE  543
Score of total rewards: 7
========================================================
EPISODE  544
Score of total rewards: 10
========================================================
EPISODE  545
Score of total rewards: 10
========================================================
EPISODE  546
Score of total rewards: 13
========================================================
EPISODE  547
Score of total rewards: 8
========================================================
EPISODE  548
Score of total rewards: 8
```

```
============================================================
EPISODE  549
Score of total rewards: 11
============================================================
EPISODE  550
Score of total rewards: 8
============================================================
EPISODE  551
Score of total rewards: 4
============================================================
EPISODE  552
Score of total rewards: 4
============================================================
EPISODE  553
Score of total rewards: 10
============================================================
EPISODE  554
Score of total rewards: 11
============================================================
EPISODE  555
Score of total rewards: 7
============================================================
EPISODE  556
Score of total rewards: 5
============================================================
EPISODE  557
Score of total rewards: 10
============================================================
EPISODE  558
Score of total rewards: 4
============================================================
EPISODE  559
Score of total rewards: 5
============================================================
EPISODE  560
Score of total rewards: 9
============================================================
EPISODE  561
Score of total rewards: 7
============================================================
EPISODE  562
Score of total rewards: 7
============================================================
EPISODE  563
Score of total rewards: 10
============================================================
EPISODE  564
Score of total rewards: 8
============================================================
EPISODE  565
Score of total rewards: 8
============================================================
EPISODE  566
Score of total rewards: 10
============================================================
EPISODE  567
Score of total rewards: 11
```

```
========================================================
EPISODE  568
Score of total rewards: 7
========================================================
EPISODE  569
Score of total rewards: 5
========================================================
EPISODE  570
Score of total rewards: 8
========================================================
EPISODE  571
Score of total rewards: 9
========================================================
EPISODE  572
Score of total rewards: 11
========================================================
EPISODE  573
Score of total rewards: 14
========================================================
EPISODE  574
Score of total rewards: 11
========================================================
EPISODE  575
Score of total rewards: 5
========================================================
EPISODE  576
Score of total rewards: 10
========================================================
EPISODE  577
Score of total rewards: 9
========================================================
EPISODE  578
Score of total rewards: 11
========================================================
EPISODE  579
Score of total rewards: 7
========================================================
EPISODE  580
Score of total rewards: 7
========================================================
EPISODE  581
Score of total rewards: 5
========================================================
EPISODE  582
Score of total rewards: 6
========================================================
EPISODE  583
Score of total rewards: 8
========================================================
EPISODE  584
Score of total rewards: 8
========================================================
EPISODE  585
Score of total rewards: 8
========================================================
EPISODE  586
Score of total rewards: 8
```

```
========================================================
EPISODE  587
Score of total rewards: 10
========================================================
EPISODE  588
Score of total rewards: 3
========================================================
EPISODE  589
Score of total rewards: 11
========================================================
EPISODE  590
Score of total rewards: 12
========================================================
EPISODE  591
Score of total rewards: 13
========================================================
EPISODE  592
Score of total rewards: 5
========================================================
EPISODE  593
Score of total rewards: 13
========================================================
EPISODE  594
Score of total rewards: 15
========================================================
EPISODE  595
Score of total rewards: 7
========================================================
EPISODE  596
Score of total rewards: 10
========================================================
EPISODE  597
Score of total rewards: 5
========================================================
EPISODE  598
Score of total rewards: 10
========================================================
EPISODE  599
Score of total rewards: 7
========================================================
EPISODE  600
Score of total rewards: 8
========================================================
EPISODE  601
Score of total rewards: 4
========================================================
EPISODE  602
Score of total rewards: 7
========================================================
EPISODE  603
Score of total rewards: 7
========================================================
EPISODE  604
Score of total rewards: 5
========================================================
EPISODE  605
Score of total rewards: 9
```

```
========================================================
EPISODE  606
Score of total rewards: 8
========================================================
EPISODE  607
Score of total rewards: 9
========================================================
EPISODE  608
Score of total rewards: 7
========================================================
EPISODE  609
Score of total rewards: 7
========================================================
EPISODE  610
Score of total rewards: 12
========================================================
EPISODE  611
Score of total rewards: 4
========================================================
EPISODE  612
Score of total rewards: 9
========================================================
EPISODE  613
Score of total rewards: 9
========================================================
EPISODE  614
Score of total rewards: 9
========================================================
EPISODE  615
Score of total rewards: 5
========================================================
EPISODE  616
Score of total rewards: 6
========================================================
EPISODE  617
Score of total rewards: 6
========================================================
EPISODE  618
Score of total rewards: 7
========================================================
EPISODE  619
Score of total rewards: 7
========================================================
EPISODE  620
Score of total rewards: 7
========================================================
EPISODE  621
Score of total rewards: 8
========================================================
EPISODE  622
Score of total rewards: 9
========================================================
EPISODE  623
Score of total rewards: 8
========================================================
EPISODE  624
Score of total rewards: 7
```

```
========================================================
EPISODE  625
Score of total rewards: 8
========================================================
EPISODE  626
Score of total rewards: 10
========================================================
EPISODE  627
Score of total rewards: 8
========================================================
EPISODE  628
Score of total rewards: 9
========================================================
EPISODE  629
Score of total rewards: 5
========================================================
EPISODE  630
Score of total rewards: 11
========================================================
EPISODE  631
Score of total rewards: 7
========================================================
EPISODE  632
Score of total rewards: 6
========================================================
EPISODE  633
Score of total rewards: 8
========================================================
EPISODE  634
Score of total rewards: 7
========================================================
EPISODE  635
Score of total rewards: 8
========================================================
EPISODE  636
Score of total rewards: 9
========================================================
EPISODE  637
Score of total rewards: 7
========================================================
EPISODE  638
Score of total rewards: 7
========================================================
EPISODE  639
Score of total rewards: 5
========================================================
EPISODE  640
Score of total rewards: 4
========================================================
EPISODE  641
Score of total rewards: 9
========================================================
EPISODE  642
Score of total rewards: 13
========================================================
EPISODE  643
Score of total rewards: 10
```

```
========================================================
EPISODE  644
Score of total rewards: 12
========================================================
EPISODE  645
Score of total rewards: 9
========================================================
EPISODE  646
Score of total rewards: 6
========================================================
EPISODE  647
Score of total rewards: 8
========================================================
EPISODE  648
Score of total rewards: 10
========================================================
EPISODE  649
Score of total rewards: 9
========================================================
EPISODE  650
Score of total rewards: 5
========================================================
EPISODE  651
Score of total rewards: 4
========================================================
EPISODE  652
Score of total rewards: 7
========================================================
EPISODE  653
Score of total rewards: 10
========================================================
EPISODE  654
Score of total rewards: 9
========================================================
EPISODE  655
Score of total rewards: 6
========================================================
EPISODE  656
Score of total rewards: 8
========================================================
EPISODE  657
Score of total rewards: 10
========================================================
EPISODE  658
Score of total rewards: 5
========================================================
EPISODE  659
Score of total rewards: 5
========================================================
EPISODE  660
Score of total rewards: 9
========================================================
EPISODE  661
Score of total rewards: 4
========================================================
EPISODE  662
Score of total rewards: 6
```

```
========================================================
EPISODE  663
Score of total rewards: 9
========================================================
EPISODE  664
Score of total rewards: 8
========================================================
EPISODE  665
Score of total rewards: 10
========================================================
EPISODE  666
Score of total rewards: 5
========================================================
EPISODE  667
Score of total rewards: 9
========================================================
EPISODE  668
Score of total rewards: 10
========================================================
EPISODE  669
Score of total rewards: 6
========================================================
EPISODE  670
Score of total rewards: 8
========================================================
EPISODE  671
Score of total rewards: 6
========================================================
EPISODE  672
Score of total rewards: 6
========================================================
EPISODE  673
Score of total rewards: 11
========================================================
EPISODE  674
Score of total rewards: 7
========================================================
EPISODE  675
Score of total rewards: 12
========================================================
EPISODE  676
Score of total rewards: 13
========================================================
EPISODE  677
Score of total rewards: 5
========================================================
EPISODE  678
Score of total rewards: 9
========================================================
EPISODE  679
Score of total rewards: 6
========================================================
EPISODE  680
Score of total rewards: 10
========================================================
EPISODE  681
Score of total rewards: 8
```

```
========================================================
EPISODE  682
Score of total rewards: 5
========================================================
EPISODE  683
Score of total rewards: 5
========================================================
EPISODE  684
Score of total rewards: 12
========================================================
EPISODE  685
Score of total rewards: 8
========================================================
EPISODE  686
Score of total rewards: 8
========================================================
EPISODE  687
Score of total rewards: 9
========================================================
EPISODE  688
Score of total rewards: 7
========================================================
EPISODE  689
Score of total rewards: 5
========================================================
EPISODE  690
Score of total rewards: 7
========================================================
EPISODE  691
Score of total rewards: 7
========================================================
EPISODE  692
Score of total rewards: 11
========================================================
EPISODE  693
Score of total rewards: 9
========================================================
EPISODE  694
Score of total rewards: 10
========================================================
EPISODE  695
Score of total rewards: 15
========================================================
EPISODE  696
Score of total rewards: 8
========================================================
EPISODE  697
Score of total rewards: 9
========================================================
EPISODE  698
Score of total rewards: 10
========================================================
EPISODE  699
Score of total rewards: 7
========================================================
EPISODE  700
Score of total rewards: 3
```

```
========================================================
EPISODE  701
Score of total rewards: 4
========================================================
EPISODE  702
Score of total rewards: 9
========================================================
EPISODE  703
Score of total rewards: 11
========================================================
EPISODE  704
Score of total rewards: 11
========================================================
EPISODE  705
Score of total rewards: 7
========================================================
EPISODE  706
Score of total rewards: 4
========================================================
EPISODE  707
Score of total rewards: 5
========================================================
EPISODE  708
Score of total rewards: 5
========================================================
EPISODE  709
Score of total rewards: 6
========================================================
EPISODE  710
Score of total rewards: 4
========================================================
EPISODE  711
Score of total rewards: 6
========================================================
EPISODE  712
Score of total rewards: 14
========================================================
EPISODE  713
Score of total rewards: 8
========================================================
EPISODE  714
Score of total rewards: 11
========================================================
EPISODE  715
Score of total rewards: 9
========================================================
EPISODE  716
Score of total rewards: 6
========================================================
EPISODE  717
Score of total rewards: 11
========================================================
EPISODE  718
Score of total rewards: 8
========================================================
EPISODE  719
Score of total rewards: 7
```

```
========================================================
EPISODE  720
Score of total rewards: 5
========================================================
EPISODE  721
Score of total rewards: 4
========================================================
EPISODE  722
Score of total rewards: 11
========================================================
EPISODE  723
Score of total rewards: 5
========================================================
EPISODE  724
Score of total rewards: 8
========================================================
EPISODE  725
Score of total rewards: 12
========================================================
EPISODE  726
Score of total rewards: 6
========================================================
EPISODE  727
Score of total rewards: 9
========================================================
EPISODE  728
Score of total rewards: 9
========================================================
EPISODE  729
Score of total rewards: 8
========================================================
EPISODE  730
Score of total rewards: 11
========================================================
EPISODE  731
Score of total rewards: 7
========================================================
EPISODE  732
Score of total rewards: 8
========================================================
EPISODE  733
Score of total rewards: 7
========================================================
EPISODE  734
Score of total rewards: 8
========================================================
EPISODE  735
Score of total rewards: 3
========================================================
EPISODE  736
Score of total rewards: 3
========================================================
EPISODE  737
Score of total rewards: 8
========================================================
EPISODE  738
Score of total rewards: 4
```

```
========================================================
EPISODE  739
Score of total rewards: 4
========================================================
EPISODE  740
Score of total rewards: 6
========================================================
EPISODE  741
Score of total rewards: 9
========================================================
EPISODE  742
Score of total rewards: 7
========================================================
EPISODE  743
Score of total rewards: 7
========================================================
EPISODE  744
Score of total rewards: 7
========================================================
EPISODE  745
Score of total rewards: 9
========================================================
EPISODE  746
Score of total rewards: 9
========================================================
EPISODE  747
Score of total rewards: 4
========================================================
EPISODE  748
Score of total rewards: 10
========================================================
EPISODE  749
Score of total rewards: 10
========================================================
EPISODE  750
Score of total rewards: 10
========================================================
EPISODE  751
Score of total rewards: 8
========================================================
EPISODE  752
Score of total rewards: 5
========================================================
EPISODE  753
Score of total rewards: 7
========================================================
EPISODE  754
Score of total rewards: 14
========================================================
EPISODE  755
Score of total rewards: 8
========================================================
EPISODE  756
Score of total rewards: 11
========================================================
EPISODE  757
Score of total rewards: 4
```

```
========================================================
EPISODE  758
Score of total rewards: 8
========================================================
EPISODE  759
Score of total rewards: 12
========================================================
EPISODE  760
Score of total rewards: 10
========================================================
EPISODE  761
Score of total rewards: 6
========================================================
EPISODE  762
Score of total rewards: 5
========================================================
EPISODE  763
Score of total rewards: 7
========================================================
EPISODE  764
Score of total rewards: 7
========================================================
EPISODE  765
Score of total rewards: 14
========================================================
EPISODE  766
Score of total rewards: 9
========================================================
EPISODE  767
Score of total rewards: 8
========================================================
EPISODE  768
Score of total rewards: 5
========================================================
EPISODE  769
Score of total rewards: 6
========================================================
EPISODE  770
Score of total rewards: 10
========================================================
EPISODE  771
Score of total rewards: 9
========================================================
EPISODE  772
Score of total rewards: 13
========================================================
EPISODE  773
Score of total rewards: 12
========================================================
EPISODE  774
Score of total rewards: 12
========================================================
EPISODE  775
Score of total rewards: 8
========================================================
EPISODE  776
Score of total rewards: 11
```

```
========================================================
EPISODE  777
Score of total rewards: 12
========================================================
EPISODE  778
Score of total rewards: 10
========================================================
EPISODE  779
Score of total rewards: 11
========================================================
EPISODE  780
Score of total rewards: 7
========================================================
EPISODE  781
Score of total rewards: 10
========================================================
EPISODE  782
Score of total rewards: 10
========================================================
EPISODE  783
Score of total rewards: 7
========================================================
EPISODE  784
Score of total rewards: 11
========================================================
EPISODE  785
Score of total rewards: 8
========================================================
EPISODE  786
Score of total rewards: 14
========================================================
EPISODE  787
Score of total rewards: 12
========================================================
EPISODE  788
Score of total rewards: 14
========================================================
EPISODE  789
Score of total rewards: 10
========================================================
EPISODE  790
Score of total rewards: 7
========================================================
EPISODE  791
Score of total rewards: 10
========================================================
EPISODE  792
Score of total rewards: 5
========================================================
EPISODE  793
Score of total rewards: 4
========================================================
EPISODE  794
Score of total rewards: 8
========================================================
EPISODE  795
Score of total rewards: 8
```

```
========================================================
EPISODE  796
Score of total rewards: 12
========================================================
EPISODE  797
Score of total rewards: 7
========================================================
EPISODE  798
Score of total rewards: 5
========================================================
EPISODE  799
Score of total rewards: 12
========================================================
EPISODE  800
Score of total rewards: 7
========================================================
EPISODE  801
Score of total rewards: 3
========================================================
EPISODE  802
Score of total rewards: 9
========================================================
EPISODE  803
Score of total rewards: 9
========================================================
EPISODE  804
Score of total rewards: 8
========================================================
EPISODE  805
Score of total rewards: 10
========================================================
EPISODE  806
Score of total rewards: 10
========================================================
EPISODE  807
Score of total rewards: 6
========================================================
EPISODE  808
Score of total rewards: 5
========================================================
EPISODE  809
Score of total rewards: 13
========================================================
EPISODE  810
Score of total rewards: 5
========================================================
EPISODE  811
Score of total rewards: 9
========================================================
EPISODE  812
Score of total rewards: 6
========================================================
EPISODE  813
Score of total rewards: 8
========================================================
EPISODE  814
Score of total rewards: 12
```

```
=========================================================
EPISODE  815
Score of total rewards: 8
=========================================================
EPISODE  816
Score of total rewards: 10
=========================================================
EPISODE  817
Score of total rewards: 9
=========================================================
EPISODE  818
Score of total rewards: 6
=========================================================
EPISODE  819
Score of total rewards: 11
=========================================================
EPISODE  820
Score of total rewards: 9
=========================================================
EPISODE  821
Score of total rewards: 6
=========================================================
EPISODE  822
Score of total rewards: 8
=========================================================
EPISODE  823
Score of total rewards: 6
=========================================================
EPISODE  824
Score of total rewards: 13
=========================================================
EPISODE  825
Score of total rewards: 4
=========================================================
EPISODE  826
Score of total rewards: 3
=========================================================
EPISODE  827
Score of total rewards: 10
=========================================================
EPISODE  828
Score of total rewards: 8
=========================================================
EPISODE  829
Score of total rewards: 5
=========================================================
EPISODE  830
Score of total rewards: 8
=========================================================
EPISODE  831
Score of total rewards: 6
=========================================================
EPISODE  832
Score of total rewards: 12
=========================================================
EPISODE  833
Score of total rewards: 8
```

```
========================================================
EPISODE  834
Score of total rewards: 9
========================================================
EPISODE  835
Score of total rewards: 7
========================================================
EPISODE  836
Score of total rewards: 11
========================================================
EPISODE  837
Score of total rewards: 8
========================================================
EPISODE  838
Score of total rewards: 5
========================================================
EPISODE  839
Score of total rewards: 5
========================================================
EPISODE  840
Score of total rewards: 11
========================================================
EPISODE  841
Score of total rewards: 7
========================================================
EPISODE  842
Score of total rewards: 14
========================================================
EPISODE  843
Score of total rewards: 7
========================================================
EPISODE  844
Score of total rewards: 10
========================================================
EPISODE  845
Score of total rewards: 6
========================================================
EPISODE  846
Score of total rewards: 6
========================================================
EPISODE  847
Score of total rewards: 6
========================================================
EPISODE  848
Score of total rewards: 6
========================================================
EPISODE  849
Score of total rewards: 7
========================================================
EPISODE  850
Score of total rewards: 5
========================================================
EPISODE  851
Score of total rewards: 7
========================================================
EPISODE  852
Score of total rewards: 11
```

```
========================================================
EPISODE  853
Score of total rewards: 6
========================================================
EPISODE  854
Score of total rewards: 11
========================================================
EPISODE  855
Score of total rewards: 5
========================================================
EPISODE  856
Score of total rewards: 4
========================================================
EPISODE  857
Score of total rewards: 8
========================================================
EPISODE  858
Score of total rewards: 4
========================================================
EPISODE  859
Score of total rewards: 12
========================================================
EPISODE  860
Score of total rewards: 9
========================================================
EPISODE  861
Score of total rewards: 5
========================================================
EPISODE  862
Score of total rewards: 5
========================================================
EPISODE  863
Score of total rewards: 6
========================================================
EPISODE  864
Score of total rewards: 7
========================================================
EPISODE  865
Score of total rewards: 5
========================================================
EPISODE  866
Score of total rewards: 6
========================================================
EPISODE  867
Score of total rewards: 7
========================================================
EPISODE  868
Score of total rewards: 5
========================================================
EPISODE  869
Score of total rewards: 12
========================================================
EPISODE  870
Score of total rewards: 5
========================================================
EPISODE  871
Score of total rewards: 6
```

```
========================================================
EPISODE  872
Score of total rewards: 3
========================================================
EPISODE  873
Score of total rewards: 8
========================================================
EPISODE  874
Score of total rewards: 12
========================================================
EPISODE  875
Score of total rewards: 8
========================================================
EPISODE  876
Score of total rewards: 7
========================================================
EPISODE  877
Score of total rewards: 7
========================================================
EPISODE  878
Score of total rewards: 10
========================================================
EPISODE  879
Score of total rewards: 10
========================================================
EPISODE  880
Score of total rewards: 10
========================================================
EPISODE  881
Score of total rewards: 8
========================================================
EPISODE  882
Score of total rewards: 7
========================================================
EPISODE  883
Score of total rewards: 7
========================================================
EPISODE  884
Score of total rewards: 5
========================================================
EPISODE  885
Score of total rewards: 6
========================================================
EPISODE  886
Score of total rewards: 7
========================================================
EPISODE  887
Score of total rewards: 5
========================================================
EPISODE  888
Score of total rewards: 8
========================================================
EPISODE  889
Score of total rewards: 4
========================================================
EPISODE  890
Score of total rewards: 3
```

```
========================================================
EPISODE  891
Score of total rewards: 9
========================================================
EPISODE  892
Score of total rewards: 11
========================================================
EPISODE  893
Score of total rewards: 11
========================================================
EPISODE  894
Score of total rewards: 9
========================================================
EPISODE  895
Score of total rewards: 7
========================================================
EPISODE  896
Score of total rewards: 6
========================================================
EPISODE  897
Score of total rewards: 6
========================================================
EPISODE  898
Score of total rewards: 14
========================================================
EPISODE  899
Score of total rewards: 11
========================================================
EPISODE  900
Score of total rewards: 8
========================================================
EPISODE  901
Score of total rewards: 8
========================================================
EPISODE  902
Score of total rewards: 10
========================================================
EPISODE  903
Score of total rewards: 3
========================================================
EPISODE  904
Score of total rewards: 7
========================================================
EPISODE  905
Score of total rewards: 10
========================================================
EPISODE  906
Score of total rewards: 8
========================================================
EPISODE  907
Score of total rewards: 7
========================================================
EPISODE  908
Score of total rewards: 7
========================================================
EPISODE  909
Score of total rewards: 8
```

```
========================================================
EPISODE  910
Score of total rewards: 9
========================================================
EPISODE  911
Score of total rewards: 11
========================================================
EPISODE  912
Score of total rewards: 9
========================================================
EPISODE  913
Score of total rewards: 5
========================================================
EPISODE  914
Score of total rewards: 12
========================================================
EPISODE  915
Score of total rewards: 8
========================================================
EPISODE  916
Score of total rewards: 9
========================================================
EPISODE  917
Score of total rewards: 6
========================================================
EPISODE  918
Score of total rewards: 11
========================================================
EPISODE  919
Score of total rewards: 14
========================================================
EPISODE  920
Score of total rewards: 8
========================================================
EPISODE  921
Score of total rewards: 3
========================================================
EPISODE  922
Score of total rewards: 5
========================================================
EPISODE  923
Score of total rewards: 7
========================================================
EPISODE  924
Score of total rewards: 12
========================================================
EPISODE  925
Score of total rewards: 7
========================================================
EPISODE  926
Score of total rewards: 6
========================================================
EPISODE  927
Score of total rewards: 9
========================================================
EPISODE  928
Score of total rewards: 7
```

```
========================================================
EPISODE  929
Score of total rewards: 7
========================================================
EPISODE  930
Score of total rewards: 3
========================================================
EPISODE  931
Score of total rewards: 4
========================================================
EPISODE  932
Score of total rewards: 11
========================================================
EPISODE  933
Score of total rewards: 9
========================================================
EPISODE  934
Score of total rewards: 13
========================================================
EPISODE  935
Score of total rewards: 12
========================================================
EPISODE  936
Score of total rewards: 4
========================================================
EPISODE  937
Score of total rewards: 4
========================================================
EPISODE  938
Score of total rewards: 10
========================================================
EPISODE  939
Score of total rewards: 5
========================================================
EPISODE  940
Score of total rewards: 5
========================================================
EPISODE  941
Score of total rewards: 10
========================================================
EPISODE  942
Score of total rewards: 10
========================================================
EPISODE  943
Score of total rewards: 6
========================================================
EPISODE  944
Score of total rewards: 9
========================================================
EPISODE  945
Score of total rewards: 4
========================================================
EPISODE  946
Score of total rewards: 9
========================================================
EPISODE  947
Score of total rewards: 5
```

```
========================================================
EPISODE  948
Score of total rewards: 12
========================================================
EPISODE  949
Score of total rewards: 9
========================================================
EPISODE  950
Score of total rewards: 10
========================================================
EPISODE  951
Score of total rewards: 10
========================================================
EPISODE  952
Score of total rewards: 5
========================================================
EPISODE  953
Score of total rewards: 6
========================================================
EPISODE  954
Score of total rewards: 8
========================================================
EPISODE  955
Score of total rewards: 11
========================================================
EPISODE  956
Score of total rewards: 6
========================================================
EPISODE  957
Score of total rewards: 5
========================================================
EPISODE  958
Score of total rewards: 10
========================================================
EPISODE  959
Score of total rewards: 3
========================================================
EPISODE  960
Score of total rewards: 8
========================================================
EPISODE  961
Score of total rewards: 7
========================================================
EPISODE  962
Score of total rewards: 8
========================================================
EPISODE  963
Score of total rewards: 6
========================================================
EPISODE  964
Score of total rewards: 12
========================================================
EPISODE  965
Score of total rewards: 6
========================================================
EPISODE  966
Score of total rewards: 4
```

```
========================================================
EPISODE  967
Score of total rewards: 9
========================================================
EPISODE  968
Score of total rewards: 11
========================================================
EPISODE  969
Score of total rewards: 8
========================================================
EPISODE  970
Score of total rewards: 6
========================================================
EPISODE  971
Score of total rewards: 5
========================================================
EPISODE  972
Score of total rewards: 7
========================================================
EPISODE  973
Score of total rewards: 5
========================================================
EPISODE  974
Score of total rewards: 6
========================================================
EPISODE  975
Score of total rewards: 11
========================================================
EPISODE  976
Score of total rewards: 9
========================================================
EPISODE  977
Score of total rewards: 8
========================================================
EPISODE  978
Score of total rewards: 11
========================================================
EPISODE  979
Score of total rewards: 8
========================================================
EPISODE  980
Score of total rewards: 4
========================================================
EPISODE  981
Score of total rewards: 7
========================================================
EPISODE  982
Score of total rewards: 6
========================================================
EPISODE  983
Score of total rewards: 7
========================================================
EPISODE  984
Score of total rewards: 7
========================================================
EPISODE  985
Score of total rewards: 8
```

