

# ShawnChumbar\_SARSA\_learner\_MountainCar

October 7, 2024

## 1 Assignment 01 - HW on Gym

CMPE 260 Homework 1 By Shawn Chumbar

### 1.1 Assignment Description

Use the Q-learning example given as a basis for your homework and develop a SARSA learning algorithm instead of Q-learning.

#### 1.1.1 Setup

```
[ ]: !pip install gymnasium[atari]
```

```
Collecting gymnasium[atari]
  Downloading gymnasium-0.29.1-py3-none-any.whl.metadata (10 kB)
Requirement already satisfied: numpy>=1.21.0 in /usr/local/lib/python3.10/dist-packages (from gymnasium[atari]) (1.26.4)
Requirement already satisfied: cloudpickle>=1.2.0 in /usr/local/lib/python3.10/dist-packages (from gymnasium[atari]) (2.2.1)
Requirement already satisfied: typing-extensions>=4.3.0 in /usr/local/lib/python3.10/dist-packages (from gymnasium[atari]) (4.12.2)
Collecting farama-notifications>=0.0.1 (from gymnasium[atari])
  Downloading Farama_Notifications-0.0.4-py3-none-any.whl.metadata (558 bytes)
Collecting shimmy<1.0,>=0.1.0 (from shimmy[atari]<1.0,>=0.1.0; extra == "atari"->gymnasium[atari])
  Downloading Shimmy-0.2.1-py3-none-any.whl.metadata (2.3 kB)
Collecting ale-py~0.8.1 (from shimmy[atari]<1.0,>=0.1.0; extra == "atari"->gymnasium[atari])
  Downloading ale_py-0.8.1-cp310-cp310-manylinux_2_17_x86_64.manylinux2014_x86_64.whl.metadata (8.1 kB)
Requirement already satisfied: importlib-resources in /usr/local/lib/python3.10/dist-packages (from ale-py~0.8.1->shimmy[atari]<1.0,>=0.1.0; extra == "atari"->gymnasium[atari]) (6.4.5)
Downloading Farama_Notifications-0.0.4-py3-none-any.whl (2.5 kB)
Downloading Shimmy-0.2.1-py3-none-any.whl (25 kB)
Downloading gymnasium-0.29.1-py3-none-any.whl (953 kB)
953.9/953.9 kB
```

34.3 MB/s eta 0:00:00

Downloading

ale\_py-0.8.1-cp310-cp310-manylinux\_2\_17\_x86\_64.manylinux2014\_x86\_64.whl (1.7 MB)  
1.7/1.7 MB

63.5 MB/s eta 0:00:00

Installing collected packages: farama-notifications, gymnasium, ale-py, shimmy

Successfully installed ale-py-0.8.1 farama-notifications-0.0.4 gymnasium-0.29.1 shimmy-0.2.1

```
[ ]: !pip install "ale-py>=0.8.0"
```

Requirement already satisfied: ale-py>=0.8.0 in /usr/local/lib/python3.10/dist-packages (0.8.1)

Requirement already satisfied: numpy in /usr/local/lib/python3.10/dist-packages (from ale-py>=0.8.0) (1.26.4)

Requirement already satisfied: importlib-resources in /usr/local/lib/python3.10/dist-packages (from ale-py>=0.8.0) (6.4.5)

Requirement already satisfied: typing-extensions in /usr/local/lib/python3.10/dist-packages (from ale-py>=0.8.0) (4.12.2)

```
[ ]: !pip install autorom[accept-rom-license]
```

Collecting autorom[accept-rom-license]

Downloading AutoROM-0.6.1-py3-none-any.whl.metadata (2.4 kB)

Requirement already satisfied: click in /usr/local/lib/python3.10/dist-packages (from autorom[accept-rom-license]) (8.1.7)

Requirement already satisfied: requests in /usr/local/lib/python3.10/dist-packages (from autorom[accept-rom-license]) (2.32.3)

Collecting AutoROM.accept-rom-license (from autorom[accept-rom-license])

Downloading AutoROM.accept-rom-license-0.6.1.tar.gz (434 kB)

434.7/434.7

kB 27.1 MB/s eta 0:00:00

Installing build dependencies ... done

Getting requirements to build wheel ... done

Preparing metadata (pyproject.toml) ... done

Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/python3.10/dist-packages (from requests->autorom[accept-rom-license]) (3.3.2)

Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/dist-packages (from requests->autorom[accept-rom-license]) (3.10)

Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3.10/dist-packages (from requests->autorom[accept-rom-license]) (2.2.3)

Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3.10/dist-packages (from requests->autorom[accept-rom-license]) (2024.8.30)

Downloading AutoROM-0.6.1-py3-none-any.whl (9.4 kB)

```

Building wheels for collected packages: AutoROM.accept-rom-license
  Building wheel for AutoROM.accept-rom-license (pyproject.toml) ...
done
  Created wheel for AutoROM.accept-rom-license:
  filename=AutoROM.accept_rom_license-0.6.1-py3-none-any.whl size=446661
  sha256=4587a0c38834593f96b438cad3ede1ef46edb6e76bc697f43f990d48c9ccb3d
  Stored in directory: /root/.cache/pip/wheels/6b/1b/ef/a43ff1a2f1736d5711faa1ba
  4c1f61be1131b8899e6a057811
Successfully built AutoROM.accept-rom-license
Installing collected packages: AutoROM.accept-rom-license, autorom
Successfully installed AutoROM.accept-rom-license-0.6.1 autorom-0.6.1

```

```
[ ]: import gymnasium as gym
      from gymnasium.wrappers import RecordVideo
```

```
[ ]: def run_gym_env(name, n):
      env = gym.make(name, render_mode="rgb_array")
      env = RecordVideo(env, "videos")
      observation, info = env.reset()
      for _ in range(int(n)):
          action = env.action_space.sample()
          observation, reward, terminated, truncated, info = env.step(action)
          if terminated or truncated:
              observation, info = env.reset()
      env.close()

      if __name__ == "__main__":
          run_gym_env('ALE/Alien-v5', 2000)
```

```

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283:
DeprecationWarning: `should_run_async` will not call `transform_cell`
automatically in the future. Please pass the result to `transformed_cell`
argument and any exception that happen during the transform in
`preprocessing_exc_tuple` in IPython 7.17 and above.
  and should_run_async(code)
/usr/local/lib/python3.10/dist-
packages/gymnasium/utils/passive_env_checker.py:335: UserWarning: WARN: No
render fps was declared in the environment (env.metadata['render_fps'] is None
or not defined), rendering may occur at inconsistent fps.
  logger.warn(

Moviepy - Building video /content/videos/rl-video-episode-0.mp4.
Moviepy - Writing video /content/videos/rl-video-episode-0.mp4

Moviepy - Done !
Moviepy - video ready /content/videos/rl-video-episode-0.mp4

```

Moviepy - Building video /content/videos/rl-video-episode-1.mp4.  
Moviepy - Writing video /content/videos/rl-video-episode-1.mp4