```
========================================================
EPISODE  986
Score of total rewards: 8
========================================================
EPISODE  987
Score of total rewards: 5
========================================================
EPISODE  988
Score of total rewards: 11
========================================================
EPISODE  989
Score of total rewards: 5
========================================================
EPISODE  990
Score of total rewards: 8
========================================================
EPISODE  991
Score of total rewards: 10
========================================================
EPISODE  992
Score of total rewards: 12
========================================================
EPISODE  993
Score of total rewards: 7
========================================================
EPISODE  994
Score of total rewards: 9
========================================================
EPISODE  995
Score of total rewards: 8
========================================================
EPISODE  996
Score of total rewards: 13
========================================================
EPISODE  997
Score of total rewards: 6
========================================================
EPISODE  998
Score of total rewards: 10
========================================================
EPISODE  999
Score of total rewards: 7
Average score over time: 7.946
```

In [79]:
```python
# only to have a look of the test play result
my_q_agent_fine_tuned.test(test_episodes=5, display=True)
```

Taxi-v3 | Episode: 4 | Step: 13
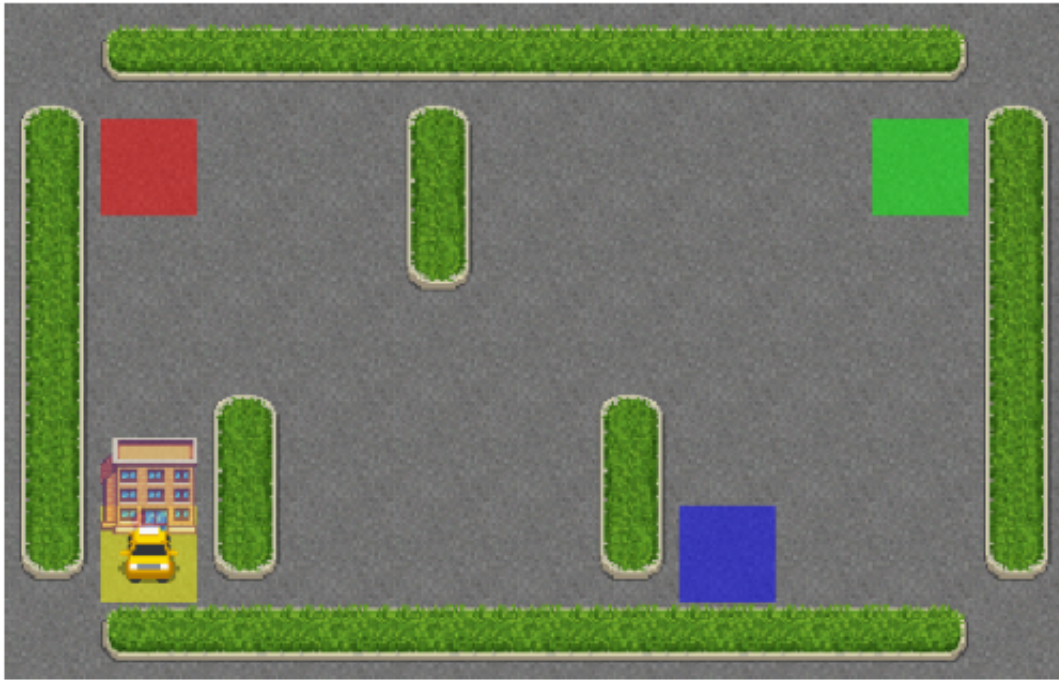

```
Score of total rewards: 8
Average score over time: 10.0
```

Out[79]: [11, 6, 10, 15, 8]

Taxi-v3 | Episode: 4 | Step: 13
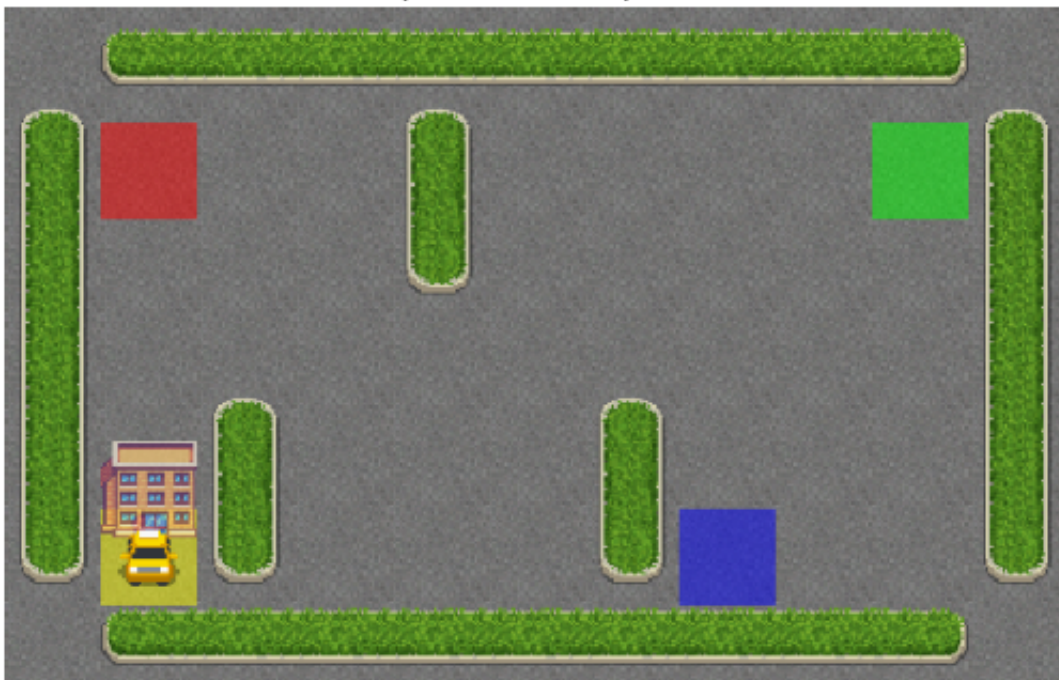
In [80]:
```
# save to GIF
my_q_agent_fine_tuned.test_to_gif(filename="taxi_q_fine_tuned",episodes=5
```

```
episodes: 0, acton: 3, state: 182
episodes: 0, acton: 3, state: 162
episodes: 0, acton: 0, state: 142
episodes: 0, acton: 3, state: 242
episodes: 0, acton: 3, state: 222
episodes: 0, acton: 1, state: 202
episodes: 0, acton: 1, state: 102
episodes: 0, acton: 4, state: 2
```

```
episodes: 0, acton: 0, state: 18
episodes: 0, acton: 0, state: 118
episodes: 0, acton: 0, state: 218
episodes: 0, acton: 0, state: 318
episodes: 0, acton: 5, state: 418
episodes: 1, acton: 3, state: 81
episodes: 1, acton: 0, state: 61
episodes: 1, acton: 0, state: 161
episodes: 1, acton: 3, state: 261
episodes: 1, acton: 3, state: 241
episodes: 1, acton: 1, state: 221
episodes: 1, acton: 1, state: 121
episodes: 1, acton: 3, state: 21
episodes: 1, acton: 4, state: 1
episodes: 1, acton: 2, state: 17
episodes: 1, acton: 0, state: 37
episodes: 1, acton: 0, state: 137
episodes: 1, acton: 2, state: 237
episodes: 1, acton: 1, state: 257
episodes: 1, acton: 2, state: 157
episodes: 1, acton: 2, state: 177
episodes: 1, acton: 1, state: 197
episodes: 1, acton: 5, state: 97
episodes: 2, acton: 0, state: 169
episodes: 2, acton: 3, state: 269
episodes: 2, acton: 3, state: 249
episodes: 2, acton: 3, state: 229
episodes: 2, acton: 0, state: 209
episodes: 2, acton: 0, state: 309
episodes: 2, acton: 4, state: 409
episodes: 2, acton: 1, state: 417
episodes: 2, acton: 1, state: 317
episodes: 2, acton: 2, state: 217
episodes: 2, acton: 2, state: 237
episodes: 2, acton: 1, state: 257
episodes: 2, acton: 2, state: 157
episodes: 2, acton: 2, state: 177
episodes: 2, acton: 1, state: 197
episodes: 2, acton: 5, state: 97
episodes: 3, acton: 0, state: 114
episodes: 3, acton: 2, state: 214
episodes: 3, acton: 2, state: 234
episodes: 3, acton: 2, state: 254
episodes: 3, acton: 0, state: 274
episodes: 3, acton: 0, state: 374
episodes: 3, acton: 4, state: 474
episodes: 3, acton: 1, state: 478
episodes: 3, acton: 1, state: 378
episodes: 3, acton: 3, state: 278
episodes: 3, acton: 3, state: 258
episodes: 3, acton: 3, state: 238
episodes: 3, acton: 0, state: 218
episodes: 3, acton: 0, state: 318
episodes: 3, acton: 5, state: 418
episodes: 4, acton: 1, state: 462
episodes: 4, acton: 1, state: 362
episodes: 4, acton: 3, state: 262
```

```
episodes: 4, acton: 3, state: 242
episodes: 4, acton: 3, state: 222
episodes: 4, acton: 1, state: 202
episodes: 4, acton: 1, state: 102
episodes: 4, acton: 4, state: 2
episodes: 4, acton: 0, state: 18
episodes: 4, acton: 0, state: 118
episodes: 4, acton: 0, state: 218
episodes: 4, acton: 0, state: 318
episodes: 4, acton: 5, state: 418
Saving gif...Done!
```

In [81]:
```python
# display the captured GIF
display.Image(data=open("./gifs/taxi_q_fine_tuned.gif",'rb').read(), form
# taxi_q_fine_tuned
```

Out[81]: `<IPython.core.display.Image object>`

If the above display function is not working, here is the previously prepared taxi_q_fine_tuned.gif

DQN Agent with Fine Tuned Parameters

In [84]:
```python
%%time
my_dqn_agent_fine_tuned_avg_reward_lst = my_dqn_agent_fine_tuned.test(tes
```

```
Trained Policy Network Model from the .pt file:
State 000, Action W, Q-values: [+3.33 +3.51 +3.39 +3.85 -4.99 -2.85]
State 001, Action P, Q-values: [-0.23 +0.28 +0.40 +0.70 +1.66 -5.24]
State 002, Action P, Q-values: [+5.54 +6.02 +5.56 +6.42 +7.76 -3.03]
State 003, Action P, Q-values: [+1.22 +1.21 +1.05 +1.86 +2.92 -6.50]
State 004, Action S, Q-values: [-4.93 -5.18 -5.00 -5.16 -14.92 -18.07]
State 005, Action W, Q-values: [+3.16 +3.31 +3.17 +3.59 -4.93 -3.14]
State 006, Action E, Q-values: [-5.20 -5.24 -5.03 -5.31 -14.88 -17.98]
```

```
State 007, Action S, Q-values: [−3.18 −3.82 −3.67 −3.68 −11.11 −12.79]
State 008, Action S, Q-values: [+1.66 +0.91 −0.14 +1.04 −7.84 −9.71]
State 009, Action S, Q-values: [−2.33 −2.77 −3.17 −2.61 −11.63 −12.20]
State 010, Action W, Q-values: [+3.49 +3.45 +3.38 +3.77 −4.79 −3.28]
State 011, Action S, Q-values: [−1.53 −2.29 −2.87 −2.23 −11.06 −12.18]
State 012, Action S, Q-values: [−3.86 −4.52 −4.24 −4.11 −12.39 −13.99]
State 013, Action S, Q-values: [−2.36 −2.99 −2.47 −2.67 −11.01 −12.98]
State 014, Action S, Q-values: [−3.68 −4.08 −3.73 −3.92 −12.44 −13.42]
State 015, Action W, Q-values: [+3.41 +3.23 +3.19 +3.60 −5.15 −2.81]
State 016, Action D, Q-values: [+13.78 +15.27 +14.37 +16.83 +7.57 +20.01]
State 017, Action E, Q-values: [+3.00 +1.74 +3.00 +2.20 −6.17 +0.29]
State 018, Action S, Q-values: [+9.76 +7.56 +7.25 +8.33 +0.27 +3.41]
State 019, Action E, Q-values: [+4.33 +3.25 +4.35 +3.58 −5.06 +0.50]
State 020, Action W, Q-values: [+3.15 +3.20 +3.09 +3.57 −5.05 −3.48]
State 021, Action W, Q-values: [−0.69 −0.13 −0.04 +0.48 −6.71 −8.06]
State 022, Action W, Q-values: [+3.22 +4.52 +4.18 +5.98 −3.59 −3.49]
State 023, Action W, Q-values: [+0.29 +1.03 +0.67 +1.62 −7.00 −7.05]
State 024, Action S, Q-values: [−4.40 −5.06 −5.20 −4.76 −14.13 −17.15]
State 025, Action W, Q-values: [+3.41 +3.45 +3.46 +3.74 −5.16 −3.01]
State 026, Action S, Q-values: [−4.46 −5.00 −4.79 −4.84 −13.77 −16.70]
State 027, Action S, Q-values: [−2.31 −2.92 −2.84 −3.00 −11.31 −13.31]
State 028, Action S, Q-values: [+0.51 −0.11 −0.61 +0.21 −9.43 −10.25]
State 029, Action S, Q-values: [−3.07 −3.29 −4.00 −3.07 −11.50 −13.37]
State 030, Action W, Q-values: [+3.32 +3.49 +3.37 +3.80 −4.76 −2.94]
State 031, Action W, Q-values: [−2.39 −2.47 −3.01 −2.30 −10.75 −12.46]
State 032, Action S, Q-values: [−3.09 −3.81 −3.73 −3.62 −11.99 −12.61]
State 033, Action S, Q-values: [−1.52 −2.16 −2.00 −2.37 −10.90 −12.92]
State 034, Action S, Q-values: [−3.17 −3.95 −4.15 −3.84 −11.63 −13.55]
State 035, Action W, Q-values: [+3.32 +3.35 +3.29 +3.65 −5.01 −3.01]
State 036, Action W, Q-values: [+13.05 +14.61 +14.53 +16.96 +3.87 +6.03]
State 037, Action S, Q-values: [+4.42 +3.54 +3.95 +2.67 −5.45 −7.21]
State 038, Action S, Q-values: [+7.47 +4.93 +5.70 +7.43 −3.20 −2.91]
State 039, Action S, Q-values: [+5.95 +4.46 +5.19 +4.91 −3.23 −2.89]
State 040, Action W, Q-values: [+3.40 +3.45 +3.36 +3.81 −4.82 −3.00]
State 041, Action S, Q-values: [−3.79 −4.22 −4.68 −4.10 −12.59 −14.59]
State 042, Action S, Q-values: [−0.47 −0.99 −1.45 −0.98 −10.18 −10.86]
State 043, Action S, Q-values: [−3.11 −3.99 −4.48 −3.72 −11.76 −13.48]
State 044, Action E, Q-values: [−1.80 −1.38 −0.53 −1.13 −9.51 −10.27]
State 045, Action W, Q-values: [+3.30 +3.41 +3.25 +3.73 −4.87 −3.08]
State 046, Action E, Q-values: [−0.89 −1.30 −0.53 −1.16 −9.90 −10.35]
State 047, Action E, Q-values: [+1.22 +2.21 +2.93 +2.56 −5.97 −5.35]
State 048, Action S, Q-values: [−0.58 −1.20 −1.57 −1.13 −10.65 −11.67]
State 049, Action S, Q-values: [−3.80 −4.17 −4.53 −4.07 −12.42 −14.36]
State 050, Action W, Q-values: [+3.38 +3.34 +3.24 +3.76 −4.98 −3.17]
State 051, Action S, Q-values: [−3.17 −3.59 −4.17 −3.39 −11.70 −12.69]
State 052, Action S, Q-values: [−2.28 −2.58 −2.34 −2.76 −10.34 −11.37]
State 053, Action S, Q-values: [−0.60 −1.41 −0.65 −1.08 −9.89 −11.34]
State 054, Action S, Q-values: [−2.42 −2.95 −2.43 −2.76 −10.74 −12.58]
State 055, Action W, Q-values: [+3.26 +3.28 +3.16 +3.62 −5.18 −3.39]
State 056, Action S, Q-values: [+6.05 +4.56 +4.13 +5.75 −3.55 −3.06]
State 057, Action E, Q-values: [+10.95 +12.65 +14.29 +11.94 +3.96 +4.41]
State 058, Action S, Q-values: [+5.99 +4.10 +4.16 +5.09 −3.85 −3.19]
State 059, Action S, Q-values: [+6.27 +4.32 +5.53 +5.83 −3.86 −3.52]
State 060, Action W, Q-values: [+3.02 +3.18 +3.11 +3.46 −5.10 −3.42]
State 061, Action W, Q-values: [−4.47 −4.68 −5.52 −4.44 −13.88 −16.70]
State 062, Action S, Q-values: [−1.44 −2.22 −2.82 −1.83 −10.30 −11.53]
State 063, Action S, Q-values: [−3.71 −4.28 −4.41 −4.01 −12.71 −14.44]
```

```
State 064, Action E, Q-values: [+0.05 -0.22 +0.46 -0.46 -8.03 -8.75]
State 065, Action W, Q-values: [+3.22 +3.16 +2.96 +3.49 -5.33 -3.67]
State 066, Action E, Q-values: [-0.51 -0.45 +0.42 -0.27 -7.07 -8.92]
State 067, Action E, Q-values: [+2.18 +3.34 +4.44 +2.80 -3.90 -3.49]
State 068, Action S, Q-values: [-1.53 -2.13 -2.65 -1.78 -10.22 -11.56]
State 069, Action W, Q-values: [-4.47 -4.69 -5.25 -4.38 -13.95 -16.32]
State 070, Action W, Q-values: [+3.34 +3.39 +3.21 +3.73 -5.14 -2.91]
State 071, Action W, Q-values: [-3.79 -4.30 -4.82 -3.75 -12.74 -14.78]
State 072, Action S, Q-values: [-1.48 -2.33 -2.70 -2.42 -11.11 -12.38]
State 073, Action S, Q-values: [+0.49 -0.18 -0.49 -0.69 -8.89 -10.15]
State 074, Action S, Q-values: [-1.53 -2.20 -2.26 -2.36 -9.89 -11.53]
State 075, Action W, Q-values: [+3.54 +3.51 +3.46 +3.83 -5.15 -3.07]
State 076, Action S, Q-values: [+4.35 +3.12 +2.43 +4.34 -4.75 -4.60]
State 077, Action E, Q-values: [+12.85 +15.35 +17.03 +13.53 +3.62 +5.89]
State 078, Action S, Q-values: [+4.40 +3.03 +2.48 +4.37 -4.83 -4.63]
State 079, Action S, Q-values: [+9.74 +8.00 +7.53 +7.26 -0.24 -0.11]
State 080, Action W, Q-values: [+3.34 +3.43 +3.22 +3.83 -5.09 -2.99]
State 081, Action S, Q-values: [-4.99 -5.51 -5.51 -5.05 -15.40 -18.64]
State 082, Action S, Q-values: [-2.26 -2.47 -3.05 -2.36 -10.08 -11.68]
State 083, Action W, Q-values: [-4.44 -4.92 -5.22 -4.42 -13.66 -16.73]
State 084, Action P, Q-values: [-0.34 +0.16 +0.32 +0.32 +1.67 -5.22]
State 085, Action W, Q-values: [+3.39 +3.41 +3.25 +3.72 -4.69 -3.03]
State 086, Action P, Q-values: [-0.34 +0.17 +0.51 +0.35 +1.60 -5.83]
State 087, Action P, Q-values: [+3.94 +4.19 +4.15 +4.27 +6.04 -2.71]
State 088, Action S, Q-values: [-2.44 -2.83 -3.15 -2.49 -10.81 -12.72]
State 089, Action W, Q-values: [-5.05 -5.48 -5.74 -4.88 -15.40 -18.46]
State 090, Action W, Q-values: [+3.28 +3.48 +3.37 +3.76 -4.94 -2.92]
State 091, Action W, Q-values: [-4.42 -4.53 -4.94 -4.37 -13.71 -16.23]
State 092, Action W, Q-values: [-2.36 -2.83 -2.94 -2.32 -10.96 -12.03]
State 093, Action S, Q-values: [-0.50 -1.02 -1.55 -0.78 -10.13 -11.43]
State 094, Action S, Q-values: [-2.11 -3.02 -2.83 -2.20 -11.24 -12.42]
State 095, Action W, Q-values: [+3.30 +3.34 +3.33 +3.76 -5.00 -3.04]
State 096, Action S, Q-values: [+2.93 +1.78 +2.04 +2.90 -6.95 -1.30]
State 097, Action D, Q-values: [+13.47 +15.36 +15.07 +15.13 +7.67 +20.01]
State 098, Action S, Q-values: [+3.01 +1.59 +2.06 +3.00 -6.40 -1.25]
State 099, Action S, Q-values: [+7.75 +5.89 +6.42 +7.70 -2.21 +3.07]
State 100, Action W, Q-values: [+3.38 +3.46 +3.36 +3.75 -5.22 -2.94]
State 101, Action N, Q-values: [-0.96 +0.45 -1.32 +0.09 -7.51 -7.72]
State 102, Action N, Q-values: [+3.10 +5.97 +4.97 +5.40 -3.41 -3.34]
State 103, Action N, Q-values: [-0.21 +1.60 -0.23 +0.97 -7.58 -7.98]
State 104, Action E, Q-values: [-4.45 -5.32 -4.40 -4.83 -13.54 -15.93]
State 105, Action W, Q-values: [+3.23 +3.46 +3.39 +3.76 -4.72 -2.95]
State 106, Action E, Q-values: [-4.44 -4.76 -4.33 -4.51 -12.99 -15.21]
State 107, Action S, Q-values: [-2.44 -2.92 -2.52 -2.73 -10.53 -11.29]
State 108, Action S, Q-values: [+2.99 +0.67 +1.38 +2.21 -6.77 -7.31]
State 109, Action S, Q-values: [-1.54 -2.66 -2.60 -2.00 -10.95 -11.54]
State 110, Action W, Q-values: [+3.29 +3.33 +3.19 +3.70 -5.24 -3.15]
State 111, Action S, Q-values: [-0.61 -2.15 -1.94 -1.24 -10.33 -10.65]
State 112, Action E, Q-values: [-3.21 -4.12 -3.12 -3.42 -11.37 -12.18]
State 113, Action S, Q-values: [-1.50 -2.83 -1.53 -2.05 -10.40 -11.05]
State 114, Action S, Q-values: [-3.17 -3.64 -3.29 -3.41 -11.99 -12.66]
State 115, Action W, Q-values: [+3.54 +3.66 +3.53 +4.02 -5.11 -2.66]
State 116, Action N, Q-values: [+12.99 +17.04 +15.56 +14.74 +4.62 +5.27]
State 117, Action S, Q-values: [+4.43 +2.29 +4.41 +3.79 -4.28 -4.48]
State 118, Action S, Q-values: [+11.93 +9.19 +9.36 +9.76 +3.25 +2.11]
State 119, Action S, Q-values: [+5.91 +3.65 +5.88 +5.03 -2.55 -2.82]
State 120, Action W, Q-values: [+3.15 +3.23 +3.12 +3.60 -4.74 -3.11]
```

```
State 121, Action N, Q-values: [−2.05 −0.59 −1.33 −0.72 −8.55 −9.34]
State 122, Action W, Q-values: [+2.05 +4.31 +3.27 +4.34 −4.96 −4.58]
State 123, Action N, Q-values: [−1.24 +0.42 −0.82 +0.18 −8.48 −8.55]
State 124, Action S, Q-values: [−3.74 −4.63 −4.60 −4.08 −12.39 −14.70]
State 125, Action W, Q-values: [+3.39 +3.39 +3.31 +3.77 −5.14 −3.29]
State 126, Action S, Q-values: [−3.80 −4.56 −4.44 −4.06 −13.03 −14.42]
State 127, Action S, Q-values: [−1.50 −2.29 −1.94 −2.26 −9.84 −10.38]
State 128, Action S, Q-values: [+1.65 −0.04 +0.50 +1.53 −7.68 −8.19]
State 129, Action S, Q-values: [−2.23 −2.78 −3.05 −2.26 −10.67 −12.06]
State 130, Action W, Q-values: [+3.35 +3.32 +3.24 +3.70 −4.87 −3.04]
State 131, Action S, Q-values: [−1.47 −1.90 −2.23 −1.53 −10.62 −11.20]
State 132, Action S, Q-values: [−2.35 −3.32 −3.09 −3.03 −10.92 −10.63]
State 133, Action S, Q-values: [−0.63 −1.70 −1.33 −1.82 −9.70 −10.53]
State 134, Action S, Q-values: [−2.41 −3.41 −2.94 −2.92 −11.12 −11.31]
State 135, Action W, Q-values: [+3.38 +3.43 +3.27 +3.73 −4.76 −2.93]
State 136, Action N, Q-values: [+10.93 +14.42 +13.17 +13.77 +3.65 +4.54]
State 137, Action S, Q-values: [+6.11 +3.35 +4.78 +4.64 −2.78 −3.50]
State 138, Action S, Q-values: [+9.76 +6.06 +7.58 +9.24 −1.35 −1.55]
State 139, Action S, Q-values: [+7.74 +4.75 +6.10 +6.74 −1.86 −2.14]
State 140, Action W, Q-values: [+3.59 +3.52 +3.35 +3.90 −4.58 −3.00]
State 141, Action S, Q-values: [−3.07 −3.91 −3.70 −3.28 −10.86 −11.83]
State 142, Action S, Q-values: [+0.46 −1.19 −1.22 −0.30 −8.25 −9.67]
State 143, Action S, Q-values: [−2.35 −3.22 −3.28 −2.69 −11.33 −11.64]
State 144, Action E, Q-values: [−2.37 −1.50 −1.49 −2.05 −9.50 −11.26]
State 145, Action W, Q-values: [+3.38 +3.37 +3.30 +3.68 −5.07 −3.09]
State 146, Action N, Q-values: [−2.86 −1.52 −1.64 −2.15 −8.92 −10.01]
State 147, Action E, Q-values: [−0.41 +1.67 +1.69 +0.62 −7.05 −7.33]
State 148, Action S, Q-values: [+0.48 −0.62 −1.32 +0.09 −8.91 −9.66]
State 149, Action S, Q-values: [−3.08 −3.33 −3.80 −3.35 −10.81 −11.56]
State 150, Action W, Q-values: [+3.39 +3.48 +3.42 +3.79 −5.01 −3.05]
State 151, Action S, Q-values: [−2.41 −3.12 −3.40 −2.71 −10.14 −10.09]
State 152, Action S, Q-values: [−1.48 −2.24 −1.51 −1.93 −10.07 −11.20]
State 153, Action S, Q-values: [+0.55 −0.77 −0.24 −0.51 −8.79 −9.62]
State 154, Action S, Q-values: [−1.55 −2.39 −1.56 −2.03 −9.81 −10.44]
State 155, Action W, Q-values: [+3.59 +3.45 +3.38 +3.82 −4.88 −2.97]
State 156, Action S, Q-values: [+7.84 +4.65 +5.01 +7.11 −2.60 −3.27]
State 157, Action E, Q-values: [+8.34 +11.89 +11.93 +10.17 +2.41 +2.00]
State 158, Action S, Q-values: [+7.76 +4.79 +5.31 +7.13 −1.26 −2.00]
State 159, Action S, Q-values: [+9.36 +5.49 +8.00 +8.77 −1.56 −1.24]
State 160, Action W, Q-values: [+3.45 +3.64 +3.55 +3.86 −4.95 −3.05]
State 161, Action W, Q-values: [−3.81 −4.39 −4.37 −3.75 −12.95 −14.12]
State 162, Action W, Q-values: [−0.60 −2.39 −1.87 −0.52 −9.23 −10.96]
State 163, Action W, Q-values: [−3.17 −4.30 −4.30 −3.06 −11.61 −13.11]
State 164, Action N, Q-values: [−1.73 −0.56 −0.56 −1.64 −8.69 −9.90]
State 165, Action W, Q-values: [+3.57 +3.45 +3.44 +3.76 −4.96 −2.83]
State 166, Action E, Q-values: [−2.09 −0.59 −0.58 −1.51 −8.65 −9.76]
State 167, Action E, Q-values: [+0.30 +2.85 +2.91 +1.01 −6.25 −6.55]
State 168, Action W, Q-values: [−0.81 −1.77 −1.89 −0.55 −9.57 −10.77]
State 169, Action W, Q-values: [−4.19 −4.14 −5.16 −3.80 −13.33 −14.99]
State 170, Action W, Q-values: [+3.35 +3.37 +3.21 +3.71 −4.94 −3.27]
State 171, Action W, Q-values: [−3.18 −3.72 −4.60 −3.10 −12.01 −13.62]
State 172, Action S, Q-values: [−0.59 −1.92 −2.15 −1.56 −10.30 −11.27]
State 173, Action S, Q-values: [+1.57 +0.71 +0.53 −0.18 −8.10 −8.92]
State 174, Action S, Q-values: [−0.59 −1.64 −1.80 −1.80 −9.55 −10.52]
State 175, Action W, Q-values: [+3.41 +3.44 +3.30 +3.81 −4.70 −2.96]
State 176, Action W, Q-values: [+6.00 +3.72 +3.81 +6.04 −3.82 −3.21]
State 177, Action N, Q-values: [+10.76 +14.26 +14.24 +11.37 +3.62 +4.30]
```

```
State 178, Action W, Q-values: [+5.84 +3.12 +3.73 +5.97 −3.36 −3.41]
State 179, Action S, Q-values: [+11.92 +8.37 +8.46 +8.90 +1.61 +1.19]
State 180, Action W, Q-values: [+3.21 +3.45 +3.35 +3.77 −4.87 −3.16]
State 181, Action W, Q-values: [−4.41 −4.70 −4.97 −4.32 −13.42 −15.85]
State 182, Action W, Q-values: [−1.41 −2.59 −2.20 −1.41 −10.69 −11.29]
State 183, Action W, Q-values: [−4.06 −4.46 −4.61 −3.76 −13.39 −14.88]
State 184, Action N, Q-values: [−1.28 +0.46 −0.40 −0.80 −8.66 −9.32]
State 185, Action W, Q-values: [+3.38 +3.57 +3.35 +3.86 −5.02 −2.86]
State 186, Action N, Q-values: [−1.02 +0.42 −0.21 −0.80 −7.59 −6.93]
State 187, Action N, Q-values: [+2.06 +4.38 +2.99 +3.43 −5.55 −4.95]
State 188, Action S, Q-values: [−1.45 −2.11 −2.29 −1.57 −10.43 −11.44]
State 189, Action W, Q-values: [−4.47 −5.06 −5.07 −4.38 −13.28 −16.36]
State 190, Action W, Q-values: [+3.45 +3.36 +3.19 +3.81 −5.05 −3.02]
State 191, Action W, Q-values: [−3.86 −4.06 −4.77 −3.76 −12.30 −15.07]
State 192, Action W, Q-values: [−1.59 −2.25 −2.42 −1.52 −10.38 −11.85]
State 193, Action S, Q-values: [+0.46 −0.93 −0.99 −0.11 −9.68 −11.06]
State 194, Action W, Q-values: [−1.52 −2.37 −2.36 −1.51 −10.53 −11.64]
State 195, Action W, Q-values: [+3.30 +3.25 +3.14 +3.52 −5.09 −3.16]
State 196, Action W, Q-values: [+4.39 +2.33 +3.09 +4.41 −4.63 −4.70]
State 197, Action N, Q-values: [+12.79 +17.03 +16.16 +13.65 +5.18 +5.62]
State 198, Action W, Q-values: [+4.42 +2.09 +2.96 +4.45 −4.52 −5.23]
State 199, Action S, Q-values: [+9.71 +5.85 +7.35 +9.68 −1.19 −1.28]
State 200, Action W, Q-values: [+3.51 +3.44 +3.29 +3.73 −5.05 −3.27]
State 201, Action N, Q-values: [−2.12 −0.56 −2.26 −1.26 −10.69 −11.12]
State 202, Action N, Q-values: [+2.05 +4.31 +2.84 +3.66 −5.13 −4.71]
State 203, Action N, Q-values: [−1.24 +0.43 −1.21 −0.21 −8.94 −9.66]
State 204, Action E, Q-values: [−4.85 −4.72 −3.78 −4.15 −12.32 −14.77]
State 205, Action W, Q-values: [+3.43 +3.38 +3.25 +3.60 −5.17 −3.08]
State 206, Action E, Q-values: [−4.75 −4.25 −3.84 −4.09 −11.99 −14.98]
State 207, Action E, Q-values: [−2.58 −2.49 −1.56 −2.03 −10.24 −11.54]
State 208, Action S, Q-values: [+4.35 +1.41 +1.74 +3.52 −5.51 −6.30]
State 209, Action S, Q-values: [−0.66 −2.05 −1.80 −1.48 −9.62 −9.81]
State 210, Action W, Q-values: [+3.32 +3.32 +3.19 +3.66 −5.16 −3.55]
State 211, Action S, Q-values: [+0.49 −0.53 −0.55 −0.38 −8.17 −8.78]
State 212, Action E, Q-values: [−3.37 −3.02 −2.36 −2.69 −10.51 −12.51]
State 213, Action E, Q-values: [−1.58 −1.85 −0.55 −1.19 −10.32 −11.32]
State 214, Action E, Q-values: [−3.07 −3.46 −2.41 −2.75 −10.77 −12.64]
State 215, Action W, Q-values: [+3.49 +3.50 +3.45 +3.83 −4.88 −3.27]
State 216, Action N, Q-values: [+10.64 +14.33 +12.65 +12.27 +3.12 +4.70]
State 217, Action E, Q-values: [+2.86 +4.08 +6.10 +4.66 −3.28 −2.88]
State 218, Action S, Q-values: [+14.30 +9.79 +10.76 +12.25 +3.29 +3.38]
State 219, Action E, Q-values: [+5.20 +5.07 +7.83 +6.54 −1.47 −1.37]
State 220, Action W, Q-values: [+3.42 +3.38 +3.25 +3.79 −5.03 −2.74]
State 221, Action W, Q-values: [−3.07 −1.45 −2.61 −1.41 −9.75 −10.33]
State 222, Action N, Q-values: [+0.71 +2.90 +1.35 +2.86 −6.34 −5.62]
State 223, Action W, Q-values: [−2.63 −0.59 −2.24 −0.56 −9.95 −10.05]
State 224, Action E, Q-values: [−4.17 −3.39 −3.10 −3.67 −11.04 −12.95]
State 225, Action W, Q-values: [+3.28 +3.41 +3.22 +3.71 −5.17 −2.88]
State 226, Action E, Q-values: [−3.82 −3.39 −3.09 −3.65 −10.84 −12.09]
State 227, Action E, Q-values: [−1.58 −1.22 −0.57 −1.33 −8.94 −10.18]
State 228, Action W, Q-values: [+1.62 +0.99 +0.81 +2.90 −5.90 −6.57]
State 229, Action W, Q-values: [−2.55 −2.91 −3.06 −1.54 −9.87 −11.26]
State 230, Action W, Q-values: [+3.40 +3.50 +3.47 +3.80 −5.30 −2.83]
State 231, Action W, Q-values: [−1.76 −1.38 −2.07 −0.53 −8.98 −10.10]
State 232, Action E, Q-values: [−2.84 −2.46 −1.56 −2.34 −10.56 −11.70]
State 233, Action E, Q-values: [+0.03 −0.91 +0.49 −0.65 −8.35 −9.93]
State 234, Action E, Q-values: [−2.46 −2.64 −1.53 −2.31 −10.32 −11.38]
```

```
State 235, Action W, Q-values: [+3.36 +3.40 +3.33 +3.69 -5.11 -3.07]
State 236, Action N, Q-values: [+7.85 +11.93 +9.99 +11.90 +1.14 +1.58]
State 237, Action E, Q-values: [+4.85 +6.73 +7.92 +5.71 -1.43 -1.27]
State 238, Action W, Q-values: [+8.37 +8.53 +8.75 +11.84 -0.45 +0.82]
State 239, Action E, Q-values: [+6.52 +6.90 +9.78 +7.08 -0.95 -1.40]
State 240, Action W, Q-values: [+3.44 +3.49 +3.37 +3.89 -5.38 -3.09]
State 241, Action W, Q-values: [-3.54 -2.89 -3.60 -2.24 -9.64 -11.07]
State 242, Action W, Q-values: [+0.32 -0.07 -0.52 +1.63 -6.82 -7.65]
State 243, Action W, Q-values: [-3.11 -2.62 -3.04 -1.49 -10.60 -11.47]
State 244, Action E, Q-values: [-3.14 -2.33 -2.31 -3.19 -9.84 -11.67]
State 245, Action W, Q-values: [+3.70 +3.46 +3.41 +3.87 -5.16 -3.14]
State 246, Action E, Q-values: [-3.43 -2.38 -2.34 -3.28 -9.54 -12.10]
State 247, Action E, Q-values: [-1.30 +0.45 +0.49 -0.81 -7.89 -8.78]
State 248, Action W, Q-values: [+0.30 -0.09 +0.07 +1.64 -6.95 -7.90]
State 249, Action W, Q-values: [-3.97 -3.55 -3.48 -2.27 -11.09 -10.66]
State 250, Action W, Q-values: [+3.43 +3.46 +3.32 +3.77 -5.05 -2.83]
State 251, Action W, Q-values: [-2.93 -2.81 -3.05 -1.50 -10.70 -11.04]
State 252, Action E, Q-values: [-1.88 -1.71 -0.61 -1.37 -9.49 -10.46]
State 253, Action E, Q-values: [+0.04 +0.25 +1.66 -0.00 -5.66 -7.17]
State 254, Action E, Q-values: [-1.45 -1.54 -0.57 -1.43 -9.09 -10.34]
State 255, Action W, Q-values: [+3.39 +3.49 +3.36 +3.77 -4.93 -2.89]
State 256, Action W, Q-values: [+6.33 +6.94 +6.81 +9.74 -1.43 -0.73]
State 257, Action N, Q-values: [+6.94 +9.74 +9.70 +6.89 +0.26 +0.86]
State 258, Action W, Q-values: [+6.54 +6.51 +6.54 +9.74 -1.06 -1.05]
State 259, Action E, Q-values: [+9.80 +9.71 +11.89 +9.12 +2.26 +1.88]
State 260, Action W, Q-values: [+3.31 +3.46 +3.34 +3.74 -4.71 -3.14]
State 261, Action W, Q-values: [-4.20 -3.94 -4.40 -3.06 -12.14 -12.93]
State 262, Action W, Q-values: [-1.53 -0.26 -1.33 +0.44 -8.16 -9.42]
State 263, Action W, Q-values: [-3.85 -3.60 -3.93 -2.39 -11.06 -11.59]
State 264, Action N, Q-values: [-2.53 -1.50 -1.51 -2.16 -8.74 -10.05]
State 265, Action W, Q-values: [+3.16 +3.19 +3.14 +3.53 -5.33 -3.35]
State 266, Action N, Q-values: [-2.59 -1.50 -1.51 -2.48 -10.05 -11.05]
State 267, Action N, Q-values: [-0.40 +1.59 +1.57 +0.11 -7.63 -8.16]
State 268, Action W, Q-values: [-1.16 -0.95 -1.74 +0.48 -8.69 -10.79]
State 269, Action W, Q-values: [-4.39 -4.46 -4.48 -3.08 -12.23 -13.55]
State 270, Action W, Q-values: [+3.31 +3.31 +3.18 +3.67 -5.02 -3.24]
State 271, Action W, Q-values: [-3.53 -3.83 -3.81 -2.38 -11.11 -11.18]
State 272, Action S, Q-values: [+0.43 -0.75 -1.04 -0.99 -8.58 -9.68]
State 273, Action S, Q-values: [+2.92 +0.77 +1.10 +0.74 -6.29 -8.85]
State 274, Action S, Q-values: [+0.44 -0.75 -1.15 -0.81 -7.65 -8.58]
State 275, Action W, Q-values: [+3.50 +3.50 +3.43 +3.87 -5.10 -3.03]
State 276, Action W, Q-values: [+3.75 +5.48 +5.18 +7.81 -2.19 -1.78]
State 277, Action N, Q-values: [+8.60 +11.89 +11.83 +9.13 +2.77 +1.95]
State 278, Action W, Q-values: [+4.79 +4.80 +5.26 +7.72 -1.63 -1.88]
State 279, Action S, Q-values: [+14.29 +9.02 +10.62 +11.38 +3.85 +4.10]
State 280, Action W, Q-values: [+3.16 +3.27 +3.13 +3.61 -5.37 -3.49]
State 281, Action W, Q-values: [-4.43 -4.24 -4.41 -3.75 -12.71 -13.77]
State 282, Action W, Q-values: [-2.05 -1.87 -1.53 -0.50 -11.11 -11.93]
State 283, Action W, Q-values: [-3.87 -4.40 -4.44 -3.10 -12.11 -13.99]
State 284, Action N, Q-values: [-1.85 -0.52 -1.42 -1.67 -10.41 -10.38]
State 285, Action W, Q-values: [+3.27 +3.35 +3.24 +3.62 -4.87 -3.38]
State 286, Action N, Q-values: [-2.26 -0.62 -1.31 -1.78 -9.82 -10.17]
State 287, Action N, Q-values: [+0.26 +2.84 +1.69 +1.39 -7.31 -7.02]
State 288, Action W, Q-values: [-1.74 -1.84 -1.65 -0.55 -10.02 -11.67]
State 289, Action W, Q-values: [-4.77 -4.08 -4.33 -3.77 -12.62 -14.38]
State 290, Action W, Q-values: [+3.42 +3.45 +3.43 +3.75 -4.97 -2.87]
State 291, Action W, Q-values: [-3.71 -3.86 -3.96 -3.09 -11.61 -12.45]
```

```
State 292, Action S, Q-values: [−0.56 −1.54 −1.61 −0.58 −9.64 −10.80]
State 293, Action S, Q-values: [+1.67 +0.11 +0.43 +1.47 −8.17 −8.08]
State 294, Action W, Q-values: [−0.74 −1.89 −1.75 −0.58 −9.28 −10.61]
State 295, Action W, Q-values: [+3.03 +3.28 +3.08 +3.66 −5.08 −3.10]
State 296, Action W, Q-values: [+2.88 +3.68 +4.40 +6.02 −4.59 −3.42]
State 297, Action N, Q-values: [+10.80 +14.31 +13.22 +11.43 +2.89 +4.51]
State 298, Action W, Q-values: [+3.09 +3.61 +4.36 +5.98 −2.59 −2.83]
State 299, Action S, Q-values: [+11.91 +9.23 +9.92 +10.77 +1.29 +1.87]
State 300, Action W, Q-values: [+3.38 +3.48 +3.28 +3.77 −5.18 −3.16]
State 301, Action N, Q-values: [−3.02 −1.48 −2.15 −2.07 −10.99 −12.14]
State 302, Action N, Q-values: [+0.80 +2.91 +1.44 +2.67 −7.21 −6.81]
State 303, Action N, Q-values: [−2.38 −0.60 −1.55 −1.28 −11.06 −11.62]
State 304, Action N, Q-values: [−5.15 −4.47 −5.23 −4.74 −14.40 −16.61]
State 305, Action W, Q-values: [+3.22 +3.24 +3.13 +3.54 −5.19 −3.32]
State 306, Action N, Q-values: [−5.39 −4.48 −5.16 −4.84 −14.29 −16.66]
State 307, Action N, Q-values: [−3.44 −2.52 −2.94 −2.97 −10.38 −11.76]
State 308, Action S, Q-values: [+5.98 +3.69 +4.73 +5.64 −3.29 −2.89]
State 309, Action S, Q-values: [+0.42 −0.70 −0.14 −0.65 −6.75 −7.63]
State 310, Action W, Q-values: [+3.46 +3.30 +3.20 +3.71 −5.24 −3.33]
State 311, Action S, Q-values: [+1.63 +0.29 +1.02 +0.60 −5.73 −6.95]
State 312, Action N, Q-values: [−3.98 −3.09 −3.73 −3.41 −11.67 −12.40]
State 313, Action N, Q-values: [−2.59 −1.54 −2.04 −2.01 −9.73 −10.57]
State 314, Action N, Q-values: [−3.83 −3.23 −4.01 −3.68 −11.81 −12.90]
State 315, Action W, Q-values: [+3.12 +3.32 +3.23 +3.68 −5.18 −2.91]
State 316, Action N, Q-values: [+8.83 +11.91 +10.42 +10.48 +2.44 +3.60]
State 317, Action N, Q-values: [+1.45 +4.42 +3.45 +3.63 −5.14 −5.95]
State 318, Action S, Q-values: [+17.02 +12.50 +14.37 +14.56 +5.26 +5.46]
State 319, Action N, Q-values: [+2.92 +5.99 +5.73 +5.34 −2.81 −3.08]
State 320, Action W, Q-values: [+3.48 +3.35 +3.26 +3.64 −4.94 −3.08]
State 321, Action N, Q-values: [−3.28 −2.31 −3.20 −3.00 −11.32 −11.74]
State 322, Action N, Q-values: [−0.08 +1.61 −0.53 +1.03 −9.09 −9.33]
State 323, Action N, Q-values: [−3.14 −1.54 −3.06 −1.92 −10.98 −12.19]
State 324, Action N, Q-values: [−4.46 −3.84 −4.30 −4.20 −12.07 −14.22]
State 325, Action W, Q-values: [+3.29 +3.54 +3.45 +3.82 −5.03 −2.77]
State 326, Action N, Q-values: [−4.51 −3.85 −4.10 −4.31 −12.04 −14.48]
State 327, Action N, Q-values: [−2.60 −1.61 −1.78 −2.16 −10.18 −10.83]
State 328, Action N, Q-values: [−0.14 +1.59 +0.03 +1.14 −8.62 −8.14]
State 329, Action N, Q-values: [−3.74 −2.40 −3.37 −2.66 −10.87 −11.73]
State 330, Action W, Q-values: [+3.03 +3.29 +3.19 +3.59 −5.02 −3.11]
State 331, Action N, Q-values: [−3.07 −1.56 −2.55 −2.01 −9.99 −10.91]
State 332, Action E, Q-values: [−3.17 −2.39 −2.35 −2.75 −9.92 −12.08]
State 333, Action N, Q-values: [−2.04 −0.57 −0.64 −1.31 −9.70 −10.87]
State 334, Action N, Q-values: [−3.16 −2.46 −2.72 −2.87 −11.04 −11.99]
State 335, Action W, Q-values: [+3.22 +3.41 +3.30 +3.69 −4.96 −3.17]
State 336, Action N, Q-values: [+5.96 +9.58 +8.23 +9.32 −1.78 −2.18]
State 337, Action E, Q-values: [+3.60 +5.48 +5.76 +5.03 −4.78 −3.85]
State 338, Action N, Q-values: [+5.78 +9.42 +8.12 +9.10 −1.84 −2.50]
State 339, Action E, Q-values: [+4.48 +7.24 +7.71 +6.77 −2.63 −2.25]
State 340, Action W, Q-values: [+3.35 +3.23 +3.19 +3.51 −5.01 −3.16]
State 341, Action N, Q-values: [−3.79 −3.09 −3.91 −3.11 −11.06 −12.84]
State 342, Action N, Q-values: [−1.30 +0.54 −0.77 +0.50 −9.04 −9.13]
State 343, Action W, Q-values: [−3.44 −2.34 −3.09 −2.26 −10.71 −11.84]
State 344, Action N, Q-values: [−4.10 −3.08 −3.74 −3.88 −11.41 −13.12]
State 345, Action W, Q-values: [+3.53 +3.58 +3.50 +3.85 −4.62 −2.66]
State 346, Action N, Q-values: [−4.00 −3.11 −3.45 −3.85 −10.90 −12.51]
State 347, Action N, Q-values: [−2.12 −0.64 −1.21 −1.35 −10.50 −10.98]
State 348, Action N, Q-values: [−1.36 +0.42 −1.03 +0.32 −9.42 −10.05]
```