Moviepy - Done !  
Moviepy - video ready /content/videos/rl-video-episode-1.mp4

### 1.1.2 Q\_Learner

The code below is the standard Q Learner class that we were given in Class.

```
[ ]: import gymnasium as gym
import numpy as np

MAX_NUM_EPISODES = 50000
STEPS_PER_EPISODE = 200
EPSILON_MIN = 0.005
max_num_steps = MAX_NUM_EPISODES * STEPS_PER_EPISODE
EPSILON_DECAY = 500 * EPSILON_MIN / max_num_steps
ALPHA = 0.05
GAMMA = 0.98
NUM_DISCRETE_BINS = 30

class Q_Learner(object):
    def __init__(self, env):
        self.obs_shape = env.observation_space.shape
        self.obs_high = env.observation_space.high
        self.obs_low = env.observation_space.low
        self.obs_bins = NUM_DISCRETE_BINS
        self.bin_width = (self.obs_high - self.obs_low) / self.obs_bins
        self.action_shape = env.action_space.n
        self.Q = np.zeros((self.obs_bins + 1, self.obs_bins + 1, self.
↪ action_shape))
        self.alpha = ALPHA
        self.gamma = GAMMA
        self.epsilon = 1.0

    def discretize(self, obs):
        return tuple(((obs - self.obs_low) / self.bin_width).astype(int))

    def get_action(self, obs):
        discretized_obs = self.discretize(obs)
        if self.epsilon > EPSILON_MIN:
            self.epsilon -= EPSILON_DECAY
        if np.random.random() > self.epsilon:
            return np.argmax(self.Q[discretized_obs])
        else:
```

```

        return np.random.choice([a for a in range(self.action_shape)])

    def learn(self, obs, action, reward, next_obs):
        discretized_obs = self.discretize(obs)
        discretized_next_obs = self.discretize(next_obs)
        td_target = reward + self.gamma * np.max(self.Q[discretized_next_obs])
        td_error = td_target - self.Q[discretized_obs][action]
        self.Q[discretized_obs][action] += self.alpha * td_error

def train(agent, env):
    best_reward = -float('inf')
    for episode in range(MAX_NUM_EPISODES):
        obs, _ = env.reset()
        total_reward = 0.0
        for _ in range(STEPS_PER_EPISODE):
            action = agent.get_action(obs)
            next_obs, reward, terminated, truncated, _ = env.step(action)
            agent.learn(obs, action, reward, next_obs)
            obs = next_obs
            total_reward += reward
            if terminated or truncated:
                break
        if total_reward > best_reward:
            best_reward = total_reward
        if episode % 100 == 0:
            print(f"Episode#{episode} reward:{total_reward} best_reward:
↪{best_reward} eps:{agent.epsilon:.3f}")
    return np.argmax(agent.Q, axis=2)

def test(agent, env, policy):
    obs, _ = env.reset()
    total_reward = 0.0
    for _ in range(STEPS_PER_EPISODE):
        action = policy[agent.discretize(obs)]
        next_obs, reward, terminated, truncated, _ = env.step(action)
        obs = next_obs
        total_reward += reward
        if terminated or truncated:
            break
    return total_reward

if __name__ == "__main__":
    env = gym.make('MountainCar-v0')
    agent = Q_Learner(env)
    learned_policy = train(agent, env)

    test_env = gym.make('MountainCar-v0', render_mode='rgb_array')

```

```

test_env = gym.wrappers.RecordVideo(test_env, "videos")
for _ in range(5): # Test for 5 episodes
    total_reward = test(agent, test_env, learned_policy)
    print(f"Test episode reward: {total_reward}")
test_env.close()

```

/usr/local/lib/python3.10/dist-packages/ipykernel/ipkernel.py:283:  
DeprecationWarning: `should\_run\_async` will not call `transform\_cell`  
automatically in the future. Please pass the result to `transformed\_cell`  
argument and any exception that happen during the transform in  
`preprocessing\_exc\_tuple` in IPython 7.17 and above.  
and should\_run\_async(code)

```

Episode#0 reward:-200.0 best_reward:-200.0 eps:1.000
Episode#100 reward:-200.0 best_reward:-200.0 eps:0.995
Episode#200 reward:-200.0 best_reward:-200.0 eps:0.990
Episode#300 reward:-200.0 best_reward:-200.0 eps:0.985
Episode#400 reward:-200.0 best_reward:-200.0 eps:0.980
Episode#500 reward:-200.0 best_reward:-200.0 eps:0.975
Episode#600 reward:-200.0 best_reward:-200.0 eps:0.970
Episode#700 reward:-200.0 best_reward:-200.0 eps:0.965
Episode#800 reward:-200.0 best_reward:-200.0 eps:0.960
Episode#900 reward:-200.0 best_reward:-200.0 eps:0.955
Episode#1000 reward:-200.0 best_reward:-200.0 eps:0.950
Episode#1100 reward:-200.0 best_reward:-200.0 eps:0.945
Episode#1200 reward:-200.0 best_reward:-200.0 eps:0.940
Episode#1300 reward:-200.0 best_reward:-200.0 eps:0.935
Episode#1400 reward:-200.0 best_reward:-200.0 eps:0.930
Episode#1500 reward:-200.0 best_reward:-200.0 eps:0.925
Episode#1600 reward:-200.0 best_reward:-200.0 eps:0.920
Episode#1700 reward:-200.0 best_reward:-200.0 eps:0.915
Episode#1800 reward:-200.0 best_reward:-200.0 eps:0.910
Episode#1900 reward:-200.0 best_reward:-200.0 eps:0.905
Episode#2000 reward:-200.0 best_reward:-200.0 eps:0.900
Episode#2100 reward:-200.0 best_reward:-200.0 eps:0.895
Episode#2200 reward:-200.0 best_reward:-200.0 eps:0.890
Episode#2300 reward:-200.0 best_reward:-200.0 eps:0.885
Episode#2400 reward:-200.0 best_reward:-200.0 eps:0.880
Episode#2500 reward:-200.0 best_reward:-200.0 eps:0.875
Episode#2600 reward:-200.0 best_reward:-200.0 eps:0.870
Episode#2700 reward:-200.0 best_reward:-200.0 eps:0.865
Episode#2800 reward:-200.0 best_reward:-200.0 eps:0.860
Episode#2900 reward:-200.0 best_reward:-200.0 eps:0.855
Episode#3000 reward:-200.0 best_reward:-200.0 eps:0.850
Episode#3100 reward:-200.0 best_reward:-200.0 eps:0.845
Episode#3200 reward:-200.0 best_reward:-200.0 eps:0.840
Episode#3300 reward:-200.0 best_reward:-200.0 eps:0.835
Episode#3400 reward:-200.0 best_reward:-200.0 eps:0.830

```

Episode#3500 reward:-200.0 best\_reward:-200.0 eps:0.825  
Episode#3600 reward:-200.0 best\_reward:-200.0 eps:0.820  
Episode#3700 reward:-200.0 best\_reward:-200.0 eps:0.815  
Episode#3800 reward:-200.0 best\_reward:-200.0 eps:0.810  
Episode#3900 reward:-200.0 best\_reward:-200.0 eps:0.805  
Episode#4000 reward:-200.0 best\_reward:-200.0 eps:0.800  
Episode#4100 reward:-200.0 best\_reward:-200.0 eps:0.795  
Episode#4200 reward:-200.0 best\_reward:-200.0 eps:0.790  
Episode#4300 reward:-200.0 best\_reward:-200.0 eps:0.785  
Episode#4400 reward:-200.0 best\_reward:-200.0 eps:0.780  
Episode#4500 reward:-200.0 best\_reward:-200.0 eps:0.775  
Episode#4600 reward:-200.0 best\_reward:-200.0 eps:0.770  
Episode#4700 reward:-200.0 best\_reward:-200.0 eps:0.765  
Episode#4800 reward:-200.0 best\_reward:-200.0 eps:0.760  
Episode#4900 reward:-200.0 best\_reward:-200.0 eps:0.755  
Episode#5000 reward:-200.0 best\_reward:-200.0 eps:0.750  
Episode#5100 reward:-200.0 best\_reward:-200.0 eps:0.745  
Episode#5200 reward:-200.0 best\_reward:-200.0 eps:0.740  
Episode#5300 reward:-200.0 best\_reward:-200.0 eps:0.735  
Episode#5400 reward:-200.0 best\_reward:-200.0 eps:0.730  
Episode#5500 reward:-200.0 best\_reward:-200.0 eps:0.725  
Episode#5600 reward:-200.0 best\_reward:-200.0 eps:0.720  
Episode#5700 reward:-200.0 best\_reward:-200.0 eps:0.715  
Episode#5800 reward:-200.0 best\_reward:-200.0 eps:0.710  
Episode#5900 reward:-200.0 best\_reward:-200.0 eps:0.705  
Episode#6000 reward:-200.0 best\_reward:-200.0 eps:0.700  
Episode#6100 reward:-200.0 best\_reward:-200.0 eps:0.695  
Episode#6200 reward:-200.0 best\_reward:-200.0 eps:0.690  
Episode#6300 reward:-200.0 best\_reward:-200.0 eps:0.685  
Episode#6400 reward:-200.0 best\_reward:-200.0 eps:0.680  
Episode#6500 reward:-200.0 best\_reward:-200.0 eps:0.675  
Episode#6600 reward:-200.0 best\_reward:-200.0 eps:0.670  
Episode#6700 reward:-200.0 best\_reward:-200.0 eps:0.665  
Episode#6800 reward:-200.0 best\_reward:-200.0 eps:0.660  
Episode#6900 reward:-200.0 best\_reward:-200.0 eps:0.655  
Episode#7000 reward:-200.0 best\_reward:-200.0 eps:0.650  
Episode#7100 reward:-200.0 best\_reward:-200.0 eps:0.645  
Episode#7200 reward:-200.0 best\_reward:-200.0 eps:0.640  
Episode#7300 reward:-200.0 best\_reward:-200.0 eps:0.635  
Episode#7400 reward:-200.0 best\_reward:-200.0 eps:0.630  
Episode#7500 reward:-200.0 best\_reward:-200.0 eps:0.625  
Episode#7600 reward:-200.0 best\_reward:-200.0 eps:0.620  
Episode#7700 reward:-200.0 best\_reward:-200.0 eps:0.615  
Episode#7800 reward:-200.0 best\_reward:-200.0 eps:0.610  
Episode#7900 reward:-200.0 best\_reward:-200.0 eps:0.605  
Episode#8000 reward:-200.0 best\_reward:-200.0 eps:0.600  
Episode#8100 reward:-200.0 best\_reward:-200.0 eps:0.595  
Episode#8200 reward:-200.0 best\_reward:-200.0 eps:0.590

Episode#8300 reward:-200.0 best\_reward:-199.0 eps:0.585  
Episode#8400 reward:-200.0 best\_reward:-199.0 eps:0.580  
Episode#8500 reward:-200.0 best\_reward:-199.0 eps:0.575  
Episode#8600 reward:-200.0 best\_reward:-199.0 eps:0.570  
Episode#8700 reward:-200.0 best\_reward:-199.0 eps:0.565  
Episode#8800 reward:-200.0 best\_reward:-199.0 eps:0.560  
Episode#8900 reward:-200.0 best\_reward:-168.0 eps:0.555  
Episode#9000 reward:-200.0 best\_reward:-168.0 eps:0.550  
Episode#9100 reward:-200.0 best\_reward:-168.0 eps:0.545  
Episode#9200 reward:-200.0 best\_reward:-168.0 eps:0.540  
Episode#9300 reward:-200.0 best\_reward:-168.0 eps:0.535  
Episode#9400 reward:-200.0 best\_reward:-168.0 eps:0.530  
Episode#9500 reward:-200.0 best\_reward:-168.0 eps:0.525  
Episode#9600 reward:-200.0 best\_reward:-167.0 eps:0.520  
Episode#9700 reward:-200.0 best\_reward:-167.0 eps:0.515  
Episode#9800 reward:-200.0 best\_reward:-167.0 eps:0.510  
Episode#9900 reward:-200.0 best\_reward:-163.0 eps:0.505  
Episode#10000 reward:-200.0 best\_reward:-163.0 eps:0.500  
Episode#10100 reward:-200.0 best\_reward:-154.0 eps:0.495  
Episode#10200 reward:-200.0 best\_reward:-154.0 eps:0.490  
Episode#10300 reward:-200.0 best\_reward:-154.0 eps:0.485  
Episode#10400 reward:-200.0 best\_reward:-154.0 eps:0.480  
Episode#10500 reward:-200.0 best\_reward:-154.0 eps:0.475  
Episode#10600 reward:-200.0 best\_reward:-154.0 eps:0.470  
Episode#10700 reward:-200.0 best\_reward:-154.0 eps:0.465  
Episode#10800 reward:-200.0 best\_reward:-154.0 eps:0.460  
Episode#10900 reward:-200.0 best\_reward:-154.0 eps:0.455  
Episode#11000 reward:-200.0 best\_reward:-154.0 eps:0.450  
Episode#11100 reward:-200.0 best\_reward:-154.0 eps:0.445  
Episode#11200 reward:-200.0 best\_reward:-154.0 eps:0.440  
Episode#11300 reward:-200.0 best\_reward:-154.0 eps:0.435  
Episode#11400 reward:-200.0 best\_reward:-154.0 eps:0.431  
Episode#11500 reward:-200.0 best\_reward:-154.0 eps:0.426  
Episode#11600 reward:-169.0 best\_reward:-154.0 eps:0.421  
Episode#11700 reward:-200.0 best\_reward:-154.0 eps:0.416  
Episode#11800 reward:-200.0 best\_reward:-154.0 eps:0.411  
Episode#11900 reward:-200.0 best\_reward:-154.0 eps:0.406  
Episode#12000 reward:-168.0 best\_reward:-154.0 eps:0.401  
Episode#12100 reward:-200.0 best\_reward:-154.0 eps:0.396  
Episode#12200 reward:-200.0 best\_reward:-154.0 eps:0.391  
Episode#12300 reward:-200.0 best\_reward:-154.0 eps:0.386  
Episode#12400 reward:-152.0 best\_reward:-152.0 eps:0.381  
Episode#12500 reward:-200.0 best\_reward:-152.0 eps:0.377  
Episode#12600 reward:-200.0 best\_reward:-152.0 eps:0.372  
Episode#12700 reward:-200.0 best\_reward:-152.0 eps:0.367  
Episode#12800 reward:-200.0 best\_reward:-152.0 eps:0.362  
Episode#12900 reward:-200.0 best\_reward:-149.0 eps:0.357  
Episode#13000 reward:-200.0 best\_reward:-149.0 eps:0.352