```
State 349, Action N, Q-values: [−3.85 −3.07 −3.68 −3.10 −10.90 −12.07]
State 350, Action W, Q-values: [+3.39 +3.51 +3.39 +3.84 −4.54 −2.96]
State 351, Action N, Q-values: [−3.46 −2.39 −3.14 −2.64 −10.55 −12.03]
State 352, Action N, Q-values: [−2.67 −1.53 −2.13 −2.32 −10.47 −10.92]
State 353, Action N, Q-values: [−1.21 +0.45 −0.12 −0.64 −8.93 −10.02]
State 354, Action N, Q-values: [−2.74 −1.51 −1.99 −2.27 −9.73 −10.79]
State 355, Action W, Q-values: [+3.56 +3.76 +3.57 +4.02 −4.69 −2.42]
State 356, Action W, Q-values: [+4.24 +7.30 +6.38 +7.64 −2.76 −1.75]
State 357, Action N, Q-values: [+4.72 +7.61 +7.16 +5.81 −2.55 −1.98]
State 358, Action W, Q-values: [+3.69 +6.66 +5.83 +7.13 −3.15 −2.58]
State 359, Action N, Q-values: [+6.52 +9.65 +8.98 +7.73 −0.12 +0.60]
State 360, Action W, Q-values: [+3.45 +3.44 +3.34 +3.81 −4.96 −3.10]
State 361, Action N, Q-values: [−4.46 −3.82 −4.64 −4.24 −13.40 −14.78]
State 362, Action N, Q-values: [−2.20 −0.51 −1.71 −0.83 −10.33 −10.74]
State 363, Action N, Q-values: [−4.00 −3.13 −3.85 −3.47 −11.35 −12.08]
State 364, Action N, Q-values: [−3.57 −2.30 −2.53 −2.87 −11.05 −12.06]
State 365, Action W, Q-values: [+3.30 +3.27 +3.10 +3.60 −5.12 −3.21]
State 366, Action N, Q-values: [−3.35 −2.14 −2.31 −2.72 −10.55 −11.64]
State 367, Action E, Q-values: [−1.72 +0.40 +0.43 −0.22 −8.84 −9.81]
State 368, Action N, Q-values: [−2.12 −0.60 −1.81 −1.07 −9.81 −10.75]
State 369, Action N, Q-values: [−4.57 −3.88 −4.92 −4.05 −12.88 −14.60]
State 370, Action W, Q-values: [+3.39 +3.46 +3.29 +3.75 −4.94 −2.94]
State 371, Action N, Q-values: [−4.19 −3.16 −4.04 −3.36 −11.24 −12.26]
State 372, Action S, Q-values: [+1.62 +1.19 −0.03 +1.01 −7.47 −7.70]
State 373, Action S, Q-values: [+4.40 +1.63 +2.17 +3.48 −4.40 −5.86]
State 374, Action S, Q-values: [+1.61 +0.34 −0.06 +0.67 −5.54 −7.35]
State 375, Action W, Q-values: [+3.62 +3.66 +3.56 +3.97 −4.60 −2.72]
State 376, Action N, Q-values: [+3.31 +6.07 +5.39 +5.56 −3.11 −2.90]
State 377, Action N, Q-values: [+6.44 +9.82 +9.73 +8.91 −1.09 −1.10]
State 378, Action N, Q-values: [+2.98 +6.02 +4.98 +5.25 −3.61 −4.02]
State 379, Action S, Q-values: [+17.01 +12.24 +13.14 +15.00 +5.46 +6.81]
State 380, Action W, Q-values: [+3.32 +3.25 +3.16 +3.49 −5.23 −3.40]
State 381, Action W, Q-values: [−4.76 −4.45 −5.07 −4.36 −13.96 −15.93]
State 382, Action N, Q-values: [−2.99 −1.45 −2.38 −1.67 −10.07 −11.09]
State 383, Action W, Q-values: [−4.27 −3.87 −4.40 −3.75 −12.22 −14.29]
State 384, Action N, Q-values: [−2.77 −1.47 −2.18 −2.37 −11.03 −12.08]
State 385, Action W, Q-values: [+3.18 +3.21 +3.10 +3.53 −4.93 −3.19]
State 386, Action N, Q-values: [−2.88 −1.54 −2.52 −2.44 −11.91 −12.82]
State 387, Action N, Q-values: [−0.68 +1.48 +0.87 +0.12 −8.12 −8.66]
State 388, Action N, Q-values: [−2.88 −1.49 −2.31 −1.56 −10.49 −11.48]
State 389, Action W, Q-values: [−4.61 −4.45 −5.03 −4.39 −13.47 −15.84]
State 390, Action W, Q-values: [+3.43 +3.45 +3.41 +3.77 −4.75 −2.73]
State 391, Action N, Q-values: [−4.42 −3.86 −4.64 −3.91 −12.30 −14.44]
State 392, Action S, Q-values: [+0.55 −1.33 −0.75 +0.10 −8.68 −9.32]
State 393, Action S, Q-values: [+2.96 +1.36 +1.82 +2.94 −6.24 −6.18]
State 394, Action W, Q-values: [+0.46 −1.06 −0.53 +0.49 −8.41 −9.93]
State 395, Action W, Q-values: [+3.26 +3.31 +3.15 +3.66 −5.29 −3.26]
State 396, Action W, Q-values: [+2.03 +4.20 +3.34 +4.23 −6.12 −5.57]
State 397, Action N, Q-values: [+8.60 +11.92 +10.68 +8.99 +1.15 +1.58]
State 398, Action W, Q-values: [+2.13 +5.09 +4.19 +5.10 −5.28 −4.96]
State 399, Action S, Q-values: [+14.36 +11.03 +12.30 +12.60 +3.77 +3.80]
State 400, Action W, Q-values: [+3.23 +3.36 +3.19 +3.53 −5.14 −2.95]
State 401, Action N, Q-values: [−3.29 −2.38 −3.20 −2.75 −12.18 −13.25]
State 402, Action N, Q-values: [+0.74 +1.66 +0.96 +0.88 −8.34 −9.03]
State 403, Action N, Q-values: [−2.36 −1.67 −2.32 −2.17 −10.65 −11.38]
State 404, Action N, Q-values: [−5.77 −4.99 −5.43 −5.36 −16.28 −18.69]
State 405, Action W, Q-values: [+3.18 +3.34 +3.23 +3.57 −5.22 −3.39]
```

```
State 406, Action N, Q-values: [−5.79 −4.99 −5.48 −5.34 −16.80 −18.78]
State 407, Action N, Q-values: [−3.71 −3.12 −3.84 −3.36 −12.63 −12.47]
State 408, Action P, Q-values: [+5.96 +5.43 +5.72 +6.47 +7.79 −3.18]
State 409, Action P, Q-values: [−0.05 −0.10 +0.39 +0.49 +1.65 −5.63]
State 410, Action W, Q-values: [+3.28 +3.51 +3.40 +3.76 −5.15 −3.16]
State 411, Action P, Q-values: [+1.30 +0.83 +1.02 +1.71 +2.91 −6.78]
State 412, Action N, Q-values: [−4.42 −3.81 −4.49 −4.02 −12.59 −14.47]
State 413, Action N, Q-values: [−3.12 −2.50 −3.14 −2.89 −11.78 −13.21]
State 414, Action N, Q-values: [−4.65 −3.99 −4.77 −4.36 −13.35 −15.03]
State 415, Action W, Q-values: [+3.39 +3.52 +3.42 +3.77 −5.12 −3.04]
State 416, Action N, Q-values: [+7.62 +9.72 +8.56 +8.63 +0.47 +5.61]
State 417, Action N, Q-values: [+1.67 +2.99 +2.52 +2.20 −7.51 −0.50]
State 418, Action D, Q-values: [+16.10 +14.53 +15.74 +17.05 +7.99 +20.01]
State 419, Action N, Q-values: [+2.64 +4.35 +3.88 +3.66 −5.98 −0.91]
State 420, Action W, Q-values: [+3.29 +3.46 +3.32 +3.75 −5.01 −3.02]
State 421, Action N, Q-values: [−3.79 −3.14 −4.17 −3.65 −12.18 −13.84]
State 422, Action N, Q-values: [−0.39 +0.48 −0.93 −0.22 −10.11 −10.75]
State 423, Action N, Q-values: [−3.10 −2.34 −3.49 −2.83 −11.67 −12.91]
State 424, Action N, Q-values: [−4.71 −4.46 −4.58 −4.74 −13.26 −15.86]
State 425, Action W, Q-values: [+3.52 +3.39 +3.33 +3.86 −4.80 −3.11]
State 426, Action N, Q-values: [−5.08 −4.40 −4.70 −4.79 −14.05 −16.63]
State 427, Action N, Q-values: [−3.02 −2.36 −2.55 −2.99 −10.76 −13.66]
State 428, Action N, Q-values: [−0.49 +0.46 −1.38 −0.06 −9.58 −9.98]
State 429, Action N, Q-values: [−4.17 −3.17 −4.31 −3.50 −12.99 −13.59]
State 430, Action W, Q-values: [+3.35 +3.21 +3.17 +3.60 −5.07 −3.17]
State 431, Action N, Q-values: [−3.04 −2.40 −3.27 −2.87 −10.71 −12.46]
State 432, Action N, Q-values: [−3.79 −3.26 −3.44 −3.64 −11.23 −13.12]
State 433, Action N, Q-values: [−2.85 −1.51 −2.00 −2.17 −11.25 −12.86]
State 434, Action N, Q-values: [−3.54 −3.05 −3.38 −3.44 −10.51 −12.15]
State 435, Action W, Q-values: [+3.61 +3.65 +3.58 +3.94 −4.77 −2.67]
State 436, Action N, Q-values: [+4.99 +6.21 +5.17 +5.65 −4.31 −3.40]
State 437, Action E, Q-values: [+3.21 +3.71 +3.78 +3.68 −6.03 −5.31]
State 438, Action N, Q-values: [+4.82 +5.78 +4.81 +5.46 −4.46 −3.53]
State 439, Action E, Q-values: [+3.71 +4.22 +4.49 +4.22 −5.30 −4.78]
State 440, Action W, Q-values: [+3.32 +3.28 +3.16 +3.64 −4.99 −3.03]
State 441, Action W, Q-values: [−3.99 −3.80 −4.33 −3.80 −12.03 −14.12]
State 442, Action N, Q-values: [−2.05 −0.56 −1.53 −0.83 −10.29 −10.60]
State 443, Action W, Q-values: [−3.11 −3.07 −3.64 −3.02 −11.03 −13.62]
State 444, Action N, Q-values: [−4.28 −3.82 −4.51 −4.22 −12.27 −14.21]
State 445, Action W, Q-values: [+3.17 +3.38 +3.23 +3.62 −5.12 −3.31]
State 446, Action N, Q-values: [−4.47 −3.80 −4.58 −4.36 −13.34 −15.14]
State 447, Action N, Q-values: [−2.50 −1.52 −2.13 −2.28 −11.29 −12.03]
State 448, Action W, Q-values: [−1.18 −0.87 −1.79 −0.55 −9.82 −11.78]
State 449, Action N, Q-values: [−4.42 −3.81 −4.55 −3.97 −12.85 −14.84]
State 450, Action W, Q-values: [+3.65 +3.68 +3.57 +3.90 −4.83 −2.40]
State 451, Action N, Q-values: [−3.73 −3.23 −3.76 −3.42 −11.43 −12.22]
State 452, Action N, Q-values: [−3.11 −2.47 −3.20 −3.15 −11.58 −12.75]
State 453, Action N, Q-values: [−1.88 −0.65 −1.15 −1.59 −10.19 −10.70]
State 454, Action N, Q-values: [−2.58 −2.40 −2.89 −2.90 −11.18 −12.94]
State 455, Action W, Q-values: [+3.29 +3.35 +3.22 +3.66 −5.43 −3.11]
State 456, Action W, Q-values: [+4.09 +4.63 +4.15 +5.24 −5.17 −4.10]
State 457, Action N, Q-values: [+3.84 +4.57 +4.26 +4.41 −6.22 −4.90]
State 458, Action N, Q-values: [+4.08 +5.30 +4.70 +5.02 −4.46 −3.63]
State 459, Action N, Q-values: [+4.33 +5.88 +5.39 +4.87 −4.22 −3.58]
State 460, Action W, Q-values: [+3.10 +3.21 +3.08 +3.46 −5.10 −3.46]
State 461, Action N, Q-values: [−5.13 −4.39 −5.25 −4.83 −14.81 −17.02]
State 462, Action N, Q-values: [−2.69 −1.53 −2.43 −2.05 −11.05 −11.42]
```

```
State 463, Action N, Q-values: [-4.57 -3.93 -4.74 -4.18 -13.64 -14.89]
State 464, Action N, Q-values: [-3.56 -3.08 -3.26 -3.37 -11.46 -12.45]
State 465, Action W, Q-values: [+3.26 +3.40 +3.26 +3.65 -4.92 -3.15]
State 466, Action N, Q-values: [-3.80 -3.24 -3.78 -3.75 -12.54 -13.28]
State 467, Action N, Q-values: [-1.29 -0.57 -0.58 -1.11 -9.61 -10.19]
State 468, Action N, Q-values: [-1.87 -1.65 -3.06 -2.03 -11.32 -12.44]
State 469, Action N, Q-values: [-5.23 -4.51 -5.34 -4.85 -14.85 -16.74]
State 470, Action W, Q-values: [+3.44 +3.52 +3.35 +3.88 -5.08 -3.06]
State 471, Action N, Q-values: [-4.48 -3.85 -4.78 -4.21 -12.74 -14.76]
State 472, Action P, Q-values: [+0.81 +0.91 +0.80 +1.79 +2.94 -6.16]
State 473, Action P, Q-values: [+4.35 +3.66 +3.89 +4.74 +6.01 -4.12]
State 474, Action P, Q-values: [+0.86 +0.96 +1.02 +1.82 +2.94 -6.18]
State 475, Action W, Q-values: [+3.15 +3.38 +3.28 +3.66 -4.97 -3.27]
State 476, Action N, Q-values: [+2.86 +4.42 +3.49 +4.13 -6.25 +1.17]
State 477, Action N, Q-values: [+6.36 +7.85 +7.77 +6.70 -1.56 +3.24]
State 478, Action N, Q-values: [+2.88 +4.44 +3.23 +3.98 -6.04 -1.04]
State 479, Action D, Q-values: [+16.10 +14.19 +14.25 +16.62 +7.91 +19.99]
State 480, Action W, Q-values: [+3.37 +3.41 +3.35 +3.79 -4.81 -2.94]
State 481, Action N, Q-values: [-5.42 -4.99 -5.23 -5.14 -15.63 -18.09]
State 482, Action N, Q-values: [-3.16 -2.33 -3.11 -2.60 -12.01 -12.61]
State 483, Action N, Q-values: [-4.96 -4.50 -5.26 -4.55 -14.31 -16.29]
State 484, Action N, Q-values: [-2.97 -2.33 -3.08 -3.03 -11.25 -13.04]
State 485, Action W, Q-values: [+3.06 +3.30 +3.14 +3.67 -5.17 -3.52]
State 486, Action N, Q-values: [-3.25 -2.34 -2.80 -2.99 -12.12 -12.98]
State 487, Action N, Q-values: [-0.59 +0.43 -0.02 -0.32 -9.46 -10.04]
State 488, Action W, Q-values: [-3.12 -2.42 -3.38 -2.40 -11.28 -13.57]
State 489, Action W, Q-values: [-5.27 -4.93 -5.60 -4.92 -15.51 -18.25]
State 490, Action W, Q-values: [+3.39 +3.52 +3.45 +3.95 -5.08 -2.77]
State 491, Action W, Q-values: [-4.87 -4.51 -4.95 -4.43 -13.81 -16.02]
State 492, Action W, Q-values: [+1.03 -0.17 +0.52 +1.69 -6.27 -7.77]
State 493, Action W, Q-values: [+3.08 +2.74 +3.15 +4.42 -5.13 -4.30]
State 494, Action W, Q-values: [+0.83 +0.67 -0.15 +1.66 -7.55 -8.54]
State 495, Action W, Q-values: [+3.21 +3.55 +3.40 +3.78 -4.87 -3.04]
State 496, Action W, Q-values: [+2.10 +3.30 +2.74 +3.38 -7.21 -6.25]
State 497, Action N, Q-values: [+7.76 +9.71 +8.81 +7.25 -0.58 +0.04]
State 498, Action W, Q-values: [+2.42 +3.12 +2.62 +3.21 -6.82 -6.17]
State 499, Action W, Q-values: [+14.82 +13.24 +14.46 +16.94 +4.45 +7.58]
Episode 1: Total Reward = 9
Episode 2: Total Reward = 7
Episode 3: Total Reward = 8
Episode 4: Total Reward = 9
Episode 5: Total Reward = 5
Episode 6: Total Reward = 7
Episode 7: Total Reward = 6
Episode 8: Total Reward = 7
Episode 9: Total Reward = 6
Episode 10: Total Reward = 7
Episode 11: Total Reward = 7
Episode 12: Total Reward = 10
Episode 13: Total Reward = 7
Episode 14: Total Reward = 11
Episode 15: Total Reward = 12
Episode 16: Total Reward = 10
Episode 17: Total Reward = 9
Episode 18: Total Reward = 11
Episode 19: Total Reward = 9
Episode 20: Total Reward = 11
```

```
Episode 21: Total Reward = 4
Episode 22: Total Reward = 7
Episode 23: Total Reward = 8
Episode 24: Total Reward = 8
Episode 25: Total Reward = 7
Episode 26: Total Reward = 11
Episode 27: Total Reward = 10
Episode 28: Total Reward = 11
Episode 29: Total Reward = 12
Episode 30: Total Reward = 7
Episode 31: Total Reward = 10
Episode 32: Total Reward = 10
Episode 33: Total Reward = 6
Episode 34: Total Reward = 11
Episode 35: Total Reward = 10
Episode 36: Total Reward = 8
Episode 37: Total Reward = 8
Episode 38: Total Reward = 8
Episode 39: Total Reward = 7
Episode 40: Total Reward = 7
Episode 41: Total Reward = 7
Episode 42: Total Reward = 12
Episode 43: Total Reward = 6
Episode 44: Total Reward = 12
Episode 45: Total Reward = 8
Episode 46: Total Reward = 12
Episode 47: Total Reward = 4
Episode 48: Total Reward = 11
Episode 49: Total Reward = 9
Episode 50: Total Reward = 5
Episode 51: Total Reward = 8
Episode 52: Total Reward = 8
Episode 53: Total Reward = 9
Episode 54: Total Reward = 6
Episode 55: Total Reward = 10
Episode 56: Total Reward = 6
Episode 57: Total Reward = 14
Episode 58: Total Reward = 5
Episode 59: Total Reward = 8
Episode 60: Total Reward = 7
Episode 61: Total Reward = 10
Episode 62: Total Reward = 10
Episode 63: Total Reward = 9
Episode 64: Total Reward = 8
Episode 65: Total Reward = 9
Episode 66: Total Reward = 8
Episode 67: Total Reward = 7
Episode 68: Total Reward = 3
Episode 69: Total Reward = 8
Episode 70: Total Reward = 9
Episode 71: Total Reward = 8
Episode 72: Total Reward = 10
Episode 73: Total Reward = 11
Episode 74: Total Reward = 10
Episode 75: Total Reward = 8
Episode 76: Total Reward = 8
Episode 77: Total Reward = 5
```

```
Episode 78: Total Reward = 6
Episode 79: Total Reward = 9
Episode 80: Total Reward = 9
Episode 81: Total Reward = 4
Episode 82: Total Reward = 7
Episode 83: Total Reward = 10
Episode 84: Total Reward = 7
Episode 85: Total Reward = 9
Episode 86: Total Reward = 9
Episode 87: Total Reward = 10
Episode 88: Total Reward = 8
Episode 89: Total Reward = 14
Episode 90: Total Reward = 7
Episode 91: Total Reward = 6
Episode 92: Total Reward = 4
Episode 93: Total Reward = 7
Episode 94: Total Reward = 3
Episode 95: Total Reward = 5
Episode 96: Total Reward = 5
Episode 97: Total Reward = 7
Episode 98: Total Reward = 6
Episode 99: Total Reward = 10
Episode 100: Total Reward = 13
Episode 101: Total Reward = 11
Episode 102: Total Reward = 8
Episode 103: Total Reward = 8
Episode 104: Total Reward = 7
Episode 105: Total Reward = 11
Episode 106: Total Reward = 6
Episode 107: Total Reward = 9
Episode 108: Total Reward = 7
Episode 109: Total Reward = 3
Episode 110: Total Reward = 9
Episode 111: Total Reward = 3
Episode 112: Total Reward = 10
Episode 113: Total Reward = 8
Episode 114: Total Reward = 7
Episode 115: Total Reward = 11
Episode 116: Total Reward = 11
Episode 117: Total Reward = 7
Episode 118: Total Reward = 9
Episode 119: Total Reward = 7
Episode 120: Total Reward = 9
Episode 121: Total Reward = 9
Episode 122: Total Reward = 8
Episode 123: Total Reward = 3
Episode 124: Total Reward = 10
Episode 125: Total Reward = 7
Episode 126: Total Reward = 6
Episode 127: Total Reward = 4
Episode 128: Total Reward = 9
Episode 129: Total Reward = 7
Episode 130: Total Reward = 6
Episode 131: Total Reward = 7
Episode 132: Total Reward = 5
Episode 133: Total Reward = 14
Episode 134: Total Reward = 6
```

```
Episode 135: Total Reward = 11
Episode 136: Total Reward = 6
Episode 137: Total Reward = 9
Episode 138: Total Reward = 9
Episode 139: Total Reward = 7
Episode 140: Total Reward = 6
Episode 141: Total Reward = 11
Episode 142: Total Reward = 5
Episode 143: Total Reward = 5
Episode 144: Total Reward = 7
Episode 145: Total Reward = 10
Episode 146: Total Reward = 4
Episode 147: Total Reward = 12
Episode 148: Total Reward = 4
Episode 149: Total Reward = 9
Episode 150: Total Reward = 7
Episode 151: Total Reward = 8
Episode 152: Total Reward = 11
Episode 153: Total Reward = 7
Episode 154: Total Reward = 9
Episode 155: Total Reward = 8
Episode 156: Total Reward = 6
Episode 157: Total Reward = 6
Episode 158: Total Reward = 9
Episode 159: Total Reward = 7
Episode 160: Total Reward = 6
Episode 161: Total Reward = 6
Episode 162: Total Reward = 9
Episode 163: Total Reward = 7
Episode 164: Total Reward = 9
Episode 165: Total Reward = 9
Episode 166: Total Reward = 7
Episode 167: Total Reward = 9
Episode 168: Total Reward = 6
Episode 169: Total Reward = 7
Episode 170: Total Reward = 5
Episode 171: Total Reward = 9
Episode 172: Total Reward = 4
Episode 173: Total Reward = 6
Episode 174: Total Reward = 9
Episode 175: Total Reward = 7
Episode 176: Total Reward = 12
Episode 177: Total Reward = 6
Episode 178: Total Reward = 10
Episode 179: Total Reward = 8
Episode 180: Total Reward = 12
Episode 181: Total Reward = 9
Episode 182: Total Reward = 9
Episode 183: Total Reward = 5
Episode 184: Total Reward = 4
Episode 185: Total Reward = 3
Episode 186: Total Reward = 11
Episode 187: Total Reward = 8
Episode 188: Total Reward = 8
Episode 189: Total Reward = 4
Episode 190: Total Reward = 10
Episode 191: Total Reward = 7
```

```
Episode 192: Total Reward = 5
Episode 193: Total Reward = 6
Episode 194: Total Reward = 6
Episode 195: Total Reward = 8
Episode 196: Total Reward = 7
Episode 197: Total Reward = 4
Episode 198: Total Reward = 4
Episode 199: Total Reward = 5
Episode 200: Total Reward = 15
Episode 201: Total Reward = 6
Episode 202: Total Reward = 8
Episode 203: Total Reward = 8
Episode 204: Total Reward = 6
Episode 205: Total Reward = 9
Episode 206: Total Reward = 9
Episode 207: Total Reward = 12
Episode 208: Total Reward = 11
Episode 209: Total Reward = 4
Episode 210: Total Reward = 10
Episode 211: Total Reward = 5
Episode 212: Total Reward = 6
Episode 213: Total Reward = 7
Episode 214: Total Reward = 10
Episode 215: Total Reward = 14
Episode 216: Total Reward = 7
Episode 217: Total Reward = 7
Episode 218: Total Reward = 13
Episode 219: Total Reward = 3
Episode 220: Total Reward = 6
Episode 221: Total Reward = 5
Episode 222: Total Reward = 8
Episode 223: Total Reward = 8
Episode 224: Total Reward = 8
Episode 225: Total Reward = 12
Episode 226: Total Reward = 4
Episode 227: Total Reward = 8
Episode 228: Total Reward = 7
Episode 229: Total Reward = 11
Episode 230: Total Reward = 11
Episode 231: Total Reward = 8
Episode 232: Total Reward = 8
Episode 233: Total Reward = 6
Episode 234: Total Reward = 6
Episode 235: Total Reward = 4
Episode 236: Total Reward = 4
Episode 237: Total Reward = 9
Episode 238: Total Reward = 8
Episode 239: Total Reward = 7
Episode 240: Total Reward = 11
Episode 241: Total Reward = 9
Episode 242: Total Reward = 11
Episode 243: Total Reward = 7
Episode 244: Total Reward = 5
Episode 245: Total Reward = 9
Episode 246: Total Reward = 8
Episode 247: Total Reward = 9
Episode 248: Total Reward = 9
```

```
Episode 249: Total Reward = 8
Episode 250: Total Reward = 6
Episode 251: Total Reward = 5
Episode 252: Total Reward = 6
Episode 253: Total Reward = 7
Episode 254: Total Reward = 6
Episode 255: Total Reward = 6
Episode 256: Total Reward = 10
Episode 257: Total Reward = 8
Episode 258: Total Reward = 14
Episode 259: Total Reward = 6
Episode 260: Total Reward = 6
Episode 261: Total Reward = 10
Episode 262: Total Reward = 10
Episode 263: Total Reward = 12
Episode 264: Total Reward = 7
Episode 265: Total Reward = 9
Episode 266: Total Reward = 5
Episode 267: Total Reward = 4
Episode 268: Total Reward = 10
Episode 269: Total Reward = 10
Episode 270: Total Reward = 3
Episode 271: Total Reward = 9
Episode 272: Total Reward = 7
Episode 273: Total Reward = 12
Episode 274: Total Reward = 9
Episode 275: Total Reward = 10
Episode 276: Total Reward = 6
Episode 277: Total Reward = 12
Episode 278: Total Reward = 11
Episode 279: Total Reward = 8
Episode 280: Total Reward = 7
Episode 281: Total Reward = 12
Episode 282: Total Reward = 9
Episode 283: Total Reward = 7
Episode 284: Total Reward = 4
Episode 285: Total Reward = 7
Episode 286: Total Reward = 7
Episode 287: Total Reward = 9
Episode 288: Total Reward = 9
Episode 289: Total Reward = 10
Episode 290: Total Reward = 14
Episode 291: Total Reward = 5
Episode 292: Total Reward = 5
Episode 293: Total Reward = 7
Episode 294: Total Reward = 11
Episode 295: Total Reward = 10
Episode 296: Total Reward = 8
Episode 297: Total Reward = 7
Episode 298: Total Reward = 12
Episode 299: Total Reward = 7
Episode 300: Total Reward = 3
Episode 301: Total Reward = 3
Episode 302: Total Reward = 9
Episode 303: Total Reward = 9
Episode 304: Total Reward = 7
Episode 305: Total Reward = 7
```

```
Episode 306: Total Reward = 6
Episode 307: Total Reward = 3
Episode 308: Total Reward = 9
Episode 309: Total Reward = 7
Episode 310: Total Reward = 5
Episode 311: Total Reward = 10
Episode 312: Total Reward = 3
Episode 313: Total Reward = 4
Episode 314: Total Reward = 8
Episode 315: Total Reward = 4
Episode 316: Total Reward = 9
Episode 317: Total Reward = 11
Episode 318: Total Reward = 7
Episode 319: Total Reward = 7
Episode 320: Total Reward = 5
Episode 321: Total Reward = 10
Episode 322: Total Reward = 7
Episode 323: Total Reward = 11
Episode 324: Total Reward = 8
Episode 325: Total Reward = 10
Episode 326: Total Reward = 8
Episode 327: Total Reward = 11
Episode 328: Total Reward = 7
Episode 329: Total Reward = 7
Episode 330: Total Reward = 7
Episode 331: Total Reward = 8
Episode 332: Total Reward = 7
Episode 333: Total Reward = 10
Episode 334: Total Reward = 9
Episode 335: Total Reward = 4
Episode 336: Total Reward = 9
Episode 337: Total Reward = 8
Episode 338: Total Reward = 4
Episode 339: Total Reward = 8
Episode 340: Total Reward = 7
Episode 341: Total Reward = 4
Episode 342: Total Reward = 8
Episode 343: Total Reward = 9
Episode 344: Total Reward = 8
Episode 345: Total Reward = 9
Episode 346: Total Reward = 7
Episode 347: Total Reward = 13
Episode 348: Total Reward = 8
Episode 349: Total Reward = 8
Episode 350: Total Reward = 13
Episode 351: Total Reward = 12
Episode 352: Total Reward = 5
Episode 353: Total Reward = 8
Episode 354: Total Reward = 12
Episode 355: Total Reward = 4
Episode 356: Total Reward = 8
Episode 357: Total Reward = 5
Episode 358: Total Reward = 10
Episode 359: Total Reward = 4
Episode 360: Total Reward = 11
Episode 361: Total Reward = 8
Episode 362: Total Reward = 7
```

```
Episode 363: Total Reward = 9
Episode 364: Total Reward = 5
Episode 365: Total Reward = 8
Episode 366: Total Reward = 13
Episode 367: Total Reward = 4
Episode 368: Total Reward = 9
Episode 369: Total Reward = 7
Episode 370: Total Reward = 8
Episode 371: Total Reward = 6
Episode 372: Total Reward = 7
Episode 373: Total Reward = 10
Episode 374: Total Reward = 3
Episode 375: Total Reward = 5
Episode 376: Total Reward = 10
Episode 377: Total Reward = 4
Episode 378: Total Reward = 13
Episode 379: Total Reward = 11
Episode 380: Total Reward = 5
Episode 381: Total Reward = 10
Episode 382: Total Reward = 8
Episode 383: Total Reward = 11
Episode 384: Total Reward = 6
Episode 385: Total Reward = 11
Episode 386: Total Reward = 8
Episode 387: Total Reward = 5
Episode 388: Total Reward = 14
Episode 389: Total Reward = 7
Episode 390: Total Reward = 6
Episode 391: Total Reward = 8
Episode 392: Total Reward = 7
Episode 393: Total Reward = 9
Episode 394: Total Reward = 10
Episode 395: Total Reward = 8
Episode 396: Total Reward = 9
Episode 397: Total Reward = 7
Episode 398: Total Reward = 12
Episode 399: Total Reward = 10
Episode 400: Total Reward = 7
Episode 401: Total Reward = 7
Episode 402: Total Reward = 10
Episode 403: Total Reward = 7
Episode 404: Total Reward = 7
Episode 405: Total Reward = 6
Episode 406: Total Reward = 7
Episode 407: Total Reward = 4
Episode 408: Total Reward = 9
Episode 409: Total Reward = 10
Episode 410: Total Reward = 6
Episode 411: Total Reward = 4
Episode 412: Total Reward = 9
Episode 413: Total Reward = 6
Episode 414: Total Reward = 11
Episode 415: Total Reward = 8
Episode 416: Total Reward = 11
Episode 417: Total Reward = 10
Episode 418: Total Reward = 12
Episode 419: Total Reward = 6
```

```
Episode 420: Total Reward = 7
Episode 421: Total Reward = 4
Episode 422: Total Reward = 6
Episode 423: Total Reward = 10
Episode 424: Total Reward = 8
Episode 425: Total Reward = 13
Episode 426: Total Reward = 6
Episode 427: Total Reward = 6
Episode 428: Total Reward = 9
Episode 429: Total Reward = 10
Episode 430: Total Reward = 8
Episode 431: Total Reward = 9
Episode 432: Total Reward = 8
Episode 433: Total Reward = 13
Episode 434: Total Reward = 8
Episode 435: Total Reward = 10
Episode 436: Total Reward = 10
Episode 437: Total Reward = 6
Episode 438: Total Reward = 9
Episode 439: Total Reward = 14
Episode 440: Total Reward = 4
Episode 441: Total Reward = 5
Episode 442: Total Reward = 13
Episode 443: Total Reward = 7
Episode 444: Total Reward = 12
Episode 445: Total Reward = 5
Episode 446: Total Reward = 5
Episode 447: Total Reward = 12
Episode 448: Total Reward = 8
Episode 449: Total Reward = 11
Episode 450: Total Reward = 9
Episode 451: Total Reward = 6
Episode 452: Total Reward = 5
Episode 453: Total Reward = 8
Episode 454: Total Reward = 6
Episode 455: Total Reward = 7
Episode 456: Total Reward = 7
Episode 457: Total Reward = 5
Episode 458: Total Reward = 8
Episode 459: Total Reward = 8
Episode 460: Total Reward = 9
Episode 461: Total Reward = 11
Episode 462: Total Reward = 14
Episode 463: Total Reward = 3
Episode 464: Total Reward = 11
Episode 465: Total Reward = 7
Episode 466: Total Reward = 15
Episode 467: Total Reward = 6
Episode 468: Total Reward = 7
Episode 469: Total Reward = 11
Episode 470: Total Reward = 11
Episode 471: Total Reward = 9
Episode 472: Total Reward = 13
Episode 473: Total Reward = 10
Episode 474: Total Reward = 12
Episode 475: Total Reward = 7
Episode 476: Total Reward = 9
```

```
Episode 477: Total Reward = 7
Episode 478: Total Reward = 9
Episode 479: Total Reward = 9
Episode 480: Total Reward = 9
Episode 481: Total Reward = 10
Episode 482: Total Reward = 5
Episode 483: Total Reward = 7
Episode 484: Total Reward = 9
Episode 485: Total Reward = 6
Episode 486: Total Reward = 11
Episode 487: Total Reward = 8
Episode 488: Total Reward = 8
Episode 489: Total Reward = 6
Episode 490: Total Reward = 10
Episode 491: Total Reward = 10
Episode 492: Total Reward = 6
Episode 493: Total Reward = 8
Episode 494: Total Reward = 12
Episode 495: Total Reward = 5
Episode 496: Total Reward = 7
Episode 497: Total Reward = 9
Episode 498: Total Reward = 9
Episode 499: Total Reward = 9
Episode 500: Total Reward = 6
Episode 501: Total Reward = 14
Episode 502: Total Reward = 9
Episode 503: Total Reward = 7
Episode 504: Total Reward = 9
Episode 505: Total Reward = 4
Episode 506: Total Reward = 8
Episode 507: Total Reward = 7
Episode 508: Total Reward = 9
Episode 509: Total Reward = 8
Episode 510: Total Reward = 8
Episode 511: Total Reward = 9
Episode 512: Total Reward = 10
Episode 513: Total Reward = 10
Episode 514: Total Reward = 11
Episode 515: Total Reward = 13
Episode 516: Total Reward = 6
Episode 517: Total Reward = 3
Episode 518: Total Reward = 9
Episode 519: Total Reward = 9
Episode 520: Total Reward = 4
Episode 521: Total Reward = 8
Episode 522: Total Reward = 8
Episode 523: Total Reward = 9
Episode 524: Total Reward = 3
Episode 525: Total Reward = 14
Episode 526: Total Reward = 9
Episode 527: Total Reward = 8
Episode 528: Total Reward = 5
Episode 529: Total Reward = 6
Episode 530: Total Reward = 8
Episode 531: Total Reward = 10
Episode 532: Total Reward = 5
Episode 533: Total Reward = 6
```

```
Episode 534: Total Reward = 7
Episode 535: Total Reward = 7
Episode 536: Total Reward = 8
Episode 537: Total Reward = 9
Episode 538: Total Reward = 4
Episode 539: Total Reward = 5
Episode 540: Total Reward = 7
Episode 541: Total Reward = 6
Episode 542: Total Reward = 10
Episode 543: Total Reward = 10
Episode 544: Total Reward = 9
Episode 545: Total Reward = 10
Episode 546: Total Reward = 8
Episode 547: Total Reward = 6
Episode 548: Total Reward = 5
Episode 549: Total Reward = 8
Episode 550: Total Reward = 5
Episode 551: Total Reward = 10
Episode 552: Total Reward = 5
Episode 553: Total Reward = 13
Episode 554: Total Reward = 7
Episode 555: Total Reward = 5
Episode 556: Total Reward = 9
Episode 557: Total Reward = 10
Episode 558: Total Reward = 10
Episode 559: Total Reward = 10
Episode 560: Total Reward = 10
Episode 561: Total Reward = 4
Episode 562: Total Reward = 8
Episode 563: Total Reward = 9
Episode 564: Total Reward = 10
Episode 565: Total Reward = 11
Episode 566: Total Reward = 8
Episode 567: Total Reward = 9
Episode 568: Total Reward = 10
Episode 569: Total Reward = 9
Episode 570: Total Reward = 9
Episode 571: Total Reward = 11
Episode 572: Total Reward = 11
Episode 573: Total Reward = 9
Episode 574: Total Reward = 6
Episode 575: Total Reward = 7
Episode 576: Total Reward = 5
Episode 577: Total Reward = 11
Episode 578: Total Reward = 13
Episode 579: Total Reward = 10
Episode 580: Total Reward = 12
Episode 581: Total Reward = 10
Episode 582: Total Reward = 7
Episode 583: Total Reward = 4
Episode 584: Total Reward = 11
Episode 585: Total Reward = 3
Episode 586: Total Reward = 11
Episode 587: Total Reward = 7
Episode 588: Total Reward = 6
Episode 589: Total Reward = 4
Episode 590: Total Reward = 6
```

```
Episode 591: Total Reward = 7
Episode 592: Total Reward = 6
Episode 593: Total Reward = 5
Episode 594: Total Reward = 6
Episode 595: Total Reward = 9
Episode 596: Total Reward = 7
Episode 597: Total Reward = 9
Episode 598: Total Reward = 5
Episode 599: Total Reward = 11
Episode 600: Total Reward = 9
Episode 601: Total Reward = 8
Episode 602: Total Reward = 8
Episode 603: Total Reward = 10
Episode 604: Total Reward = 5
Episode 605: Total Reward = 6
Episode 606: Total Reward = 10
Episode 607: Total Reward = 8
Episode 608: Total Reward = 4
Episode 609: Total Reward = 3
Episode 610: Total Reward = 4
Episode 611: Total Reward = 6
Episode 612: Total Reward = 11
Episode 613: Total Reward = 9
Episode 614: Total Reward = 9
Episode 615: Total Reward = 5
Episode 616: Total Reward = 12
Episode 617: Total Reward = 13
Episode 618: Total Reward = 12
Episode 619: Total Reward = 7
Episode 620: Total Reward = 9
Episode 621: Total Reward = 11
Episode 622: Total Reward = 7
Episode 623: Total Reward = 9
Episode 624: Total Reward = 5
Episode 625: Total Reward = 9
Episode 626: Total Reward = 11
Episode 627: Total Reward = 9
Episode 628: Total Reward = 6
Episode 629: Total Reward = 8
Episode 630: Total Reward = 6
Episode 631: Total Reward = 4
Episode 632: Total Reward = 8
Episode 633: Total Reward = 15
Episode 634: Total Reward = 8
Episode 635: Total Reward = 6
Episode 636: Total Reward = 14
Episode 637: Total Reward = 11
Episode 638: Total Reward = 3
Episode 639: Total Reward = 4
Episode 640: Total Reward = 8
Episode 641: Total Reward = 11
Episode 642: Total Reward = 9
Episode 643: Total Reward = 14
Episode 644: Total Reward = 8
Episode 645: Total Reward = 12
Episode 646: Total Reward = 13
Episode 647: Total Reward = 12
```

```
Episode 648: Total Reward = 5
Episode 649: Total Reward = 9
Episode 650: Total Reward = 7
Episode 651: Total Reward = 6
Episode 652: Total Reward = 8
Episode 653: Total Reward = 12
Episode 654: Total Reward = 10
Episode 655: Total Reward = 8
Episode 656: Total Reward = 7
Episode 657: Total Reward = 7
Episode 658: Total Reward = 5
Episode 659: Total Reward = 8
Episode 660: Total Reward = 5
Episode 661: Total Reward = 4
Episode 662: Total Reward = 6
Episode 663: Total Reward = 7
Episode 664: Total Reward = 10
Episode 665: Total Reward = 3
Episode 666: Total Reward = 4
Episode 667: Total Reward = 3
Episode 668: Total Reward = 8
Episode 669: Total Reward = 10
Episode 670: Total Reward = 3
Episode 671: Total Reward = 4
Episode 672: Total Reward = 9
Episode 673: Total Reward = 5
Episode 674: Total Reward = 3
Episode 675: Total Reward = 10
Episode 676: Total Reward = 11
Episode 677: Total Reward = 7
Episode 678: Total Reward = 9
Episode 679: Total Reward = 3
Episode 680: Total Reward = 7
Episode 681: Total Reward = 5
Episode 682: Total Reward = 9
Episode 683: Total Reward = 8
Episode 684: Total Reward = 4
Episode 685: Total Reward = 9
Episode 686: Total Reward = 7
Episode 687: Total Reward = 7
Episode 688: Total Reward = 6
Episode 689: Total Reward = 10
Episode 690: Total Reward = 10
Episode 691: Total Reward = 9
Episode 692: Total Reward = 6
Episode 693: Total Reward = 10
Episode 694: Total Reward = 12
Episode 695: Total Reward = 5
Episode 696: Total Reward = 8
Episode 697: Total Reward = 8
Episode 698: Total Reward = 6
Episode 699: Total Reward = 9
Episode 700: Total Reward = 10
Episode 701: Total Reward = 6
Episode 702: Total Reward = 9
Episode 703: Total Reward = 9
Episode 704: Total Reward = 5
```

```
Episode 705: Total Reward = 9
Episode 706: Total Reward = 9
Episode 707: Total Reward = 8
Episode 708: Total Reward = 8
Episode 709: Total Reward = 6
Episode 710: Total Reward = 8
Episode 711: Total Reward = 7
Episode 712: Total Reward = 6
Episode 713: Total Reward = 5
Episode 714: Total Reward = 12
Episode 715: Total Reward = 3
Episode 716: Total Reward = 13
Episode 717: Total Reward = 11
Episode 718: Total Reward = 7
Episode 719: Total Reward = 9
Episode 720: Total Reward = 11
Episode 721: Total Reward = 11
Episode 722: Total Reward = 3
Episode 723: Total Reward = 7
Episode 724: Total Reward = 7
Episode 725: Total Reward = 10
Episode 726: Total Reward = 8
Episode 727: Total Reward = 10
Episode 728: Total Reward = 10
Episode 729: Total Reward = 9
Episode 730: Total Reward = 9
Episode 731: Total Reward = 7
Episode 732: Total Reward = 9
Episode 733: Total Reward = 12
Episode 734: Total Reward = 7
Episode 735: Total Reward = 5
Episode 736: Total Reward = 7
Episode 737: Total Reward = 7
Episode 738: Total Reward = 7
Episode 739: Total Reward = 8
Episode 740: Total Reward = 9
Episode 741: Total Reward = 7
Episode 742: Total Reward = 8
Episode 743: Total Reward = 6
Episode 744: Total Reward = 6
Episode 745: Total Reward = 8
Episode 746: Total Reward = 7
Episode 747: Total Reward = 9
Episode 748: Total Reward = 4
Episode 749: Total Reward = 12
Episode 750: Total Reward = 5
Episode 751: Total Reward = 7
Episode 752: Total Reward = 5
Episode 753: Total Reward = 11
Episode 754: Total Reward = 3
Episode 755: Total Reward = 11
Episode 756: Total Reward = 4
Episode 757: Total Reward = 7
Episode 758: Total Reward = 7
Episode 759: Total Reward = 9
Episode 760: Total Reward = 10
Episode 761: Total Reward = 4
```

```
Episode 762: Total Reward = 10
Episode 763: Total Reward = 6
Episode 764: Total Reward = 7
Episode 765: Total Reward = 4
Episode 766: Total Reward = 7
Episode 767: Total Reward = 10
Episode 768: Total Reward = 9
Episode 769: Total Reward = 9
Episode 770: Total Reward = 10
Episode 771: Total Reward = 8
Episode 772: Total Reward = 12
Episode 773: Total Reward = 9
Episode 774: Total Reward = 10
Episode 775: Total Reward = 9
Episode 776: Total Reward = 10
Episode 777: Total Reward = 8
Episode 778: Total Reward = 7
Episode 779: Total Reward = 4
Episode 780: Total Reward = 7
Episode 781: Total Reward = 8
Episode 782: Total Reward = 11
Episode 783: Total Reward = 5
Episode 784: Total Reward = 5
Episode 785: Total Reward = 13
Episode 786: Total Reward = 7
Episode 787: Total Reward = 4
Episode 788: Total Reward = 5
Episode 789: Total Reward = 5
Episode 790: Total Reward = 9
Episode 791: Total Reward = 9
Episode 792: Total Reward = 4
Episode 793: Total Reward = 12
Episode 794: Total Reward = 7
Episode 795: Total Reward = 8
Episode 796: Total Reward = 8
Episode 797: Total Reward = 10
Episode 798: Total Reward = 10
Episode 799: Total Reward = 10
Episode 800: Total Reward = 7
Episode 801: Total Reward = 5
Episode 802: Total Reward = 12
Episode 803: Total Reward = 9
Episode 804: Total Reward = 6
Episode 805: Total Reward = 6
Episode 806: Total Reward = 8
Episode 807: Total Reward = 7
Episode 808: Total Reward = 7
Episode 809: Total Reward = 7
Episode 810: Total Reward = 7
Episode 811: Total Reward = 11
Episode 812: Total Reward = 8
Episode 813: Total Reward = 6
Episode 814: Total Reward = 9
Episode 815: Total Reward = 9
Episode 816: Total Reward = 4
Episode 817: Total Reward = 10
Episode 818: Total Reward = 6
```

```
Episode 819: Total Reward = 10
Episode 820: Total Reward = 10
Episode 821: Total Reward = 5
Episode 822: Total Reward = 8
Episode 823: Total Reward = 11
Episode 824: Total Reward = 9
Episode 825: Total Reward = 11
Episode 826: Total Reward = 10
Episode 827: Total Reward = 7
Episode 828: Total Reward = 8
Episode 829: Total Reward = 10
Episode 830: Total Reward = 8
Episode 831: Total Reward = 6
Episode 832: Total Reward = 5
Episode 833: Total Reward = 9
Episode 834: Total Reward = 8
Episode 835: Total Reward = 12
Episode 836: Total Reward = 4
Episode 837: Total Reward = 5
Episode 838: Total Reward = 11
Episode 839: Total Reward = 9
Episode 840: Total Reward = 10
Episode 841: Total Reward = 9
Episode 842: Total Reward = 8
Episode 843: Total Reward = 8
Episode 844: Total Reward = 6
Episode 845: Total Reward = 8
Episode 846: Total Reward = 7
Episode 847: Total Reward = 6
Episode 848: Total Reward = 5
Episode 849: Total Reward = 7
Episode 850: Total Reward = 12
Episode 851: Total Reward = 5
Episode 852: Total Reward = 10
Episode 853: Total Reward = 7
Episode 854: Total Reward = 8
Episode 855: Total Reward = 14
Episode 856: Total Reward = 14
Episode 857: Total Reward = 4
Episode 858: Total Reward = 4
Episode 859: Total Reward = 9
Episode 860: Total Reward = 7
Episode 861: Total Reward = 8
Episode 862: Total Reward = 11
Episode 863: Total Reward = 9
Episode 864: Total Reward = 13
Episode 865: Total Reward = 3
Episode 866: Total Reward = 10
Episode 867: Total Reward = 10
Episode 868: Total Reward = 4
Episode 869: Total Reward = 7
Episode 870: Total Reward = 5
Episode 871: Total Reward = 5
Episode 872: Total Reward = 11
Episode 873: Total Reward = 8
Episode 874: Total Reward = 6
Episode 875: Total Reward = 7
```

```
Episode 876: Total Reward = 4
Episode 877: Total Reward = 3
Episode 878: Total Reward = 7
Episode 879: Total Reward = 11
Episode 880: Total Reward = 9
Episode 881: Total Reward = 8
Episode 882: Total Reward = 9
Episode 883: Total Reward = 3
Episode 884: Total Reward = 7
Episode 885: Total Reward = 4
Episode 886: Total Reward = 9
Episode 887: Total Reward = 9
Episode 888: Total Reward = 7
Episode 889: Total Reward = 6
Episode 890: Total Reward = 9
Episode 891: Total Reward = 6
Episode 892: Total Reward = 10
Episode 893: Total Reward = 8
Episode 894: Total Reward = 7
Episode 895: Total Reward = 12
Episode 896: Total Reward = 5
Episode 897: Total Reward = 5
Episode 898: Total Reward = 5
Episode 899: Total Reward = 5
Episode 900: Total Reward = 6
Episode 901: Total Reward = 5
Episode 902: Total Reward = 11
Episode 903: Total Reward = 14
Episode 904: Total Reward = 12
Episode 905: Total Reward = 8
Episode 906: Total Reward = 9
Episode 907: Total Reward = 5
Episode 908: Total Reward = 11
Episode 909: Total Reward = 8
Episode 910: Total Reward = 10
Episode 911: Total Reward = 10
Episode 912: Total Reward = 6
Episode 913: Total Reward = 5
Episode 914: Total Reward = 8
Episode 915: Total Reward = 12
Episode 916: Total Reward = 10
Episode 917: Total Reward = 9
Episode 918: Total Reward = 7
Episode 919: Total Reward = 10
Episode 920: Total Reward = 5
Episode 921: Total Reward = 7
Episode 922: Total Reward = 11
Episode 923: Total Reward = 9
Episode 924: Total Reward = 5
Episode 925: Total Reward = 11
Episode 926: Total Reward = 10
Episode 927: Total Reward = 5
Episode 928: Total Reward = 6
Episode 929: Total Reward = 11
Episode 930: Total Reward = 12
Episode 931: Total Reward = 11
Episode 932: Total Reward = 11
```

```
Episode 933: Total Reward = 10
Episode 934: Total Reward = 10
Episode 935: Total Reward = 9
Episode 936: Total Reward = 15
Episode 937: Total Reward = 8
Episode 938: Total Reward = 6
Episode 939: Total Reward = 8
Episode 940: Total Reward = 10
Episode 941: Total Reward = 8
Episode 942: Total Reward = 10
Episode 943: Total Reward = 6
Episode 944: Total Reward = 4
Episode 945: Total Reward = 4
Episode 946: Total Reward = 7
Episode 947: Total Reward = 7
Episode 948: Total Reward = 9
Episode 949: Total Reward = 6
Episode 950: Total Reward = 9
Episode 951: Total Reward = 9
Episode 952: Total Reward = 7
Episode 953: Total Reward = 7
Episode 954: Total Reward = 3
Episode 955: Total Reward = 7
Episode 956: Total Reward = 5
Episode 957: Total Reward = 7
Episode 958: Total Reward = 10
Episode 959: Total Reward = 8
Episode 960: Total Reward = 8
Episode 961: Total Reward = 9
Episode 962: Total Reward = 4
Episode 963: Total Reward = 7
Episode 964: Total Reward = 8
Episode 965: Total Reward = 6
Episode 966: Total Reward = 6
Episode 967: Total Reward = 13
Episode 968: Total Reward = 4
Episode 969: Total Reward = 13
Episode 970: Total Reward = 12
Episode 971: Total Reward = 10
Episode 972: Total Reward = 8
Episode 973: Total Reward = 11
Episode 974: Total Reward = 6
Episode 975: Total Reward = 7
Episode 976: Total Reward = 14
Episode 977: Total Reward = 9
Episode 978: Total Reward = 14
Episode 979: Total Reward = 9
Episode 980: Total Reward = 10
Episode 981: Total Reward = 8
Episode 982: Total Reward = 14
Episode 983: Total Reward = 15
Episode 984: Total Reward = 12
Episode 985: Total Reward = 6
Episode 986: Total Reward = 8
Episode 987: Total Reward = 6
Episode 988: Total Reward = 9
Episode 989: Total Reward = 12
```