Episode#13100 reward:-200.0 best\_reward:-149.0 eps:0.347  
Episode#13200 reward:-200.0 best\_reward:-149.0 eps:0.342  
Episode#13300 reward:-190.0 best\_reward:-149.0 eps:0.337  
Episode#13400 reward:-200.0 best\_reward:-149.0 eps:0.333  
Episode#13500 reward:-200.0 best\_reward:-149.0 eps:0.328  
Episode#13600 reward:-161.0 best\_reward:-149.0 eps:0.323  
Episode#13700 reward:-200.0 best\_reward:-149.0 eps:0.319  
Episode#13800 reward:-200.0 best\_reward:-149.0 eps:0.314  
Episode#13900 reward:-163.0 best\_reward:-149.0 eps:0.310  
Episode#14000 reward:-158.0 best\_reward:-149.0 eps:0.305  
Episode#14100 reward:-200.0 best\_reward:-149.0 eps:0.300  
Episode#14200 reward:-200.0 best\_reward:-149.0 eps:0.295  
Episode#14300 reward:-200.0 best\_reward:-149.0 eps:0.290  
Episode#14400 reward:-200.0 best\_reward:-149.0 eps:0.285  
Episode#14500 reward:-200.0 best\_reward:-139.0 eps:0.281  
Episode#14600 reward:-200.0 best\_reward:-139.0 eps:0.276  
Episode#14700 reward:-157.0 best\_reward:-139.0 eps:0.271  
Episode#14800 reward:-161.0 best\_reward:-139.0 eps:0.266  
Episode#14900 reward:-200.0 best\_reward:-139.0 eps:0.261  
Episode#15000 reward:-191.0 best\_reward:-139.0 eps:0.257  
Episode#15100 reward:-200.0 best\_reward:-139.0 eps:0.252  
Episode#15200 reward:-164.0 best\_reward:-139.0 eps:0.248  
Episode#15300 reward:-200.0 best\_reward:-139.0 eps:0.243  
Episode#15400 reward:-200.0 best\_reward:-139.0 eps:0.239  
Episode#15500 reward:-200.0 best\_reward:-139.0 eps:0.234  
Episode#15600 reward:-200.0 best\_reward:-139.0 eps:0.229  
Episode#15700 reward:-200.0 best\_reward:-139.0 eps:0.224  
Episode#15800 reward:-180.0 best\_reward:-128.0 eps:0.219  
Episode#15900 reward:-172.0 best\_reward:-128.0 eps:0.215  
Episode#16000 reward:-189.0 best\_reward:-119.0 eps:0.210  
Episode#16100 reward:-162.0 best\_reward:-119.0 eps:0.206  
Episode#16200 reward:-160.0 best\_reward:-119.0 eps:0.202  
Episode#16300 reward:-187.0 best\_reward:-119.0 eps:0.197  
Episode#16400 reward:-152.0 best\_reward:-119.0 eps:0.193  
Episode#16500 reward:-156.0 best\_reward:-119.0 eps:0.189  
Episode#16600 reward:-180.0 best\_reward:-119.0 eps:0.185  
Episode#16700 reward:-153.0 best\_reward:-119.0 eps:0.180  
Episode#16800 reward:-166.0 best\_reward:-114.0 eps:0.176  
Episode#16900 reward:-125.0 best\_reward:-114.0 eps:0.172  
Episode#17000 reward:-200.0 best\_reward:-114.0 eps:0.168  
Episode#17100 reward:-190.0 best\_reward:-114.0 eps:0.163  
Episode#17200 reward:-151.0 best\_reward:-114.0 eps:0.159  
Episode#17300 reward:-158.0 best\_reward:-114.0 eps:0.155  
Episode#17400 reward:-162.0 best\_reward:-114.0 eps:0.151  
Episode#17500 reward:-151.0 best\_reward:-114.0 eps:0.147  
Episode#17600 reward:-165.0 best\_reward:-114.0 eps:0.142  
Episode#17700 reward:-118.0 best\_reward:-114.0 eps:0.138  
Episode#17800 reward:-177.0 best\_reward:-114.0 eps:0.134

Episode#17900 reward:-200.0 best\_reward:-114.0 eps:0.129  
Episode#18000 reward:-200.0 best\_reward:-114.0 eps:0.124  
Episode#18100 reward:-197.0 best\_reward:-114.0 eps:0.119  
Episode#18200 reward:-160.0 best\_reward:-114.0 eps:0.115  
Episode#18300 reward:-155.0 best\_reward:-114.0 eps:0.111  
Episode#18400 reward:-153.0 best\_reward:-114.0 eps:0.107  
Episode#18500 reward:-160.0 best\_reward:-114.0 eps:0.102  
Episode#18600 reward:-192.0 best\_reward:-114.0 eps:0.098  
Episode#18700 reward:-158.0 best\_reward:-114.0 eps:0.093  
Episode#18800 reward:-159.0 best\_reward:-114.0 eps:0.089  
Episode#18900 reward:-163.0 best\_reward:-114.0 eps:0.085  
Episode#19000 reward:-200.0 best\_reward:-114.0 eps:0.080  
Episode#19100 reward:-200.0 best\_reward:-114.0 eps:0.076  
Episode#19200 reward:-200.0 best\_reward:-114.0 eps:0.071  
Episode#19300 reward:-200.0 best\_reward:-114.0 eps:0.066  
Episode#19400 reward:-180.0 best\_reward:-114.0 eps:0.062  
Episode#19500 reward:-158.0 best\_reward:-114.0 eps:0.057  
Episode#19600 reward:-151.0 best\_reward:-114.0 eps:0.053  
Episode#19700 reward:-200.0 best\_reward:-114.0 eps:0.049  
Episode#19800 reward:-158.0 best\_reward:-114.0 eps:0.045  
Episode#19900 reward:-200.0 best\_reward:-114.0 eps:0.041  
Episode#20000 reward:-157.0 best\_reward:-114.0 eps:0.036  
Episode#20100 reward:-161.0 best\_reward:-114.0 eps:0.032  
Episode#20200 reward:-158.0 best\_reward:-91.0 eps:0.028  
Episode#20300 reward:-188.0 best\_reward:-91.0 eps:0.023  
Episode#20400 reward:-200.0 best\_reward:-91.0 eps:0.019  
Episode#20500 reward:-158.0 best\_reward:-91.0 eps:0.015  
Episode#20600 reward:-172.0 best\_reward:-91.0 eps:0.011  
Episode#20700 reward:-147.0 best\_reward:-91.0 eps:0.006  
Episode#20800 reward:-190.0 best\_reward:-91.0 eps:0.005  
Episode#20900 reward:-149.0 best\_reward:-91.0 eps:0.005  
Episode#21000 reward:-156.0 best\_reward:-91.0 eps:0.005  
Episode#21100 reward:-181.0 best\_reward:-91.0 eps:0.005  
Episode#21200 reward:-151.0 best\_reward:-91.0 eps:0.005  
Episode#21300 reward:-199.0 best\_reward:-91.0 eps:0.005  
Episode#21400 reward:-191.0 best\_reward:-91.0 eps:0.005  
Episode#21500 reward:-154.0 best\_reward:-91.0 eps:0.005  
Episode#21600 reward:-157.0 best\_reward:-91.0 eps:0.005  
Episode#21700 reward:-195.0 best\_reward:-91.0 eps:0.005  
Episode#21800 reward:-153.0 best\_reward:-91.0 eps:0.005  
Episode#21900 reward:-159.0 best\_reward:-91.0 eps:0.005  
Episode#22000 reward:-148.0 best\_reward:-91.0 eps:0.005  
Episode#22100 reward:-169.0 best\_reward:-91.0 eps:0.005  
Episode#22200 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#22300 reward:-149.0 best\_reward:-91.0 eps:0.005  
Episode#22400 reward:-148.0 best\_reward:-91.0 eps:0.005  
Episode#22500 reward:-115.0 best\_reward:-91.0 eps:0.005  
Episode#22600 reward:-196.0 best\_reward:-91.0 eps:0.005