```
Episode 990: Total Reward = 6
Episode 991: Total Reward = 6
Episode 992: Total Reward = 6
Episode 993: Total Reward = 4
Episode 994: Total Reward = 9
Episode 995: Total Reward = 4
Episode 996: Total Reward = 11
Episode 997: Total Reward = 9
Episode 998: Total Reward = 5
Episode 999: Total Reward = 13
Episode 1000: Total Reward = 8
Average Reward across all Episodes: = 7.994
CPU times: user 1.1 s, sys: 49.7 ms, total: 1.15 s
Wall time: 567 ms
```

In [85]:
```python
# # only to have a look of the test play result
my_dqn_agent_fine_tuned.test(test_episodes=5, display=True)
```


Taxi-v3 | Episode: 4 | Step: 10
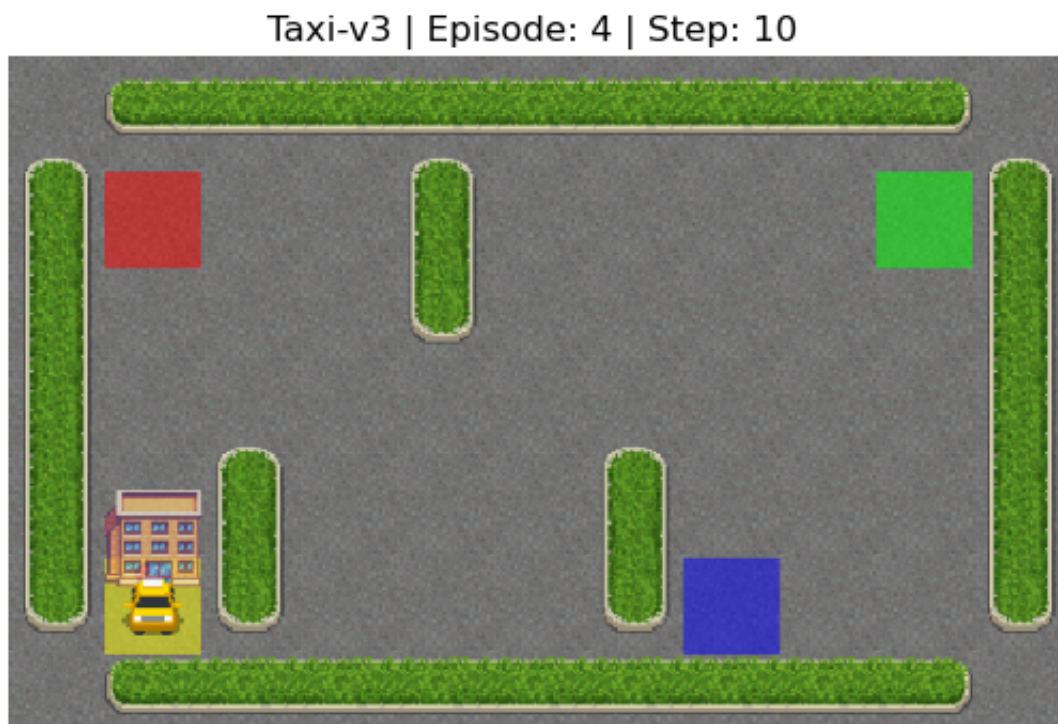
```
Episode 5: Total Reward = 11
Average Reward across all Episodes: = 9.0
```

Out[85]: [6, 8, 10, 10, 11]

Taxi-v3 | Episode: 4 | Step: 10
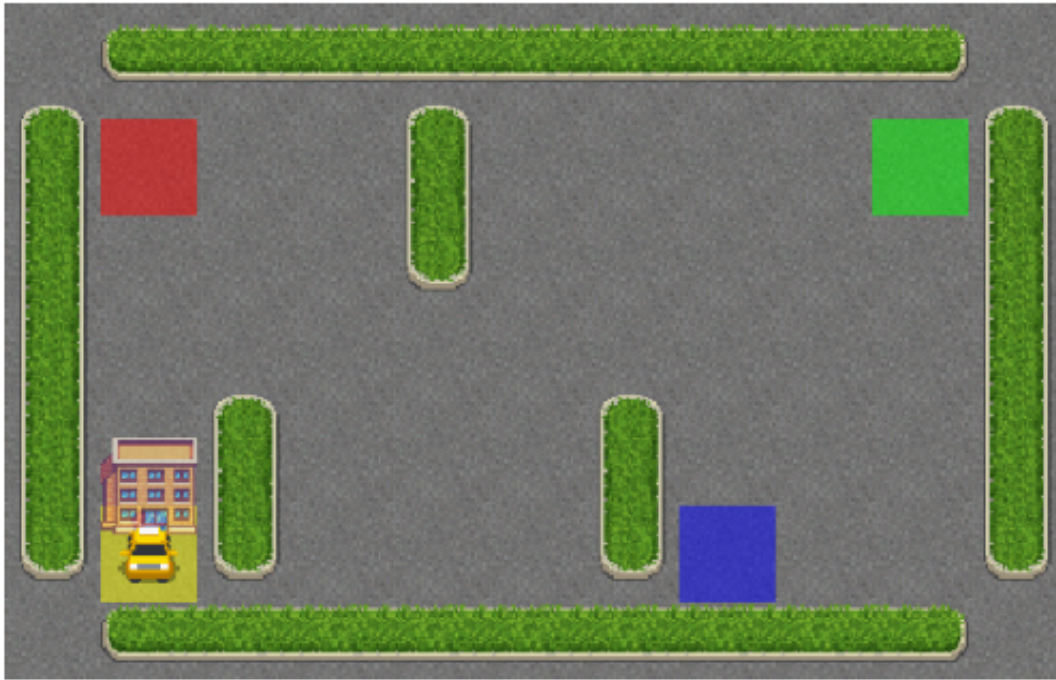
```
In [86]: my_dqn_agent_fine_tuned.test_to_gif(filename="taxi_dqn_fine_tuned",episod
```

```
===== Trained Policy ===== :
State 000, Action W, Q-values: [+3.33 +3.51 +3.39 +3.85 −4.99 −2.85]
State 001, Action P, Q-values: [−0.23 +0.28 +0.40 +0.70 +1.66 −5.24]
State 002, Action P, Q-values: [+5.54 +6.02 +5.56 +6.42 +7.76 −3.03]
State 003, Action P, Q-values: [+1.22 +1.21 +1.05 +1.86 +2.92 −6.50]
State 004, Action S, Q-values: [−4.93 −5.18 −5.00 −5.16 −14.92 −18.07]
State 005, Action W, Q-values: [+3.16 +3.31 +3.17 +3.59 −4.93 −3.14]
State 006, Action E, Q-values: [−5.20 −5.24 −5.03 −5.31 −14.88 −17.98]
State 007, Action S, Q-values: [−3.18 −3.82 −3.67 −3.68 −11.11 −12.79]
State 008, Action S, Q-values: [+1.66 +0.91 −0.14 +1.04 −7.84 −9.71]
State 009, Action S, Q-values: [−2.33 −2.77 −3.17 −2.61 −11.63 −12.20]
State 010, Action W, Q-values: [+3.49 +3.45 +3.38 +3.77 −4.79 −3.28]
State 011, Action S, Q-values: [−1.53 −2.29 −2.87 −2.23 −11.06 −12.18]
State 012, Action S, Q-values: [−3.86 −4.52 −4.24 −4.11 −12.39 −13.99]
State 013, Action S, Q-values: [−2.36 −2.99 −2.47 −2.67 −11.01 −12.98]
State 014, Action S, Q-values: [−3.68 −4.08 −3.73 −3.92 −12.44 −13.42]
State 015, Action W, Q-values: [+3.41 +3.23 +3.19 +3.60 −5.15 −2.81]
State 016, Action D, Q-values: [+13.78 +15.27 +14.37 +16.83 +7.57 +20.01]
State 017, Action E, Q-values: [+3.00 +1.74 +3.00 +2.20 −6.17 +0.29]
State 018, Action S, Q-values: [+9.76 +7.56 +7.25 +8.33 +0.27 +3.41]
State 019, Action E, Q-values: [+4.33 +3.25 +4.35 +3.58 −5.06 +0.50]
State 020, Action W, Q-values: [+3.15 +3.20 +3.09 +3.57 −5.05 −3.48]
State 021, Action W, Q-values: [−0.69 −0.13 −0.04 +0.48 −6.71 −8.06]
State 022, Action W, Q-values: [+3.22 +4.52 +4.18 +5.98 −3.59 −3.49]
State 023, Action W, Q-values: [+0.29 +1.03 +0.67 +1.62 −7.00 −7.05]
State 024, Action S, Q-values: [−4.40 −5.06 −5.20 −4.76 −14.13 −17.15]
State 025, Action W, Q-values: [+3.41 +3.45 +3.46 +3.74 −5.16 −3.01]
State 026, Action S, Q-values: [−4.46 −5.00 −4.79 −4.84 −13.77 −16.70]
State 027, Action S, Q-values: [−2.31 −2.92 −2.84 −3.00 −11.31 −13.31]
State 028, Action S, Q-values: [+0.51 −0.11 −0.61 +0.21 −9.43 −10.25]
State 029, Action S, Q-values: [−3.07 −3.29 −4.00 −3.07 −11.50 −13.37]
State 030, Action W, Q-values: [+3.32 +3.49 +3.37 +3.80 −4.76 −2.94]
State 031, Action W, Q-values: [−2.39 −2.47 −3.01 −2.30 −10.75 −12.46]
```

```
State 032, Action S, Q-values: [−3.09 −3.81 −3.73 −3.62 −11.99 −12.61]
State 033, Action S, Q-values: [−1.52 −2.16 −2.00 −2.37 −10.90 −12.92]
State 034, Action S, Q-values: [−3.17 −3.95 −4.15 −3.84 −11.63 −13.55]
State 035, Action W, Q-values: [+3.32 +3.35 +3.29 +3.65 −5.01 −3.01]
State 036, Action W, Q-values: [+13.05 +14.61 +14.53 +16.96 +3.87 +6.03]
State 037, Action S, Q-values: [+4.42 +3.54 +3.95 +2.67 −5.45 −7.21]
State 038, Action S, Q-values: [+7.47 +4.93 +5.70 +7.43 −3.20 −2.91]
State 039, Action S, Q-values: [+5.95 +4.46 +5.19 +4.91 −3.23 −2.89]
State 040, Action W, Q-values: [+3.40 +3.45 +3.36 +3.81 −4.82 −3.00]
State 041, Action S, Q-values: [−3.79 −4.22 −4.68 −4.10 −12.59 −14.59]
State 042, Action S, Q-values: [−0.47 −0.99 −1.45 −0.98 −10.18 −10.86]
State 043, Action S, Q-values: [−3.11 −3.99 −4.48 −3.72 −11.76 −13.48]
State 044, Action E, Q-values: [−1.80 −1.38 −0.53 −1.13 −9.51 −10.27]
State 045, Action W, Q-values: [+3.30 +3.41 +3.25 +3.73 −4.87 −3.08]
State 046, Action E, Q-values: [−0.89 −1.30 −0.53 −1.16 −9.90 −10.35]
State 047, Action E, Q-values: [+1.22 +2.21 +2.93 +2.56 −5.97 −5.35]
State 048, Action S, Q-values: [−0.58 −1.20 −1.57 −1.13 −10.65 −11.67]
State 049, Action S, Q-values: [−3.80 −4.17 −4.53 −4.07 −12.42 −14.36]
State 050, Action W, Q-values: [+3.38 +3.34 +3.24 +3.76 −4.98 −3.17]
State 051, Action S, Q-values: [−3.17 −3.59 −4.17 −3.39 −11.70 −12.69]
State 052, Action S, Q-values: [−2.28 −2.58 −2.34 −2.76 −10.34 −11.37]
State 053, Action S, Q-values: [−0.60 −1.41 −0.65 −1.08 −9.89 −11.34]
State 054, Action S, Q-values: [−2.42 −2.95 −2.43 −2.76 −10.74 −12.58]
State 055, Action W, Q-values: [+3.26 +3.28 +3.16 +3.62 −5.18 −3.39]
State 056, Action S, Q-values: [+6.05 +4.56 +4.13 +5.75 −3.55 −3.06]
State 057, Action E, Q-values: [+10.95 +12.65 +14.29 +11.94 +3.96 +4.41]
State 058, Action S, Q-values: [+5.99 +4.10 +4.16 +5.09 −3.85 −3.19]
State 059, Action S, Q-values: [+6.27 +4.32 +5.53 +5.83 −3.86 −3.52]
State 060, Action W, Q-values: [+3.02 +3.18 +3.11 +3.46 −5.10 −3.42]
State 061, Action W, Q-values: [−4.47 −4.68 −5.52 −4.44 −13.88 −16.70]
State 062, Action S, Q-values: [−1.44 −2.22 −2.82 −1.83 −10.30 −11.53]
State 063, Action S, Q-values: [−3.71 −4.28 −4.41 −4.01 −12.71 −14.44]
State 064, Action E, Q-values: [+0.05 −0.22 +0.46 −0.46 −8.03 −8.75]
State 065, Action W, Q-values: [+3.22 +3.16 +2.96 +3.49 −5.33 −3.67]
State 066, Action E, Q-values: [−0.51 −0.45 +0.42 −0.27 −7.07 −8.92]
State 067, Action E, Q-values: [+2.18 +3.34 +4.44 +2.80 −3.90 −3.49]
State 068, Action S, Q-values: [−1.53 −2.13 −2.65 −1.78 −10.22 −11.56]
State 069, Action W, Q-values: [−4.47 −4.69 −5.25 −4.38 −13.95 −16.32]
State 070, Action W, Q-values: [+3.34 +3.39 +3.21 +3.73 −5.14 −2.91]
State 071, Action W, Q-values: [−3.79 −4.30 −4.82 −3.75 −12.74 −14.78]
State 072, Action S, Q-values: [−1.48 −2.33 −2.70 −2.42 −11.11 −12.38]
State 073, Action S, Q-values: [+0.49 −0.18 −0.49 −0.69 −8.89 −10.15]
State 074, Action S, Q-values: [−1.53 −2.20 −2.26 −2.36 −9.89 −11.53]
State 075, Action W, Q-values: [+3.54 +3.51 +3.46 +3.83 −5.15 −3.07]
State 076, Action S, Q-values: [+4.35 +3.12 +2.43 +4.34 −4.75 −4.60]
State 077, Action E, Q-values: [+12.85 +15.35 +17.03 +13.53 +3.62 +5.89]
State 078, Action S, Q-values: [+4.40 +3.03 +2.48 +4.37 −4.83 −4.63]
State 079, Action S, Q-values: [+9.74 +8.00 +7.53 +7.26 −0.24 −0.11]
State 080, Action W, Q-values: [+3.34 +3.43 +3.22 +3.83 −5.09 −2.99]
State 081, Action S, Q-values: [−4.99 −5.51 −5.51 −5.05 −15.40 −18.64]
State 082, Action S, Q-values: [−2.26 −2.47 −3.05 −2.36 −10.08 −11.68]
State 083, Action W, Q-values: [−4.44 −4.92 −5.22 −4.42 −13.66 −16.73]
State 084, Action P, Q-values: [−0.34 +0.16 +0.32 +0.32 +1.67 −5.22]
State 085, Action W, Q-values: [+3.39 +3.41 +3.25 +3.72 −4.69 −3.03]
State 086, Action P, Q-values: [−0.34 +0.17 +0.51 +0.35 +1.60 −5.83]
State 087, Action P, Q-values: [+3.94 +4.19 +4.15 +4.27 +6.04 −2.71]
State 088, Action S, Q-values: [−2.44 −2.83 −3.15 −2.49 −10.81 −12.72]
```

```
State 089, Action W, Q-values: [-5.05 -5.48 -5.74 -4.88 -15.40 -18.46]
State 090, Action W, Q-values: [+3.28 +3.48 +3.37 +3.76 -4.94 -2.92]
State 091, Action W, Q-values: [-4.42 -4.53 -4.94 -4.37 -13.71 -16.23]
State 092, Action W, Q-values: [-2.36 -2.83 -2.94 -2.32 -10.96 -12.03]
State 093, Action S, Q-values: [-0.50 -1.02 -1.55 -0.78 -10.13 -11.43]
State 094, Action S, Q-values: [-2.11 -3.02 -2.83 -2.20 -11.24 -12.42]
State 095, Action W, Q-values: [+3.30 +3.34 +3.33 +3.76 -5.00 -3.04]
State 096, Action S, Q-values: [+2.93 +1.78 +2.04 +2.90 -6.95 -1.30]
State 097, Action D, Q-values: [+13.47 +15.36 +15.07 +15.13 +7.67 +20.01]
State 098, Action S, Q-values: [+3.01 +1.59 +2.06 +3.00 -6.40 -1.25]
State 099, Action S, Q-values: [+7.75 +5.89 +6.42 +7.70 -2.21 +3.07]
State 100, Action W, Q-values: [+3.38 +3.46 +3.36 +3.75 -5.22 -2.94]
State 101, Action N, Q-values: [-0.96 +0.45 -1.32 +0.09 -7.51 -7.72]
State 102, Action N, Q-values: [+3.10 +5.97 +4.97 +5.40 -3.41 -3.34]
State 103, Action N, Q-values: [-0.21 +1.60 -0.23 +0.97 -7.58 -7.98]
State 104, Action E, Q-values: [-4.45 -5.32 -4.40 -4.83 -13.54 -15.93]
State 105, Action W, Q-values: [+3.23 +3.46 +3.39 +3.76 -4.72 -2.95]
State 106, Action E, Q-values: [-4.44 -4.76 -4.33 -4.51 -12.99 -15.21]
State 107, Action S, Q-values: [-2.44 -2.92 -2.52 -2.73 -10.53 -11.29]
State 108, Action S, Q-values: [+2.99 +0.67 +1.38 +2.21 -6.77 -7.31]
State 109, Action S, Q-values: [-1.54 -2.66 -2.60 -2.00 -10.95 -11.54]
State 110, Action W, Q-values: [+3.29 +3.33 +3.19 +3.70 -5.24 -3.15]
State 111, Action S, Q-values: [-0.61 -2.15 -1.94 -1.24 -10.33 -10.65]
State 112, Action E, Q-values: [-3.21 -4.12 -3.12 -3.42 -11.37 -12.18]
State 113, Action S, Q-values: [-1.50 -2.83 -1.53 -2.05 -10.40 -11.05]
State 114, Action S, Q-values: [-3.17 -3.64 -3.29 -3.41 -11.99 -12.66]
State 115, Action W, Q-values: [+3.54 +3.66 +3.53 +4.02 -5.11 -2.66]
State 116, Action N, Q-values: [+12.99 +17.04 +15.56 +14.74 +4.62 +5.27]
State 117, Action S, Q-values: [+4.43 +2.29 +4.41 +3.79 -4.28 -4.48]
State 118, Action S, Q-values: [+11.93 +9.19 +9.36 +9.76 +3.25 +2.11]
State 119, Action S, Q-values: [+5.91 +3.65 +5.88 +5.03 -2.55 -2.82]
State 120, Action W, Q-values: [+3.15 +3.23 +3.12 +3.60 -4.74 -3.11]
State 121, Action N, Q-values: [-2.05 -0.59 -1.33 -0.72 -8.55 -9.34]
State 122, Action W, Q-values: [+2.05 +4.31 +3.27 +4.34 -4.96 -4.58]
State 123, Action N, Q-values: [-1.24 +0.42 -0.82 +0.18 -8.48 -8.55]
State 124, Action S, Q-values: [-3.74 -4.63 -4.60 -4.08 -12.39 -14.70]
State 125, Action W, Q-values: [+3.39 +3.39 +3.31 +3.77 -5.14 -3.29]
State 126, Action S, Q-values: [-3.80 -4.56 -4.44 -4.06 -13.03 -14.42]
State 127, Action S, Q-values: [-1.50 -2.29 -1.94 -2.26 -9.84 -10.38]
State 128, Action S, Q-values: [+1.65 -0.04 +0.50 +1.53 -7.68 -8.19]
State 129, Action S, Q-values: [-2.23 -2.78 -3.05 -2.26 -10.67 -12.06]
State 130, Action W, Q-values: [+3.35 +3.32 +3.24 +3.70 -4.87 -3.04]
State 131, Action S, Q-values: [-1.47 -1.90 -2.23 -1.53 -10.62 -11.20]
State 132, Action S, Q-values: [-2.35 -3.32 -3.09 -3.03 -10.92 -10.63]
State 133, Action S, Q-values: [-0.63 -1.70 -1.33 -1.82 -9.70 -10.53]
State 134, Action S, Q-values: [-2.41 -3.41 -2.94 -2.92 -11.12 -11.31]
State 135, Action W, Q-values: [+3.38 +3.43 +3.27 +3.73 -4.76 -2.93]
State 136, Action N, Q-values: [+10.93 +14.42 +13.17 +13.77 +3.65 +4.54]
State 137, Action S, Q-values: [+6.11 +3.35 +4.78 +4.64 -2.78 -3.50]
State 138, Action S, Q-values: [+9.76 +6.06 +7.58 +9.24 -1.35 -1.55]
State 139, Action S, Q-values: [+7.74 +4.75 +6.10 +6.74 -1.86 -2.14]
State 140, Action W, Q-values: [+3.59 +3.52 +3.35 +3.90 -4.58 -3.00]
State 141, Action S, Q-values: [-3.07 -3.91 -3.70 -3.28 -10.86 -11.83]
State 142, Action S, Q-values: [+0.46 -1.19 -1.22 -0.30 -8.25 -9.67]
State 143, Action S, Q-values: [-2.35 -3.22 -3.28 -2.69 -11.33 -11.64]
State 144, Action E, Q-values: [-2.37 -1.50 -1.49 -2.05 -9.50 -11.26]
State 145, Action W, Q-values: [+3.38 +3.37 +3.30 +3.68 -5.07 -3.09]
```

```
State 146, Action N, Q-values: [−2.86 −1.52 −1.64 −2.15 −8.92 −10.01]
State 147, Action E, Q-values: [−0.41 +1.67 +1.69 +0.62 −7.05 −7.33]
State 148, Action S, Q-values: [+0.48 −0.62 −1.32 +0.09 −8.91 −9.66]
State 149, Action S, Q-values: [−3.08 −3.33 −3.80 −3.35 −10.81 −11.56]
State 150, Action W, Q-values: [+3.39 +3.48 +3.42 +3.79 −5.01 −3.05]
State 151, Action S, Q-values: [−2.41 −3.12 −3.40 −2.71 −10.14 −10.09]
State 152, Action S, Q-values: [−1.48 −2.24 −1.51 −1.93 −10.07 −11.20]
State 153, Action S, Q-values: [+0.55 −0.77 −0.24 −0.51 −8.79 −9.62]
State 154, Action S, Q-values: [−1.55 −2.39 −1.56 −2.03 −9.81 −10.44]
State 155, Action W, Q-values: [+3.59 +3.45 +3.38 +3.82 −4.88 −2.97]
State 156, Action S, Q-values: [+7.84 +4.65 +5.01 +7.11 −2.60 −3.27]
State 157, Action E, Q-values: [+8.34 +11.89 +11.93 +10.17 +2.41 +2.00]
State 158, Action S, Q-values: [+7.76 +4.79 +5.31 +7.13 −1.26 −2.00]
State 159, Action S, Q-values: [+9.36 +5.49 +8.00 +8.77 −1.56 −1.24]
State 160, Action W, Q-values: [+3.45 +3.64 +3.55 +3.86 −4.95 −3.05]
State 161, Action W, Q-values: [−3.81 −4.39 −4.37 −3.75 −12.95 −14.12]
State 162, Action W, Q-values: [−0.60 −2.39 −1.87 −0.52 −9.23 −10.96]
State 163, Action W, Q-values: [−3.17 −4.30 −4.30 −3.06 −11.61 −13.11]
State 164, Action N, Q-values: [−1.73 −0.56 −0.56 −1.64 −8.69 −9.90]
State 165, Action W, Q-values: [+3.57 +3.45 +3.44 +3.76 −4.96 −2.83]
State 166, Action E, Q-values: [−2.09 −0.59 −0.58 −1.51 −8.65 −9.76]
State 167, Action E, Q-values: [+0.30 +2.85 +2.91 +1.01 −6.25 −6.55]
State 168, Action W, Q-values: [−0.81 −1.77 −1.89 −0.55 −9.57 −10.77]
State 169, Action W, Q-values: [−4.19 −4.14 −5.16 −3.80 −13.33 −14.99]
State 170, Action W, Q-values: [+3.35 +3.37 +3.21 +3.71 −4.94 −3.27]
State 171, Action W, Q-values: [−3.18 −3.72 −4.60 −3.10 −12.01 −13.62]
State 172, Action S, Q-values: [−0.59 −1.92 −2.15 −1.56 −10.30 −11.27]
State 173, Action S, Q-values: [+1.57 +0.71 +0.53 −0.18 −8.10 −8.92]
State 174, Action S, Q-values: [−0.59 −1.64 −1.80 −1.80 −9.55 −10.52]
State 175, Action W, Q-values: [+3.41 +3.44 +3.30 +3.81 −4.70 −2.96]
State 176, Action W, Q-values: [+6.00 +3.72 +3.81 +6.04 −3.82 −3.21]
State 177, Action N, Q-values: [+10.76 +14.26 +14.24 +11.37 +3.62 +4.30]
State 178, Action W, Q-values: [+5.84 +3.12 +3.73 +5.97 −3.36 −3.41]
State 179, Action S, Q-values: [+11.92 +8.37 +8.46 +8.90 +1.61 +1.19]
State 180, Action W, Q-values: [+3.21 +3.45 +3.35 +3.77 −4.87 −3.16]
State 181, Action W, Q-values: [−4.41 −4.70 −4.97 −4.32 −13.42 −15.85]
State 182, Action W, Q-values: [−1.41 −2.59 −2.20 −1.41 −10.69 −11.29]
State 183, Action W, Q-values: [−4.06 −4.46 −4.61 −3.76 −13.39 −14.88]
State 184, Action N, Q-values: [−1.28 +0.46 −0.40 −0.80 −8.66 −9.32]
State 185, Action W, Q-values: [+3.38 +3.57 +3.35 +3.86 −5.02 −2.86]
State 186, Action N, Q-values: [−1.02 +0.42 −0.21 −0.80 −7.59 −6.93]
State 187, Action N, Q-values: [+2.06 +4.38 +2.99 +3.43 −5.55 −4.95]
State 188, Action S, Q-values: [−1.45 −2.11 −2.29 −1.57 −10.43 −11.44]
State 189, Action W, Q-values: [−4.47 −5.06 −5.07 −4.38 −13.28 −16.36]
State 190, Action W, Q-values: [+3.45 +3.36 +3.19 +3.81 −5.05 −3.02]
State 191, Action W, Q-values: [−3.86 −4.06 −4.77 −3.76 −12.30 −15.07]
State 192, Action W, Q-values: [−1.59 −2.25 −2.42 −1.52 −10.38 −11.85]
State 193, Action S, Q-values: [+0.46 −0.93 −0.99 −0.11 −9.68 −11.06]
State 194, Action W, Q-values: [−1.52 −2.37 −2.36 −1.51 −10.53 −11.64]
State 195, Action W, Q-values: [+3.30 +3.25 +3.14 +3.52 −5.09 −3.16]
State 196, Action W, Q-values: [+4.39 +2.33 +3.09 +4.41 −4.63 −4.70]
State 197, Action N, Q-values: [+12.79 +17.03 +16.16 +13.65 +5.18 +5.62]
State 198, Action W, Q-values: [+4.42 +2.09 +2.96 +4.45 −4.52 −5.23]
State 199, Action S, Q-values: [+9.71 +5.85 +7.35 +9.68 −1.19 −1.28]
State 200, Action W, Q-values: [+3.51 +3.44 +3.29 +3.73 −5.05 −3.27]
State 201, Action N, Q-values: [−2.12 −0.56 −2.26 −1.26 −10.69 −11.12]
State 202, Action N, Q-values: [+2.05 +4.31 +2.84 +3.66 −5.13 −4.71]
```