Episode#22700 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#22800 reward:-186.0 best\_reward:-91.0 eps:0.005  
Episode#22900 reward:-163.0 best\_reward:-91.0 eps:0.005  
Episode#23000 reward:-151.0 best\_reward:-91.0 eps:0.005  
Episode#23100 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#23200 reward:-125.0 best\_reward:-91.0 eps:0.005  
Episode#23300 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#23400 reward:-152.0 best\_reward:-91.0 eps:0.005  
Episode#23500 reward:-189.0 best\_reward:-91.0 eps:0.005  
Episode#23600 reward:-158.0 best\_reward:-91.0 eps:0.005  
Episode#23700 reward:-150.0 best\_reward:-91.0 eps:0.005  
Episode#23800 reward:-152.0 best\_reward:-91.0 eps:0.005  
Episode#23900 reward:-177.0 best\_reward:-91.0 eps:0.005  
Episode#24000 reward:-154.0 best\_reward:-91.0 eps:0.005  
Episode#24100 reward:-154.0 best\_reward:-91.0 eps:0.005  
Episode#24200 reward:-162.0 best\_reward:-91.0 eps:0.005  
Episode#24300 reward:-147.0 best\_reward:-91.0 eps:0.005  
Episode#24400 reward:-147.0 best\_reward:-91.0 eps:0.005  
Episode#24500 reward:-150.0 best\_reward:-91.0 eps:0.005  
Episode#24600 reward:-191.0 best\_reward:-91.0 eps:0.005  
Episode#24700 reward:-159.0 best\_reward:-91.0 eps:0.005  
Episode#24800 reward:-142.0 best\_reward:-91.0 eps:0.005  
Episode#24900 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#25000 reward:-149.0 best\_reward:-91.0 eps:0.005  
Episode#25100 reward:-158.0 best\_reward:-91.0 eps:0.005  
Episode#25200 reward:-160.0 best\_reward:-91.0 eps:0.005  
Episode#25300 reward:-153.0 best\_reward:-91.0 eps:0.005  
Episode#25400 reward:-149.0 best\_reward:-91.0 eps:0.005  
Episode#25500 reward:-152.0 best\_reward:-91.0 eps:0.005  
Episode#25600 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#25700 reward:-148.0 best\_reward:-91.0 eps:0.005  
Episode#25800 reward:-146.0 best\_reward:-91.0 eps:0.005  
Episode#25900 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#26000 reward:-157.0 best\_reward:-91.0 eps:0.005  
Episode#26100 reward:-148.0 best\_reward:-91.0 eps:0.005  
Episode#26200 reward:-200.0 best\_reward:-91.0 eps:0.005  
Episode#26300 reward:-136.0 best\_reward:-91.0 eps:0.005  
Episode#26400 reward:-152.0 best\_reward:-91.0 eps:0.005  
Episode#26500 reward:-145.0 best\_reward:-91.0 eps:0.005  
Episode#26600 reward:-155.0 best\_reward:-91.0 eps:0.005  
Episode#26700 reward:-158.0 best\_reward:-91.0 eps:0.005  
Episode#26800 reward:-131.0 best\_reward:-91.0 eps:0.005  
Episode#26900 reward:-152.0 best\_reward:-91.0 eps:0.005  
Episode#27000 reward:-113.0 best\_reward:-91.0 eps:0.005  
Episode#27100 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#27200 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#27300 reward:-139.0 best\_reward:-88.0 eps:0.005  
Episode#27400 reward:-144.0 best\_reward:-88.0 eps:0.005

Episode#27500 reward:-159.0 best\_reward:-88.0 eps:0.005  
Episode#27600 reward:-176.0 best\_reward:-88.0 eps:0.005  
Episode#27700 reward:-125.0 best\_reward:-88.0 eps:0.005  
Episode#27800 reward:-156.0 best\_reward:-88.0 eps:0.005  
Episode#27900 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#28000 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#28100 reward:-194.0 best\_reward:-88.0 eps:0.005  
Episode#28200 reward:-155.0 best\_reward:-88.0 eps:0.005  
Episode#28300 reward:-184.0 best\_reward:-88.0 eps:0.005  
Episode#28400 reward:-154.0 best\_reward:-88.0 eps:0.005  
Episode#28500 reward:-147.0 best\_reward:-88.0 eps:0.005  
Episode#28600 reward:-151.0 best\_reward:-88.0 eps:0.005  
Episode#28700 reward:-149.0 best\_reward:-88.0 eps:0.005  
Episode#28800 reward:-149.0 best\_reward:-88.0 eps:0.005  
Episode#28900 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#29000 reward:-147.0 best\_reward:-88.0 eps:0.005  
Episode#29100 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#29200 reward:-179.0 best\_reward:-88.0 eps:0.005  
Episode#29300 reward:-152.0 best\_reward:-88.0 eps:0.005  
Episode#29400 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#29500 reward:-114.0 best\_reward:-88.0 eps:0.005  
Episode#29600 reward:-115.0 best\_reward:-88.0 eps:0.005  
Episode#29700 reward:-146.0 best\_reward:-88.0 eps:0.005  
Episode#29800 reward:-154.0 best\_reward:-88.0 eps:0.005  
Episode#29900 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#30000 reward:-151.0 best\_reward:-88.0 eps:0.005  
Episode#30100 reward:-145.0 best\_reward:-88.0 eps:0.005  
Episode#30200 reward:-151.0 best\_reward:-88.0 eps:0.005  
Episode#30300 reward:-148.0 best\_reward:-88.0 eps:0.005  
Episode#30400 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#30500 reward:-141.0 best\_reward:-88.0 eps:0.005  
Episode#30600 reward:-149.0 best\_reward:-88.0 eps:0.005  
Episode#30700 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#30800 reward:-148.0 best\_reward:-88.0 eps:0.005  
Episode#30900 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#31000 reward:-142.0 best\_reward:-88.0 eps:0.005  
Episode#31100 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#31200 reward:-196.0 best\_reward:-88.0 eps:0.005  
Episode#31300 reward:-192.0 best\_reward:-88.0 eps:0.005  
Episode#31400 reward:-156.0 best\_reward:-88.0 eps:0.005  
Episode#31500 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#31600 reward:-144.0 best\_reward:-88.0 eps:0.005  
Episode#31700 reward:-157.0 best\_reward:-88.0 eps:0.005  
Episode#31800 reward:-151.0 best\_reward:-88.0 eps:0.005  
Episode#31900 reward:-156.0 best\_reward:-88.0 eps:0.005  
Episode#32000 reward:-113.0 best\_reward:-88.0 eps:0.005  
Episode#32100 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#32200 reward:-112.0 best\_reward:-88.0 eps:0.005

Episode#32300 reward:-146.0 best\_reward:-88.0 eps:0.005  
Episode#32400 reward:-140.0 best\_reward:-88.0 eps:0.005  
Episode#32500 reward:-144.0 best\_reward:-88.0 eps:0.005  
Episode#32600 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#32700 reward:-139.0 best\_reward:-88.0 eps:0.005  
Episode#32800 reward:-142.0 best\_reward:-88.0 eps:0.005  
Episode#32900 reward:-136.0 best\_reward:-88.0 eps:0.005  
Episode#33000 reward:-148.0 best\_reward:-88.0 eps:0.005  
Episode#33100 reward:-112.0 best\_reward:-88.0 eps:0.005  
Episode#33200 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#33300 reward:-113.0 best\_reward:-88.0 eps:0.005  
Episode#33400 reward:-142.0 best\_reward:-88.0 eps:0.005  
Episode#33500 reward:-111.0 best\_reward:-88.0 eps:0.005  
Episode#33600 reward:-140.0 best\_reward:-88.0 eps:0.005  
Episode#33700 reward:-158.0 best\_reward:-88.0 eps:0.005  
Episode#33800 reward:-113.0 best\_reward:-88.0 eps:0.005  
Episode#33900 reward:-141.0 best\_reward:-88.0 eps:0.005  
Episode#34000 reward:-141.0 best\_reward:-88.0 eps:0.005  
Episode#34100 reward:-117.0 best\_reward:-88.0 eps:0.005  
Episode#34200 reward:-138.0 best\_reward:-88.0 eps:0.005  
Episode#34300 reward:-146.0 best\_reward:-88.0 eps:0.005  
Episode#34400 reward:-139.0 best\_reward:-88.0 eps:0.005  
Episode#34500 reward:-122.0 best\_reward:-88.0 eps:0.005  
Episode#34600 reward:-121.0 best\_reward:-88.0 eps:0.005  
Episode#34700 reward:-141.0 best\_reward:-88.0 eps:0.005  
Episode#34800 reward:-117.0 best\_reward:-88.0 eps:0.005  
Episode#34900 reward:-116.0 best\_reward:-88.0 eps:0.005  
Episode#35000 reward:-116.0 best\_reward:-88.0 eps:0.005  
Episode#35100 reward:-114.0 best\_reward:-88.0 eps:0.005  
Episode#35200 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#35300 reward:-135.0 best\_reward:-88.0 eps:0.005  
Episode#35400 reward:-140.0 best\_reward:-88.0 eps:0.005  
Episode#35500 reward:-127.0 best\_reward:-88.0 eps:0.005  
Episode#35600 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#35700 reward:-198.0 best\_reward:-88.0 eps:0.005  
Episode#35800 reward:-157.0 best\_reward:-88.0 eps:0.005  
Episode#35900 reward:-115.0 best\_reward:-88.0 eps:0.005  
Episode#36000 reward:-125.0 best\_reward:-88.0 eps:0.005  
Episode#36100 reward:-188.0 best\_reward:-88.0 eps:0.005  
Episode#36200 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#36300 reward:-121.0 best\_reward:-88.0 eps:0.005  
Episode#36400 reward:-114.0 best\_reward:-88.0 eps:0.005  
Episode#36500 reward:-121.0 best\_reward:-88.0 eps:0.005  
Episode#36600 reward:-128.0 best\_reward:-88.0 eps:0.005  
Episode#36700 reward:-125.0 best\_reward:-88.0 eps:0.005  
Episode#36800 reward:-191.0 best\_reward:-88.0 eps:0.005  
Episode#36900 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#37000 reward:-189.0 best\_reward:-88.0 eps:0.005

Episode#37100 reward:-188.0 best\_reward:-88.0 eps:0.005  
Episode#37200 reward:-194.0 best\_reward:-88.0 eps:0.005  
Episode#37300 reward:-166.0 best\_reward:-88.0 eps:0.005  
Episode#37400 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#37500 reward:-147.0 best\_reward:-88.0 eps:0.005  
Episode#37600 reward:-189.0 best\_reward:-88.0 eps:0.005  
Episode#37700 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#37800 reward:-184.0 best\_reward:-88.0 eps:0.005  
Episode#37900 reward:-176.0 best\_reward:-88.0 eps:0.005  
Episode#38000 reward:-165.0 best\_reward:-88.0 eps:0.005  
Episode#38100 reward:-194.0 best\_reward:-88.0 eps:0.005  
Episode#38200 reward:-195.0 best\_reward:-88.0 eps:0.005  
Episode#38300 reward:-148.0 best\_reward:-88.0 eps:0.005  
Episode#38400 reward:-142.0 best\_reward:-88.0 eps:0.005  
Episode#38500 reward:-121.0 best\_reward:-88.0 eps:0.005  
Episode#38600 reward:-161.0 best\_reward:-88.0 eps:0.005  
Episode#38700 reward:-170.0 best\_reward:-88.0 eps:0.005  
Episode#38800 reward:-126.0 best\_reward:-88.0 eps:0.005  
Episode#38900 reward:-131.0 best\_reward:-88.0 eps:0.005  
Episode#39000 reward:-160.0 best\_reward:-88.0 eps:0.005  
Episode#39100 reward:-154.0 best\_reward:-88.0 eps:0.005  
Episode#39200 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#39300 reward:-129.0 best\_reward:-88.0 eps:0.005  
Episode#39400 reward:-146.0 best\_reward:-88.0 eps:0.005  
Episode#39500 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#39600 reward:-121.0 best\_reward:-88.0 eps:0.005  
Episode#39700 reward:-140.0 best\_reward:-88.0 eps:0.005  
Episode#39800 reward:-147.0 best\_reward:-88.0 eps:0.005  
Episode#39900 reward:-146.0 best\_reward:-88.0 eps:0.005  
Episode#40000 reward:-145.0 best\_reward:-88.0 eps:0.005  
Episode#40100 reward:-139.0 best\_reward:-88.0 eps:0.005  
Episode#40200 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#40300 reward:-116.0 best\_reward:-88.0 eps:0.005  
Episode#40400 reward:-145.0 best\_reward:-88.0 eps:0.005  
Episode#40500 reward:-148.0 best\_reward:-88.0 eps:0.005  
Episode#40600 reward:-116.0 best\_reward:-88.0 eps:0.005  
Episode#40700 reward:-145.0 best\_reward:-88.0 eps:0.005  
Episode#40800 reward:-146.0 best\_reward:-88.0 eps:0.005  
Episode#40900 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#41000 reward:-137.0 best\_reward:-88.0 eps:0.005  
Episode#41100 reward:-113.0 best\_reward:-88.0 eps:0.005  
Episode#41200 reward:-158.0 best\_reward:-88.0 eps:0.005  
Episode#41300 reward:-145.0 best\_reward:-88.0 eps:0.005  
Episode#41400 reward:-152.0 best\_reward:-88.0 eps:0.005  
Episode#41500 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#41600 reward:-112.0 best\_reward:-88.0 eps:0.005  
Episode#41700 reward:-141.0 best\_reward:-88.0 eps:0.005  
Episode#41800 reward:-153.0 best\_reward:-88.0 eps:0.005