```
State 203, Action N, Q-values: [-1.24 +0.43 -1.21 -0.21 -8.94 -9.66]
State 204, Action E, Q-values: [-4.85 -4.72 -3.78 -4.15 -12.32 -14.77]
State 205, Action W, Q-values: [+3.43 +3.38 +3.25 +3.60 -5.17 -3.08]
State 206, Action E, Q-values: [-4.75 -4.25 -3.84 -4.09 -11.99 -14.98]
State 207, Action E, Q-values: [-2.58 -2.49 -1.56 -2.03 -10.24 -11.54]
State 208, Action S, Q-values: [+4.35 +1.41 +1.74 +3.52 -5.51 -6.30]
State 209, Action S, Q-values: [-0.66 -2.05 -1.80 -1.48 -9.62 -9.81]
State 210, Action W, Q-values: [+3.32 +3.32 +3.19 +3.66 -5.16 -3.55]
State 211, Action S, Q-values: [+0.49 -0.53 -0.55 -0.38 -8.17 -8.78]
State 212, Action E, Q-values: [-3.37 -3.02 -2.36 -2.69 -10.51 -12.51]
State 213, Action E, Q-values: [-1.58 -1.85 -0.55 -1.19 -10.32 -11.32]
State 214, Action E, Q-values: [-3.07 -3.46 -2.41 -2.75 -10.77 -12.64]
State 215, Action W, Q-values: [+3.49 +3.50 +3.45 +3.83 -4.88 -3.27]
State 216, Action N, Q-values: [+10.64 +14.33 +12.65 +12.27 +3.12 +4.70]
State 217, Action E, Q-values: [+2.86 +4.08 +6.10 +4.66 -3.28 -2.88]
State 218, Action S, Q-values: [+14.30 +9.79 +10.76 +12.25 +3.29 +3.38]
State 219, Action E, Q-values: [+5.20 +5.07 +7.83 +6.54 -1.47 -1.37]
State 220, Action W, Q-values: [+3.42 +3.38 +3.25 +3.79 -5.03 -2.74]
State 221, Action W, Q-values: [-3.07 -1.45 -2.61 -1.41 -9.75 -10.33]
State 222, Action N, Q-values: [+0.71 +2.90 +1.35 +2.86 -6.34 -5.62]
State 223, Action W, Q-values: [-2.63 -0.59 -2.24 -0.56 -9.95 -10.05]
State 224, Action E, Q-values: [-4.17 -3.39 -3.10 -3.67 -11.04 -12.95]
State 225, Action W, Q-values: [+3.28 +3.41 +3.22 +3.71 -5.17 -2.88]
State 226, Action E, Q-values: [-3.82 -3.39 -3.09 -3.65 -10.84 -12.09]
State 227, Action E, Q-values: [-1.58 -1.22 -0.57 -1.33 -8.94 -10.18]
State 228, Action W, Q-values: [+1.62 +0.99 +0.81 +2.90 -5.90 -6.57]
State 229, Action W, Q-values: [-2.55 -2.91 -3.06 -1.54 -9.87 -11.26]
State 230, Action W, Q-values: [+3.40 +3.50 +3.47 +3.80 -5.30 -2.83]
State 231, Action W, Q-values: [-1.76 -1.38 -2.07 -0.53 -8.98 -10.10]
State 232, Action E, Q-values: [-2.84 -2.46 -1.56 -2.34 -10.56 -11.70]
State 233, Action E, Q-values: [+0.03 -0.91 +0.49 -0.65 -8.35 -9.93]
State 234, Action E, Q-values: [-2.46 -2.64 -1.53 -2.31 -10.32 -11.38]
State 235, Action W, Q-values: [+3.36 +3.40 +3.33 +3.69 -5.11 -3.07]
State 236, Action N, Q-values: [+7.85 +11.93 +9.99 +11.90 +1.14 +1.58]
State 237, Action E, Q-values: [+4.85 +6.73 +7.92 +5.71 -1.43 -1.27]
State 238, Action W, Q-values: [+8.37 +8.53 +8.75 +11.84 -0.45 +0.82]
State 239, Action E, Q-values: [+6.52 +6.90 +9.78 +7.08 -0.95 -1.40]
State 240, Action W, Q-values: [+3.44 +3.49 +3.37 +3.89 -5.38 -3.09]
State 241, Action W, Q-values: [-3.54 -2.89 -3.60 -2.24 -9.64 -11.07]
State 242, Action W, Q-values: [+0.32 -0.07 -0.52 +1.63 -6.82 -7.65]
State 243, Action W, Q-values: [-3.11 -2.62 -3.04 -1.49 -10.60 -11.47]
State 244, Action E, Q-values: [-3.14 -2.33 -2.31 -3.19 -9.84 -11.67]
State 245, Action W, Q-values: [+3.70 +3.46 +3.41 +3.87 -5.16 -3.14]
State 246, Action E, Q-values: [-3.43 -2.38 -2.34 -3.28 -9.54 -12.10]
State 247, Action E, Q-values: [-1.30 +0.45 +0.49 -0.81 -7.89 -8.78]
State 248, Action W, Q-values: [+0.30 -0.09 +0.07 +1.64 -6.95 -7.90]
State 249, Action W, Q-values: [-3.97 -3.55 -3.48 -2.27 -11.09 -10.66]
State 250, Action W, Q-values: [+3.43 +3.46 +3.32 +3.77 -5.05 -2.83]
State 251, Action W, Q-values: [-2.93 -2.81 -3.05 -1.50 -10.70 -11.04]
State 252, Action E, Q-values: [-1.88 -1.71 -0.61 -1.37 -9.49 -10.46]
State 253, Action E, Q-values: [+0.04 +0.25 +1.66 -0.00 -5.66 -7.17]
State 254, Action E, Q-values: [-1.45 -1.54 -0.57 -1.43 -9.09 -10.34]
State 255, Action W, Q-values: [+3.39 +3.49 +3.36 +3.77 -4.93 -2.89]
State 256, Action W, Q-values: [+6.33 +6.94 +6.81 +9.74 -1.43 -0.73]
State 257, Action N, Q-values: [+6.94 +9.74 +9.70 +6.89 +0.26 +0.86]
State 258, Action W, Q-values: [+6.54 +6.51 +6.54 +9.74 -1.06 -1.05]
State 259, Action E, Q-values: [+9.80 +9.71 +11.89 +9.12 +2.26 +1.88]
```

```
State 260, Action W, Q-values: [+3.31 +3.46 +3.34 +3.74 −4.71 −3.14]
State 261, Action W, Q-values: [−4.20 −3.94 −4.40 −3.06 −12.14 −12.93]
State 262, Action W, Q-values: [−1.53 −0.26 −1.33 +0.44 −8.16 −9.42]
State 263, Action W, Q-values: [−3.85 −3.60 −3.93 −2.39 −11.06 −11.59]
State 264, Action N, Q-values: [−2.53 −1.50 −1.51 −2.16 −8.74 −10.05]
State 265, Action W, Q-values: [+3.16 +3.19 +3.14 +3.53 −5.33 −3.35]
State 266, Action N, Q-values: [−2.59 −1.50 −1.51 −2.48 −10.05 −11.05]
State 267, Action N, Q-values: [−0.40 +1.59 +1.57 +0.11 −7.63 −8.16]
State 268, Action W, Q-values: [−1.16 −0.95 −1.74 +0.48 −8.69 −10.79]
State 269, Action W, Q-values: [−4.39 −4.46 −4.48 −3.08 −12.23 −13.55]
State 270, Action W, Q-values: [+3.31 +3.31 +3.18 +3.67 −5.02 −3.24]
State 271, Action W, Q-values: [−3.53 −3.83 −3.81 −2.38 −11.11 −11.18]
State 272, Action S, Q-values: [+0.43 −0.75 −1.04 −0.99 −8.58 −9.68]
State 273, Action S, Q-values: [+2.92 +0.77 +1.10 +0.74 −6.29 −8.85]
State 274, Action S, Q-values: [+0.44 −0.75 −1.15 −0.81 −7.65 −8.58]
State 275, Action W, Q-values: [+3.50 +3.50 +3.43 +3.87 −5.10 −3.03]
State 276, Action W, Q-values: [+3.75 +5.48 +5.18 +7.81 −2.19 −1.78]
State 277, Action N, Q-values: [+8.60 +11.89 +11.83 +9.13 +2.77 +1.95]
State 278, Action W, Q-values: [+4.79 +4.80 +5.26 +7.72 −1.63 −1.88]
State 279, Action S, Q-values: [+14.29 +9.02 +10.62 +11.38 +3.85 +4.10]
State 280, Action W, Q-values: [+3.16 +3.27 +3.13 +3.61 −5.37 −3.49]
State 281, Action W, Q-values: [−4.43 −4.24 −4.41 −3.75 −12.71 −13.77]
State 282, Action W, Q-values: [−2.05 −1.87 −1.53 −0.50 −11.11 −11.93]
State 283, Action W, Q-values: [−3.87 −4.40 −4.44 −3.10 −12.11 −13.99]
State 284, Action N, Q-values: [−1.85 −0.52 −1.42 −1.67 −10.41 −10.38]
State 285, Action W, Q-values: [+3.27 +3.35 +3.24 +3.62 −4.87 −3.38]
State 286, Action N, Q-values: [−2.26 −0.62 −1.31 −1.78 −9.82 −10.17]
State 287, Action N, Q-values: [+0.26 +2.84 +1.69 +1.39 −7.31 −7.02]
State 288, Action W, Q-values: [−1.74 −1.84 −1.65 −0.55 −10.02 −11.67]
State 289, Action W, Q-values: [−4.77 −4.08 −4.33 −3.77 −12.62 −14.38]
State 290, Action W, Q-values: [+3.42 +3.45 +3.43 +3.75 −4.97 −2.87]
State 291, Action W, Q-values: [−3.71 −3.86 −3.96 −3.09 −11.61 −12.45]
State 292, Action S, Q-values: [−0.56 −1.54 −1.61 −0.58 −9.64 −10.80]
State 293, Action S, Q-values: [+1.67 +0.11 +0.43 +1.47 −8.17 −8.08]
State 294, Action W, Q-values: [−0.74 −1.89 −1.75 −0.58 −9.28 −10.61]
State 295, Action W, Q-values: [+3.03 +3.28 +3.08 +3.66 −5.08 −3.10]
State 296, Action W, Q-values: [+2.88 +3.68 +4.40 +6.02 −4.59 −3.42]
State 297, Action N, Q-values: [+10.80 +14.31 +13.22 +11.43 +2.89 +4.51]
State 298, Action W, Q-values: [+3.09 +3.61 +4.36 +5.98 −2.59 −2.83]
State 299, Action S, Q-values: [+11.91 +9.23 +9.92 +10.77 +1.29 +1.87]
State 300, Action W, Q-values: [+3.38 +3.48 +3.28 +3.77 −5.18 −3.16]
State 301, Action N, Q-values: [−3.02 −1.48 −2.15 −2.07 −10.99 −12.14]
State 302, Action N, Q-values: [+0.80 +2.91 +1.44 +2.67 −7.21 −6.81]
State 303, Action N, Q-values: [−2.38 −0.60 −1.55 −1.28 −11.06 −11.62]
State 304, Action N, Q-values: [−5.15 −4.47 −5.23 −4.74 −14.40 −16.61]
State 305, Action W, Q-values: [+3.22 +3.24 +3.13 +3.54 −5.19 −3.32]
State 306, Action N, Q-values: [−5.39 −4.48 −5.16 −4.84 −14.29 −16.66]
State 307, Action N, Q-values: [−3.44 −2.52 −2.94 −2.97 −10.38 −11.76]
State 308, Action S, Q-values: [+5.98 +3.69 +4.73 +5.64 −3.29 −2.89]
State 309, Action S, Q-values: [+0.42 −0.70 −0.14 −0.65 −6.75 −7.63]
State 310, Action W, Q-values: [+3.46 +3.30 +3.20 +3.71 −5.24 −3.33]
State 311, Action S, Q-values: [+1.63 +0.29 +1.02 +0.60 −5.73 −6.95]
State 312, Action N, Q-values: [−3.98 −3.09 −3.73 −3.41 −11.67 −12.40]
State 313, Action N, Q-values: [−2.59 −1.54 −2.04 −2.01 −9.73 −10.57]
State 314, Action N, Q-values: [−3.83 −3.23 −4.01 −3.68 −11.81 −12.90]
State 315, Action W, Q-values: [+3.12 +3.32 +3.23 +3.68 −5.18 −2.91]
State 316, Action N, Q-values: [+8.83 +11.91 +10.42 +10.48 +2.44 +3.60]
```

```
State 317, Action N, Q-values: [+1.45 +4.42 +3.45 +3.63 -5.14 -5.95]
State 318, Action S, Q-values: [+17.02 +12.50 +14.37 +14.56 +5.26 +5.46]
State 319, Action N, Q-values: [+2.92 +5.99 +5.73 +5.34 -2.81 -3.08]
State 320, Action W, Q-values: [+3.48 +3.35 +3.26 +3.64 -4.94 -3.08]
State 321, Action N, Q-values: [-3.28 -2.31 -3.20 -3.00 -11.32 -11.74]
State 322, Action N, Q-values: [-0.08 +1.61 -0.53 +1.03 -9.09 -9.33]
State 323, Action N, Q-values: [-3.14 -1.54 -3.06 -1.92 -10.98 -12.19]
State 324, Action N, Q-values: [-4.46 -3.84 -4.30 -4.20 -12.07 -14.22]
State 325, Action W, Q-values: [+3.29 +3.54 +3.45 +3.82 -5.03 -2.77]
State 326, Action N, Q-values: [-4.51 -3.85 -4.10 -4.31 -12.04 -14.48]
State 327, Action N, Q-values: [-2.60 -1.61 -1.78 -2.16 -10.18 -10.83]
State 328, Action N, Q-values: [-0.14 +1.59 +0.03 +1.14 -8.62 -8.14]
State 329, Action N, Q-values: [-3.74 -2.40 -3.37 -2.66 -10.87 -11.73]
State 330, Action W, Q-values: [+3.03 +3.29 +3.19 +3.59 -5.02 -3.11]
State 331, Action N, Q-values: [-3.07 -1.56 -2.55 -2.01 -9.99 -10.91]
State 332, Action E, Q-values: [-3.17 -2.39 -2.35 -2.75 -9.92 -12.08]
State 333, Action N, Q-values: [-2.04 -0.57 -0.64 -1.31 -9.70 -10.87]
State 334, Action N, Q-values: [-3.16 -2.46 -2.72 -2.87 -11.04 -11.99]
State 335, Action W, Q-values: [+3.22 +3.41 +3.30 +3.69 -4.96 -3.17]
State 336, Action N, Q-values: [+5.96 +9.58 +8.23 +9.32 -1.78 -2.18]
State 337, Action E, Q-values: [+3.60 +5.48 +5.76 +5.03 -4.78 -3.85]
State 338, Action N, Q-values: [+5.78 +9.42 +8.12 +9.10 -1.84 -2.50]
State 339, Action E, Q-values: [+4.48 +7.24 +7.71 +6.77 -2.63 -2.25]
State 340, Action W, Q-values: [+3.35 +3.23 +3.19 +3.51 -5.01 -3.16]
State 341, Action N, Q-values: [-3.79 -3.09 -3.91 -3.11 -11.06 -12.84]
State 342, Action N, Q-values: [-1.30 +0.54 -0.77 +0.50 -9.04 -9.13]
State 343, Action W, Q-values: [-3.44 -2.34 -3.09 -2.26 -10.71 -11.84]
State 344, Action N, Q-values: [-4.10 -3.08 -3.74 -3.88 -11.41 -13.12]
State 345, Action W, Q-values: [+3.53 +3.58 +3.50 +3.85 -4.62 -2.66]
State 346, Action N, Q-values: [-4.00 -3.11 -3.45 -3.85 -10.90 -12.51]
State 347, Action N, Q-values: [-2.12 -0.64 -1.21 -1.35 -10.50 -10.98]
State 348, Action N, Q-values: [-1.36 +0.42 -1.03 +0.32 -9.42 -10.05]
State 349, Action N, Q-values: [-3.85 -3.07 -3.68 -3.10 -10.90 -12.07]
State 350, Action W, Q-values: [+3.39 +3.51 +3.39 +3.84 -4.54 -2.96]
State 351, Action N, Q-values: [-3.46 -2.39 -3.14 -2.64 -10.55 -12.03]
State 352, Action N, Q-values: [-2.67 -1.53 -2.13 -2.32 -10.47 -10.92]
State 353, Action N, Q-values: [-1.21 +0.45 -0.12 -0.64 -8.93 -10.02]
State 354, Action N, Q-values: [-2.74 -1.51 -1.99 -2.27 -9.73 -10.79]
State 355, Action W, Q-values: [+3.56 +3.76 +3.57 +4.02 -4.69 -2.42]
State 356, Action W, Q-values: [+4.24 +7.30 +6.38 +7.64 -2.76 -1.75]
State 357, Action N, Q-values: [+4.72 +7.61 +7.16 +5.81 -2.55 -1.98]
State 358, Action W, Q-values: [+3.69 +6.66 +5.83 +7.13 -3.15 -2.58]
State 359, Action N, Q-values: [+6.52 +9.65 +8.98 +7.73 -0.12 +0.60]
State 360, Action W, Q-values: [+3.45 +3.44 +3.34 +3.81 -4.96 -3.10]
State 361, Action N, Q-values: [-4.46 -3.82 -4.64 -4.24 -13.40 -14.78]
State 362, Action N, Q-values: [-2.20 -0.51 -1.71 -0.83 -10.33 -10.74]
State 363, Action N, Q-values: [-4.00 -3.13 -3.85 -3.47 -11.35 -12.08]
State 364, Action N, Q-values: [-3.57 -2.30 -2.53 -2.87 -11.05 -12.06]
State 365, Action W, Q-values: [+3.30 +3.27 +3.10 +3.60 -5.12 -3.21]
State 366, Action N, Q-values: [-3.35 -2.14 -2.31 -2.72 -10.55 -11.64]
State 367, Action E, Q-values: [-1.72 +0.40 +0.43 -0.22 -8.84 -9.81]
State 368, Action N, Q-values: [-2.12 -0.60 -1.81 -1.07 -9.81 -10.75]
State 369, Action N, Q-values: [-4.57 -3.88 -4.92 -4.05 -12.88 -14.60]
State 370, Action W, Q-values: [+3.39 +3.46 +3.29 +3.75 -4.94 -2.94]
State 371, Action N, Q-values: [-4.19 -3.16 -4.04 -3.36 -11.24 -12.26]
State 372, Action S, Q-values: [+1.62 +1.19 -0.03 +1.01 -7.47 -7.70]
State 373, Action S, Q-values: [+4.40 +1.63 +2.17 +3.48 -4.40 -5.86]
```

```
State 374, Action S, Q-values: [+1.61 +0.34 −0.06 +0.67 −5.54 −7.35]
State 375, Action W, Q-values: [+3.62 +3.66 +3.56 +3.97 −4.60 −2.72]
State 376, Action N, Q-values: [+3.31 +6.07 +5.39 +5.56 −3.11 −2.90]
State 377, Action N, Q-values: [+6.44 +9.82 +9.73 +8.91 −1.09 −1.10]
State 378, Action N, Q-values: [+2.98 +6.02 +4.98 +5.25 −3.61 −4.02]
State 379, Action S, Q-values: [+17.01 +12.24 +13.14 +15.00 +5.46 +6.81]
State 380, Action W, Q-values: [+3.32 +3.25 +3.16 +3.49 −5.23 −3.40]
State 381, Action W, Q-values: [−4.76 −4.45 −5.07 −4.36 −13.96 −15.93]
State 382, Action N, Q-values: [−2.99 −1.45 −2.38 −1.67 −10.07 −11.09]
State 383, Action W, Q-values: [−4.27 −3.87 −4.40 −3.75 −12.22 −14.29]
State 384, Action N, Q-values: [−2.77 −1.47 −2.18 −2.37 −11.03 −12.08]
State 385, Action W, Q-values: [+3.18 +3.21 +3.10 +3.53 −4.93 −3.19]
State 386, Action N, Q-values: [−2.88 −1.54 −2.52 −2.44 −11.91 −12.82]
State 387, Action N, Q-values: [−0.68 +1.48 +0.87 +0.12 −8.12 −8.66]
State 388, Action N, Q-values: [−2.88 −1.49 −2.31 −1.56 −10.49 −11.48]
State 389, Action W, Q-values: [−4.61 −4.45 −5.03 −4.39 −13.47 −15.84]
State 390, Action W, Q-values: [+3.43 +3.45 +3.41 +3.77 −4.75 −2.73]
State 391, Action N, Q-values: [−4.42 −3.86 −4.64 −3.91 −12.30 −14.44]
State 392, Action S, Q-values: [+0.55 −1.33 −0.75 +0.10 −8.68 −9.32]
State 393, Action S, Q-values: [+2.96 +1.36 +1.82 +2.94 −6.24 −6.18]
State 394, Action W, Q-values: [+0.46 −1.06 −0.53 +0.49 −8.41 −9.93]
State 395, Action W, Q-values: [+3.26 +3.31 +3.15 +3.66 −5.29 −3.26]
State 396, Action W, Q-values: [+2.03 +4.20 +3.34 +4.23 −6.12 −5.57]
State 397, Action N, Q-values: [+8.60 +11.92 +10.68 +8.99 +1.15 +1.58]
State 398, Action W, Q-values: [+2.13 +5.09 +4.19 +5.10 −5.28 −4.96]
State 399, Action S, Q-values: [+14.36 +11.03 +12.30 +12.60 +3.77 +3.80]
State 400, Action W, Q-values: [+3.23 +3.36 +3.19 +3.53 −5.14 −2.95]
State 401, Action N, Q-values: [−3.29 −2.38 −3.20 −2.75 −12.18 −13.25]
State 402, Action N, Q-values: [+0.74 +1.66 +0.96 +0.88 −8.34 −9.03]
State 403, Action N, Q-values: [−2.36 −1.67 −2.32 −2.17 −10.65 −11.38]
State 404, Action N, Q-values: [−5.77 −4.99 −5.43 −5.36 −16.28 −18.69]
State 405, Action W, Q-values: [+3.18 +3.34 +3.23 +3.57 −5.22 −3.39]
State 406, Action N, Q-values: [−5.79 −4.99 −5.48 −5.34 −16.80 −18.78]
State 407, Action N, Q-values: [−3.71 −3.12 −3.84 −3.36 −12.63 −12.47]
State 408, Action P, Q-values: [+5.96 +5.43 +5.72 +6.47 +7.79 −3.18]
State 409, Action P, Q-values: [−0.05 −0.10 +0.39 +0.49 +1.65 −5.63]
State 410, Action W, Q-values: [+3.28 +3.51 +3.40 +3.76 −5.15 −3.16]
State 411, Action P, Q-values: [+1.30 +0.83 +1.02 +1.71 +2.91 −6.78]
State 412, Action N, Q-values: [−4.42 −3.81 −4.49 −4.02 −12.59 −14.47]
State 413, Action N, Q-values: [−3.12 −2.50 −3.14 −2.89 −11.78 −13.21]
State 414, Action N, Q-values: [−4.65 −3.99 −4.77 −4.36 −13.35 −15.03]
State 415, Action W, Q-values: [+3.39 +3.52 +3.42 +3.77 −5.12 −3.04]
State 416, Action N, Q-values: [+7.62 +9.72 +8.56 +8.63 +0.47 +5.61]
State 417, Action N, Q-values: [+1.67 +2.99 +2.52 +2.20 −7.51 −0.50]
State 418, Action D, Q-values: [+16.10 +14.53 +15.74 +17.05 +7.99 +20.01]
State 419, Action N, Q-values: [+2.64 +4.35 +3.88 +3.66 −5.98 −0.91]
State 420, Action W, Q-values: [+3.29 +3.46 +3.32 +3.75 −5.01 −3.02]
State 421, Action N, Q-values: [−3.79 −3.14 −4.17 −3.65 −12.18 −13.84]
State 422, Action N, Q-values: [−0.39 +0.48 −0.93 −0.22 −10.11 −10.75]
State 423, Action N, Q-values: [−3.10 −2.34 −3.49 −2.83 −11.67 −12.91]
State 424, Action N, Q-values: [−4.71 −4.46 −4.58 −4.74 −13.26 −15.86]
State 425, Action W, Q-values: [+3.52 +3.39 +3.33 +3.86 −4.80 −3.11]
State 426, Action N, Q-values: [−5.08 −4.40 −4.70 −4.79 −14.05 −16.63]
State 427, Action N, Q-values: [−3.02 −2.36 −2.55 −2.99 −10.76 −13.66]
State 428, Action N, Q-values: [−0.49 +0.46 −1.38 −0.06 −9.58 −9.98]
State 429, Action N, Q-values: [−4.17 −3.17 −4.31 −3.50 −12.99 −13.59]
State 430, Action W, Q-values: [+3.35 +3.21 +3.17 +3.60 −5.07 −3.17]
```