Episode#41900 reward:-145.0 best\_reward:-88.0 eps:0.005  
Episode#42000 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#42100 reward:-134.0 best\_reward:-88.0 eps:0.005  
Episode#42200 reward:-115.0 best\_reward:-88.0 eps:0.005  
Episode#42300 reward:-153.0 best\_reward:-88.0 eps:0.005  
Episode#42400 reward:-149.0 best\_reward:-88.0 eps:0.005  
Episode#42500 reward:-114.0 best\_reward:-88.0 eps:0.005  
Episode#42600 reward:-113.0 best\_reward:-88.0 eps:0.005  
Episode#42700 reward:-158.0 best\_reward:-88.0 eps:0.005  
Episode#42800 reward:-155.0 best\_reward:-88.0 eps:0.005  
Episode#42900 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#43000 reward:-186.0 best\_reward:-88.0 eps:0.005  
Episode#43100 reward:-155.0 best\_reward:-88.0 eps:0.005  
Episode#43200 reward:-120.0 best\_reward:-88.0 eps:0.005  
Episode#43300 reward:-157.0 best\_reward:-88.0 eps:0.005  
Episode#43400 reward:-118.0 best\_reward:-88.0 eps:0.005  
Episode#43500 reward:-140.0 best\_reward:-88.0 eps:0.005  
Episode#43600 reward:-123.0 best\_reward:-88.0 eps:0.005  
Episode#43700 reward:-142.0 best\_reward:-88.0 eps:0.005  
Episode#43800 reward:-157.0 best\_reward:-88.0 eps:0.005  
Episode#43900 reward:-126.0 best\_reward:-88.0 eps:0.005  
Episode#44000 reward:-150.0 best\_reward:-88.0 eps:0.005  
Episode#44100 reward:-155.0 best\_reward:-88.0 eps:0.005  
Episode#44200 reward:-149.0 best\_reward:-88.0 eps:0.005  
Episode#44300 reward:-120.0 best\_reward:-88.0 eps:0.005  
Episode#44400 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#44500 reward:-190.0 best\_reward:-88.0 eps:0.005  
Episode#44600 reward:-143.0 best\_reward:-88.0 eps:0.005  
Episode#44700 reward:-157.0 best\_reward:-88.0 eps:0.005  
Episode#44800 reward:-126.0 best\_reward:-88.0 eps:0.005  
Episode#44900 reward:-146.0 best\_reward:-88.0 eps:0.005  
Episode#45000 reward:-175.0 best\_reward:-88.0 eps:0.005  
Episode#45100 reward:-200.0 best\_reward:-88.0 eps:0.005  
Episode#45200 reward:-120.0 best\_reward:-88.0 eps:0.005  
Episode#45300 reward:-122.0 best\_reward:-88.0 eps:0.005  
Episode#45400 reward:-95.0 best\_reward:-88.0 eps:0.005  
Episode#45500 reward:-94.0 best\_reward:-88.0 eps:0.005  
Episode#45600 reward:-186.0 best\_reward:-88.0 eps:0.005  
Episode#45700 reward:-120.0 best\_reward:-88.0 eps:0.005  
Episode#45800 reward:-119.0 best\_reward:-88.0 eps:0.005  
Episode#45900 reward:-112.0 best\_reward:-88.0 eps:0.005  
Episode#46000 reward:-111.0 best\_reward:-88.0 eps:0.005  
Episode#46100 reward:-109.0 best\_reward:-88.0 eps:0.005  
Episode#46200 reward:-114.0 best\_reward:-88.0 eps:0.005  
Episode#46300 reward:-115.0 best\_reward:-88.0 eps:0.005  
Episode#46400 reward:-110.0 best\_reward:-88.0 eps:0.005  
Episode#46500 reward:-114.0 best\_reward:-88.0 eps:0.005  
Episode#46600 reward:-130.0 best\_reward:-88.0 eps:0.005

```
Episode#46700 reward:-154.0 best_reward:-88.0 eps:0.005
Episode#46800 reward:-150.0 best_reward:-88.0 eps:0.005
Episode#46900 reward:-183.0 best_reward:-88.0 eps:0.005
Episode#47000 reward:-158.0 best_reward:-88.0 eps:0.005
Episode#47100 reward:-189.0 best_reward:-88.0 eps:0.005
Episode#47200 reward:-142.0 best_reward:-88.0 eps:0.005
Episode#47300 reward:-169.0 best_reward:-88.0 eps:0.005
Episode#47400 reward:-194.0 best_reward:-88.0 eps:0.005
Episode#47500 reward:-152.0 best_reward:-88.0 eps:0.005
Episode#47600 reward:-153.0 best_reward:-88.0 eps:0.005
Episode#47700 reward:-155.0 best_reward:-88.0 eps:0.005
Episode#47800 reward:-179.0 best_reward:-88.0 eps:0.005
Episode#47900 reward:-183.0 best_reward:-88.0 eps:0.005
Episode#48000 reward:-167.0 best_reward:-88.0 eps:0.005
Episode#48100 reward:-155.0 best_reward:-88.0 eps:0.005
Episode#48200 reward:-200.0 best_reward:-88.0 eps:0.005
Episode#48300 reward:-152.0 best_reward:-88.0 eps:0.005
Episode#48400 reward:-143.0 best_reward:-88.0 eps:0.005
Episode#48500 reward:-184.0 best_reward:-88.0 eps:0.005
Episode#48600 reward:-144.0 best_reward:-88.0 eps:0.005
Episode#48700 reward:-190.0 best_reward:-88.0 eps:0.005
Episode#48800 reward:-167.0 best_reward:-88.0 eps:0.005
Episode#48900 reward:-140.0 best_reward:-88.0 eps:0.005
Episode#49000 reward:-152.0 best_reward:-88.0 eps:0.005
Episode#49100 reward:-188.0 best_reward:-88.0 eps:0.005
Episode#49200 reward:-143.0 best_reward:-88.0 eps:0.005
Episode#49300 reward:-144.0 best_reward:-88.0 eps:0.005
Episode#49400 reward:-133.0 best_reward:-88.0 eps:0.005
Episode#49500 reward:-142.0 best_reward:-88.0 eps:0.005
Episode#49600 reward:-152.0 best_reward:-88.0 eps:0.005
Episode#49700 reward:-118.0 best_reward:-88.0 eps:0.005
Episode#49800 reward:-141.0 best_reward:-88.0 eps:0.005
Episode#49900 reward:-142.0 best_reward:-88.0 eps:0.005
```

```
/usr/local/lib/python3.10/dist-packages/gymnasium/wrappers/record_video.py:94:
```

```
UserWarning: WARN: Overwriting existing videos at /content/videos folder
```

```
(try specifying a different `video_folder` for the `RecordVideo` wrapper if this
is not desired)
```

```
logger.warn(
```

```
Moviepy - Building video /content/videos/rl-video-episode-0.mp4.
```

```
Moviepy - Writing video /content/videos/rl-video-episode-0.mp4
```

```
Moviepy - Done !
```

```
Moviepy - video ready /content/videos/rl-video-episode-0.mp4
```



```
Test episode reward: -141.0
Moviepy - Building video /content/videos/rl-video-episode-1.mp4.
Moviepy - Writing video /content/videos/rl-video-episode-1.mp4
```

```
Moviepy - Done !
Moviepy - video ready /content/videos/rl-video-episode-1.mp4
Test episode reward: -140.0
Test episode reward: -142.0
Test episode reward: -150.0
Test episode reward: -124.0
```