```
State 431, Action N, Q-values: [-3.04 -2.40 -3.27 -2.87 -10.71 -12.46]
State 432, Action N, Q-values: [-3.79 -3.26 -3.44 -3.64 -11.23 -13.12]
State 433, Action N, Q-values: [-2.85 -1.51 -2.00 -2.17 -11.25 -12.86]
State 434, Action N, Q-values: [-3.54 -3.05 -3.38 -3.44 -10.51 -12.15]
State 435, Action W, Q-values: [+3.61 +3.65 +3.58 +3.94 -4.77 -2.67]
State 436, Action N, Q-values: [+4.99 +6.21 +5.17 +5.65 -4.31 -3.40]
State 437, Action E, Q-values: [+3.21 +3.71 +3.78 +3.68 -6.03 -5.31]
State 438, Action N, Q-values: [+4.82 +5.78 +4.81 +5.46 -4.46 -3.53]
State 439, Action E, Q-values: [+3.71 +4.22 +4.49 +4.22 -5.30 -4.78]
State 440, Action W, Q-values: [+3.32 +3.28 +3.16 +3.64 -4.99 -3.03]
State 441, Action W, Q-values: [-3.99 -3.80 -4.33 -3.80 -12.03 -14.12]
State 442, Action N, Q-values: [-2.05 -0.56 -1.53 -0.83 -10.29 -10.60]
State 443, Action W, Q-values: [-3.11 -3.07 -3.64 -3.02 -11.03 -13.62]
State 444, Action N, Q-values: [-4.28 -3.82 -4.51 -4.22 -12.27 -14.21]
State 445, Action W, Q-values: [+3.17 +3.38 +3.23 +3.62 -5.12 -3.31]
State 446, Action N, Q-values: [-4.47 -3.80 -4.58 -4.36 -13.34 -15.14]
State 447, Action N, Q-values: [-2.50 -1.52 -2.13 -2.28 -11.29 -12.03]
State 448, Action W, Q-values: [-1.18 -0.87 -1.79 -0.55 -9.82 -11.78]
State 449, Action N, Q-values: [-4.42 -3.81 -4.55 -3.97 -12.85 -14.84]
State 450, Action W, Q-values: [+3.65 +3.68 +3.57 +3.90 -4.83 -2.40]
State 451, Action N, Q-values: [-3.73 -3.23 -3.76 -3.42 -11.43 -12.22]
State 452, Action N, Q-values: [-3.11 -2.47 -3.20 -3.15 -11.58 -12.75]
State 453, Action N, Q-values: [-1.88 -0.65 -1.15 -1.59 -10.19 -10.70]
State 454, Action N, Q-values: [-2.58 -2.40 -2.89 -2.90 -11.18 -12.94]
State 455, Action W, Q-values: [+3.29 +3.35 +3.22 +3.66 -5.43 -3.11]
State 456, Action W, Q-values: [+4.09 +4.63 +4.15 +5.24 -5.17 -4.10]
State 457, Action N, Q-values: [+3.84 +4.57 +4.26 +4.41 -6.22 -4.90]
State 458, Action N, Q-values: [+4.08 +5.30 +4.70 +5.02 -4.46 -3.63]
State 459, Action N, Q-values: [+4.33 +5.88 +5.39 +4.87 -4.22 -3.58]
State 460, Action W, Q-values: [+3.10 +3.21 +3.08 +3.46 -5.10 -3.46]
State 461, Action N, Q-values: [-5.13 -4.39 -5.25 -4.83 -14.81 -17.02]
State 462, Action N, Q-values: [-2.69 -1.53 -2.43 -2.05 -11.05 -11.42]
State 463, Action N, Q-values: [-4.57 -3.93 -4.74 -4.18 -13.64 -14.89]
State 464, Action N, Q-values: [-3.56 -3.08 -3.26 -3.37 -11.46 -12.45]
State 465, Action W, Q-values: [+3.26 +3.40 +3.26 +3.65 -4.92 -3.15]
State 466, Action N, Q-values: [-3.80 -3.24 -3.78 -3.75 -12.54 -13.28]
State 467, Action N, Q-values: [-1.29 -0.57 -0.58 -1.11 -9.61 -10.19]
State 468, Action N, Q-values: [-1.87 -1.65 -3.06 -2.03 -11.32 -12.44]
State 469, Action N, Q-values: [-5.23 -4.51 -5.34 -4.85 -14.85 -16.74]
State 470, Action W, Q-values: [+3.44 +3.52 +3.35 +3.88 -5.08 -3.06]
State 471, Action N, Q-values: [-4.48 -3.85 -4.78 -4.21 -12.74 -14.76]
State 472, Action P, Q-values: [+0.81 +0.91 +0.80 +1.79 +2.94 -6.16]
State 473, Action P, Q-values: [+4.35 +3.66 +3.89 +4.74 +6.01 -4.12]
State 474, Action P, Q-values: [+0.86 +0.96 +1.02 +1.82 +2.94 -6.18]
State 475, Action W, Q-values: [+3.15 +3.38 +3.28 +3.66 -4.97 -3.27]
State 476, Action N, Q-values: [+2.86 +4.42 +3.49 +4.13 -6.25 +1.17]
State 477, Action N, Q-values: [+6.36 +7.85 +7.77 +6.70 -1.56 +3.24]
State 478, Action N, Q-values: [+2.88 +4.44 +3.23 +3.98 -6.04 -1.04]
State 479, Action D, Q-values: [+16.10 +14.19 +14.25 +16.62 +7.91 +19.99]
State 480, Action W, Q-values: [+3.37 +3.41 +3.35 +3.79 -4.81 -2.94]
State 481, Action N, Q-values: [-5.42 -4.99 -5.23 -5.14 -15.63 -18.09]
State 482, Action N, Q-values: [-3.16 -2.33 -3.11 -2.60 -12.01 -12.61]
State 483, Action N, Q-values: [-4.96 -4.50 -5.26 -4.55 -14.31 -16.29]
State 484, Action N, Q-values: [-2.97 -2.33 -3.08 -3.03 -11.25 -13.04]
State 485, Action W, Q-values: [+3.06 +3.30 +3.14 +3.67 -5.17 -3.52]
State 486, Action N, Q-values: [-3.25 -2.34 -2.80 -2.99 -12.12 -12.98]
State 487, Action N, Q-values: [-0.59 +0.43 -0.02 -0.32 -9.46 -10.04]
```

```
State 488, Action W, Q-values: [-3.12 -2.42 -3.38 -2.40 -11.28 -13.57]
State 489, Action W, Q-values: [-5.27 -4.93 -5.60 -4.92 -15.51 -18.25]
State 490, Action W, Q-values: [+3.39 +3.52 +3.45 +3.95 -5.08 -2.77]
State 491, Action W, Q-values: [-4.87 -4.51 -4.95 -4.43 -13.81 -16.02]
State 492, Action W, Q-values: [+1.03 -0.17 +0.52 +1.69 -6.27 -7.77]
State 493, Action W, Q-values: [+3.08 +2.74 +3.15 +4.42 -5.13 -4.30]
State 494, Action W, Q-values: [+0.83 +0.67 -0.15 +1.66 -7.55 -8.54]
State 495, Action W, Q-values: [+3.21 +3.55 +3.40 +3.78 -4.87 -3.04]
State 496, Action W, Q-values: [+2.10 +3.30 +2.74 +3.38 -7.21 -6.25]
State 497, Action N, Q-values: [+7.76 +9.71 +8.81 +7.25 -0.58 +0.04]
State 498, Action W, Q-values: [+2.42 +3.12 +2.62 +3.21 -6.82 -6.17]
State 499, Action W, Q-values: [+14.82 +13.24 +14.46 +16.94 +4.45 +7.58]
episodes: 0, acton: 1, state: 454
episodes: 0, acton: 1, state: 354
episodes: 0, acton: 2, state: 254
episodes: 0, acton: 0, state: 274
episodes: 0, acton: 0, state: 374
episodes: 0, acton: 4, state: 474
episodes: 0, acton: 1, state: 478
episodes: 0, acton: 1, state: 378
episodes: 0, acton: 3, state: 278
episodes: 0, acton: 3, state: 258
episodes: 0, acton: 3, state: 238
episodes: 0, acton: 0, state: 218
episodes: 0, acton: 0, state: 318
episodes: 0, acton: 5, state: 418
Episode 1: Total Reward = 7
episodes: 1, acton: 0, state: 12
episodes: 1, acton: 2, state: 112
episodes: 1, acton: 0, state: 132
episodes: 1, acton: 2, state: 232
episodes: 1, acton: 2, state: 252
episodes: 1, acton: 0, state: 272
episodes: 1, acton: 0, state: 372
episodes: 1, acton: 4, state: 472
episodes: 1, acton: 1, state: 476
episodes: 1, acton: 1, state: 376
episodes: 1, acton: 3, state: 276
episodes: 1, acton: 3, state: 256
episodes: 1, acton: 1, state: 236
episodes: 1, acton: 1, state: 136
episodes: 1, acton: 3, state: 36
episodes: 1, acton: 5, state: 16
Episode 2: Total Reward = 5
episodes: 2, acton: 3, state: 189
episodes: 2, acton: 3, state: 169
episodes: 2, acton: 0, state: 149
episodes: 2, acton: 3, state: 249
episodes: 2, acton: 3, state: 229
episodes: 2, acton: 0, state: 209
episodes: 2, acton: 0, state: 309
episodes: 2, acton: 4, state: 409
episodes: 2, acton: 1, state: 417
episodes: 2, acton: 1, state: 317
episodes: 2, acton: 2, state: 217
episodes: 2, acton: 2, state: 237
episodes: 2, acton: 1, state: 257
```

```
episodes: 2, acton: 2, state: 157
episodes: 2, acton: 1, state: 177
episodes: 2, acton: 2, state: 77
episodes: 2, acton: 5, state: 97
Episode 3: Total Reward = 4
episodes: 3, acton: 1, state: 307
episodes: 3, acton: 2, state: 207
episodes: 3, acton: 2, state: 227
episodes: 3, acton: 2, state: 247
episodes: 3, acton: 1, state: 267
episodes: 3, acton: 2, state: 167
episodes: 3, acton: 1, state: 187
episodes: 3, acton: 4, state: 87
episodes: 3, acton: 0, state: 99
episodes: 3, acton: 0, state: 199
episodes: 3, acton: 0, state: 299
episodes: 3, acton: 0, state: 399
episodes: 3, acton: 3, state: 499
episodes: 3, acton: 5, state: 479
Episode 4: Total Reward = 7
episodes: 4, acton: 3, state: 261
episodes: 4, acton: 3, state: 241
episodes: 4, acton: 3, state: 221
episodes: 4, acton: 1, state: 201
episodes: 4, acton: 1, state: 101
episodes: 4, acton: 4, state: 1
episodes: 4, acton: 2, state: 17
episodes: 4, acton: 0, state: 37
episodes: 4, acton: 0, state: 137
episodes: 4, acton: 2, state: 237
episodes: 4, acton: 1, state: 257
episodes: 4, acton: 2, state: 157
episodes: 4, acton: 1, state: 177
episodes: 4, acton: 2, state: 77
episodes: 4, acton: 5, state: 97
Episode 5: Total Reward = 6
Saving gif...Done!
Average Reward across all Episodes: = 5.8
```

In [87]:
```python
# display the captured GIF
display.Image(data=open("./gifs/taxi_dqn_fine_tuned.gif",'rb').read(), fo
# "taxi_dqn_fine_tuned"
```

Out[87]:   `<IPython.core.display.Image object>`

If the above display function is not working, here is the previously prepared taxi_dqn_fine_tuned.gif
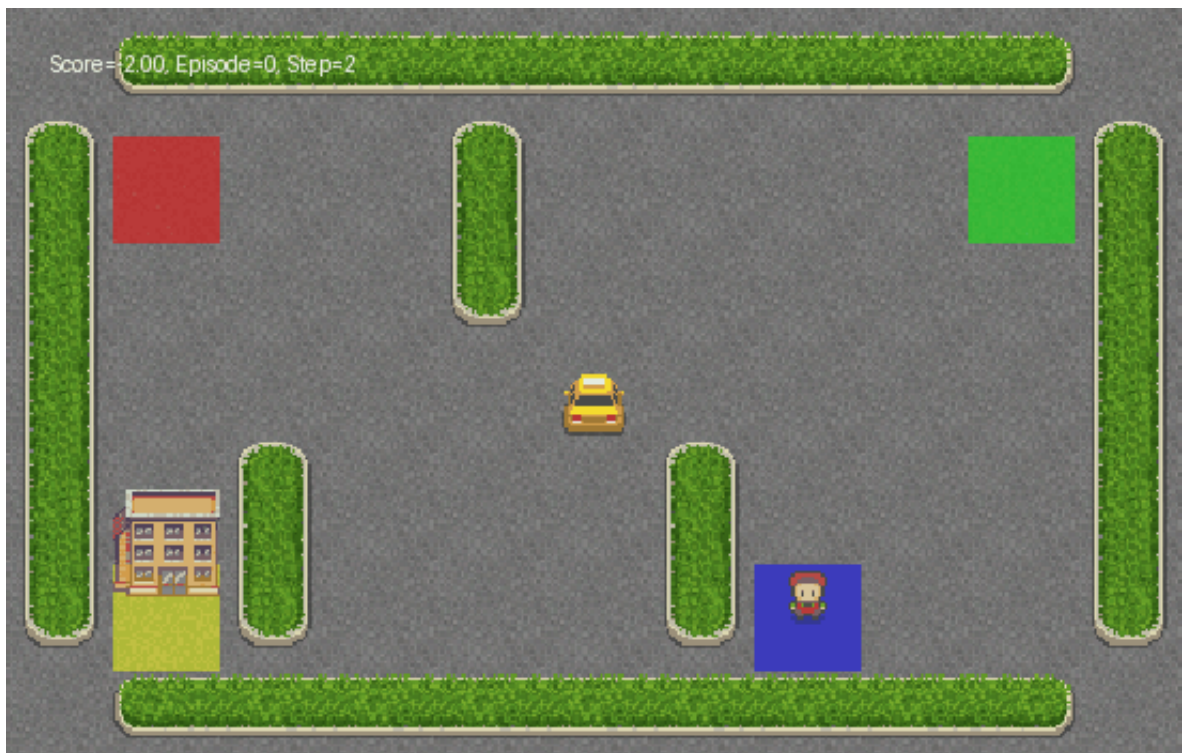
Compare the prediction of fine tuned models

```
In [90]:  # my_q_agent_fine_tuned_avg_reward_lst )
          # my_dqn_agent_fine_tuned_avg_reward_lst)
```

```
In [91]:  x = range(1,1000 + 1)

          plt.figure(figsize=(10, 6))
          plt.plot(x, my_dqn_agent_fine_tuned_avg_reward_lst, label='DQN', marker='
          plt.plot(x, my_q_agent_fine_tuned_avg_reward_lst , label='Q-learning', ma

          # titles and labels
          plt.title('Prediction Comparison (Fine Tuned Unique Parameters)')
          plt.xlabel('Time Episodes')
          plt.ylabel('Average Reward')
          plt.legend()
          # display
          plt.grid(True)
          plt.show()
```

Prediction Comparison (Fine Tuned Unique Parameters)
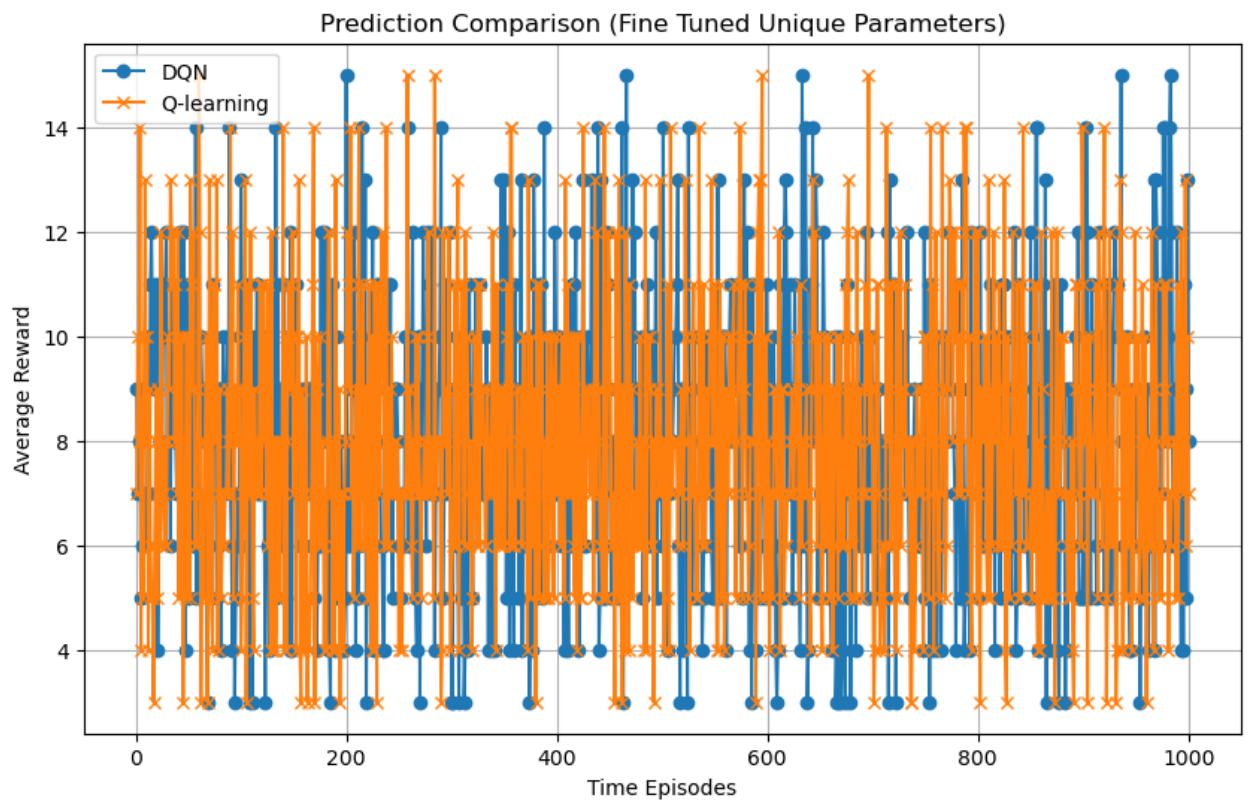
Discussion

Controlled Parameters Limitations in Training and Performance

The initial approach of using controlled parameters for both Q-learning and Deep Q-Network (DQN) algorithms showed significant differences in learning patterns and performance. Q-learning demonstrated rapid improvement during the early episodes of the training phase, converging to an average score of 7-8 after around 5,000 episodes. This quick convergence highlights the fact that Q-learning is rather efficient in this Taxi game environment.

However, it is important to note that a one-to-one direct comparison using controlled parameters for both algorithms is not a good idea at all. The learning rate α and discount factor γ impact Q-learning and DQN differently, probably due to their design difference. Especially, DQN has additional parameters such as network synchronization rate τ, batch size, and the neural network's weights and biases, which have no counterparts in the traditional Q-learning algorithm.

There is drastic difference between the performance of Q-learning and DQN with the controlled parameters $\alpha = 0.5, \gamma = 0.95$. DQN showed a gradual improvement up to 50,000 episodes but then plateaued at around -200 average reward which is significantly underperforming if compared to Q-learning. This stagnation in DQN training and performance suggests that the selected values might be suitable for Q-learning but only suboptimal for DQN.

Optimized Performance through Unique Parameter Tuning

Q-learning and DQN show more promising result with manually fine-tuned unique

parameters. With the tailored parameters, both Q-learning and DQN optimized to similar performance levels, converging to an average score of 7-8 over 1,000 test runs.

Q-learning shows a rapid learning curve just like before, stabilizing at the optimal performance level after approximately 5,000 training episodes.

DQN, in contrast, when properly fine-tuned, exhibited a different learning trajectory than before. DQN initially varies in its average training score but began to improve significantly after 10,000 episodes. The DQN algorithm gradually converged toward higher scores, eventually matching Q-learning's performance after about 75,000 episodes of training.

A few highlights:

- Algorithm specific parameters matter for performance (controlled parameters comparison simply does not work)
- Under the Taxi environment, DQN and Q-learning can be on par in performance when given sufficient and appropriate training time, episodes, and parameters
- Q-learning's rapid convergence shows a great fit for small environment like Taxi game and DQN properly has more potential in handling more complex environments

Conclusion

Convergence
Q-learning demonstrated consistent convergence to an optimal policy which align with its theoretical design. In contrast, DQN, by the nature of its design, does not necessarily guarantee convergence to an optimal policy. Despite that, our fine-tuned DQN model achieved performance on par with Q-learning in the Taxi environment. That is, DQN can provide a "good enough" policy for prediction.

Small-Scale Environments

For problem with limited action and observation spaces, such as the Taxi game, the performance difference between Q-learning and DQN was minimal. This result suggests that, perhaps in a Discrete environment the added complexity of DQN is probably not advantages. Where the Q-learning, which guarantees optimal policy, should be a better tool of choice.

Efficiency

Q-learning exhibited rapid convergence within thousands of episodes in the training phase. DQN, while capable of reaching similar performance as Q-learning eventually, required substantially more episodes and computational resources. The difference in efficiency is an important consideration for applications, especially in resource-

constrained situations.

Setting a Benchmark

The project highlighted the need for more robust benchmarking methodology to compare the performance of the reinforcement learning models. Data Structure and Algorithmic Analysis can always rely on metric like the Big O notation for performance comparison. In contrast, the reinforcement learning performance comparison in this project lacks some kind of standard metric as the performance can be influenced by various factors, including hardware specifications, system load, and even environmental conditions. Further study in reinforcement learning is needed in order to learn how to utilize standardized metrics to measure the models.

Potential Improvement

I think there are still room for me to improve and optimize the Q-learning and DQN implementations and comparison. Maybe like fine-tuning hyperparameters, optimizing network architecture for DQN, and exploring different network synchronization values and strategies. Also, testing the algorithms on multiple Gymnasium environment, for example, compare between Discrete Space and Continuous Space. Or to put the algorithms to test and compare between different games of Discrete space like Taxi game vs Cliff walking. Further, the algorithms can also be tested in a more complex environment, like a bigger and more complex map for the Taxi agent to perform in.

Final Thoughts

Overall, I found the project to be interesting, challenging yet rewarding, especially when I get to see the Taxi agent navigating its way through the map and picking-up/dropping-off the passenger correctly.

Reference: Please include all relevant links (git, video, etc)

https://github.com/yuchen071/DQN-for-LunarLander-v2/blob/main/LunarLander.ipynb

https://pytorch.org/tutorials/intermediate/reinforcement_q_learning.html

https://colab.research.google.com/drive/1peYBWia1IQzzgrbDIXwYYVUf7qiPHair?usp=sharing

https://stackoverflow.com/questions/53198503/epsilon-and-learning-rate-decay-in-epsilon-greedy-q-learning

https://medium.com/@CalebMBowyer/strategies-for-decaying-epsilon-in-epsilon-greedy-9b500ad9171d

https://youtu.be/EUrWGTCGzlA?si=fXl8PEnb0TrV9Vwt

https://www.stochasticbard.com/blog/deep_dive_reinforcement_learning/

https://stackoverflow.com/questions/40195740/how-to-run-openai-gym-render-over-a-server

https://colab.research.google.com/github/erodola/ML-s2-2024/blob/main/labs/07_qlearning.ipynb

https://wikidocs.net/174570

https://huggingface.co/learn/deep-rl-course/unit3/introduction

https://huggingface.co/learn/deep-rl-course/unit2/hands-on

https://medium.com/@hsinhungw/intro-to-reinforcement-learning-monte-carlo-to-policy-gradient-1c7ede4eed6e

https://drawar.github.io/blog/2019/05/12/lunar-lander-dqn.html

https://colab.research.google.com/drive/117s1JRxFTWHJXQiWOmp6Hhi4pbBFzhVE?usp=sharing

https://guandroz.medium.com/deep-q-learning-with-pytorch-and-openai-gym-the-taxi-cab-puzzle-e7a3028f732

https://tiewkh.github.io/blog/deepqlearning-openaitaxi/

https://rubikscode.net/2021/07/13/deep-q-learning-with-python-and-tensorflow-2-0/

https://www.youtube.com/watch?v=6kOvmZDEMdc

https://www.youtube.com/watch?v=EUrWGTCGzlA&list=PL58zEckBH8fCt_lYkmayZoR9XfDCW9hte

https://github.com/openai/gym/wiki/Leaderboard

https://aakash94.github.io/Reward-Based-Epsilon-Decay/

https://medium.com/@CalebMBowyer/strategies-for-decaying-epsilon-in-epsilon-greedy

https://towardsdatascience.com/convenient-reinforcement-learning-with-stable-baselines3-dccf466b7585