### 1.1.3 SARSA Learner

```
[ ]: import gymnasium as gym
import numpy as np

MAX_NUM_EPISODES = 50000
STEPS_PER_EPISODE = 200
EPSILON_MIN = 0.005
max_num_steps = MAX_NUM_EPISODES * STEPS_PER_EPISODE
EPSILON_DECAY = 500 * EPSILON_MIN / max_num_steps
ALPHA = 0.05
GAMMA = 0.98
NUM_DISCRETE_BINS = 30

class SARSA_Learner(object):
    def __init__(self, env):
        self.obs_shape = env.observation_space.shape
        self.obs_high = env.observation_space.high
        self.obs_low = env.observation_space.low
        self.obs_bins = NUM_DISCRETE_BINS
        self.bin_width = (self.obs_high - self.obs_low) / self.obs_bins
        self.action_shape = env.action_space.n
        self.Q = np.zeros((self.obs_bins + 1, self.obs_bins + 1, self.
↪ action_shape))
        self.alpha = ALPHA
        self.gamma = GAMMA
        self.epsilon = 1.0

    def discretize(self, obs):
        return tuple(((obs - self.obs_low) / self.bin_width).astype(int))

    def get_action(self, obs):
        discretized_obs = self.discretize(obs)
        if self.epsilon > EPSILON_MIN:
```

```

        self.epsilon -= EPSILON_DECAY
    if np.random.random() > self.epsilon:
        return np.argmax(self.Q[discretized_obs])
    else:
        return np.random.choice([a for a in range(self.action_shape)])

# Modified learn method for SARSA
def learn(self, obs, action, reward, next_obs, next_action):
    discretized_obs = self.discretize(obs)
    discretized_next_obs = self.discretize(next_obs)
    # SARSA Update to use Q-value of next state-action pair instead of max_
    ↪ Q-value
    td_target = reward + self.gamma * self.
    ↪ Q[discretized_next_obs][next_action]
    td_error = td_target - self.Q[discretized_obs][action]
    self.Q[discretized_obs][action] += self.alpha * td_error

def train(agent, env):
    best_reward = -float('inf')
    for episode in range(MAX_NUM_EPISODES):
        obs, _ = env.reset()
        total_reward = 0.0
        # Choose initial action
        action = agent.get_action(obs) # Added this line for SARSA
        for _ in range(STEPS_PER_EPISODE):
            next_obs, reward, terminated, truncated, _ = env.step(action)
            next_action = agent.get_action(next_obs) # Choose next action
            # SARSA learning: pass current state-action and next state-action
            agent.learn(obs, action, reward, next_obs, next_action)
            obs = next_obs
            action = next_action # Update action for next iteration
            total_reward += reward
            if terminated or truncated:
                break
            if total_reward > best_reward:
                best_reward = total_reward
            if episode % 100 == 0:
                print(f"Episode#{episode} reward:{total_reward} best_reward:
    ↪ {best_reward} eps:{agent.epsilon:.3f}")
        return np.argmax(agent.Q, axis=2)

def test(agent, env, policy):
    obs, _ = env.reset()
    total_reward = 0.0
    for _ in range(STEPS_PER_EPISODE):
        action = policy[agent.discretize(obs)]
        next_obs, reward, terminated, truncated, _ = env.step(action)

```

```

        obs = next_obs
        total_reward += reward
        if terminated or truncated:
            break
    return total_reward

if __name__ == "__main__":
    env = gym.make('MountainCar-v0')
    agent = SARSA_Learner(env) # Changed to SARSA_Learner
    learned_policy = train(agent, env)

    test_env = gym.make('MountainCar-v0', render_mode='rgb_array')
    test_env = gym.wrappers.RecordVideo(test_env, "videos")
    for _ in range(20): # Test for 5 episodes
        total_reward = test(agent, test_env, learned_policy)
        print(f"Test episode reward: {total_reward}")
    test_env.close()

```

```

Episode#0 reward:-200.0 best_reward:-200.0 eps:1.000
Episode#100 reward:-200.0 best_reward:-200.0 eps:0.995
Episode#200 reward:-200.0 best_reward:-200.0 eps:0.990
Episode#300 reward:-200.0 best_reward:-200.0 eps:0.985
Episode#400 reward:-200.0 best_reward:-200.0 eps:0.980
Episode#500 reward:-200.0 best_reward:-200.0 eps:0.975
Episode#600 reward:-200.0 best_reward:-200.0 eps:0.970
Episode#700 reward:-200.0 best_reward:-200.0 eps:0.965
Episode#800 reward:-200.0 best_reward:-200.0 eps:0.960
Episode#900 reward:-200.0 best_reward:-200.0 eps:0.955
Episode#1000 reward:-200.0 best_reward:-200.0 eps:0.950
Episode#1100 reward:-200.0 best_reward:-200.0 eps:0.945
Episode#1200 reward:-200.0 best_reward:-200.0 eps:0.940
Episode#1300 reward:-200.0 best_reward:-200.0 eps:0.935
Episode#1400 reward:-200.0 best_reward:-200.0 eps:0.930
Episode#1500 reward:-200.0 best_reward:-200.0 eps:0.925
Episode#1600 reward:-200.0 best_reward:-200.0 eps:0.920
Episode#1700 reward:-200.0 best_reward:-200.0 eps:0.915
Episode#1800 reward:-200.0 best_reward:-200.0 eps:0.909
Episode#1900 reward:-200.0 best_reward:-200.0 eps:0.904
Episode#2000 reward:-200.0 best_reward:-200.0 eps:0.899
Episode#2100 reward:-200.0 best_reward:-200.0 eps:0.894
Episode#2200 reward:-200.0 best_reward:-200.0 eps:0.889
Episode#2300 reward:-200.0 best_reward:-200.0 eps:0.884
Episode#2400 reward:-200.0 best_reward:-200.0 eps:0.879
Episode#2500 reward:-200.0 best_reward:-200.0 eps:0.874
Episode#2600 reward:-200.0 best_reward:-200.0 eps:0.869
Episode#2700 reward:-200.0 best_reward:-200.0 eps:0.864
Episode#2800 reward:-200.0 best_reward:-200.0 eps:0.859
Episode#2900 reward:-200.0 best_reward:-200.0 eps:0.854

```

Episode#3000 reward:-200.0 best\_reward:-200.0 eps:0.849  
Episode#3100 reward:-200.0 best\_reward:-200.0 eps:0.844  
Episode#3200 reward:-200.0 best\_reward:-200.0 eps:0.839  
Episode#3300 reward:-200.0 best\_reward:-200.0 eps:0.834  
Episode#3400 reward:-200.0 best\_reward:-200.0 eps:0.829  
Episode#3500 reward:-200.0 best\_reward:-200.0 eps:0.824  
Episode#3600 reward:-200.0 best\_reward:-200.0 eps:0.819  
Episode#3700 reward:-200.0 best\_reward:-200.0 eps:0.814  
Episode#3800 reward:-200.0 best\_reward:-200.0 eps:0.809  
Episode#3900 reward:-200.0 best\_reward:-200.0 eps:0.804  
Episode#4000 reward:-200.0 best\_reward:-200.0 eps:0.799  
Episode#4100 reward:-200.0 best\_reward:-200.0 eps:0.794  
Episode#4200 reward:-200.0 best\_reward:-200.0 eps:0.789  
Episode#4300 reward:-200.0 best\_reward:-200.0 eps:0.784  
Episode#4400 reward:-200.0 best\_reward:-200.0 eps:0.779  
Episode#4500 reward:-200.0 best\_reward:-200.0 eps:0.774  
Episode#4600 reward:-200.0 best\_reward:-200.0 eps:0.769  
Episode#4700 reward:-200.0 best\_reward:-200.0 eps:0.764  
Episode#4800 reward:-200.0 best\_reward:-200.0 eps:0.759  
Episode#4900 reward:-200.0 best\_reward:-200.0 eps:0.754  
Episode#5000 reward:-200.0 best\_reward:-200.0 eps:0.749  
Episode#5100 reward:-200.0 best\_reward:-200.0 eps:0.744  
Episode#5200 reward:-200.0 best\_reward:-200.0 eps:0.739  
Episode#5300 reward:-200.0 best\_reward:-200.0 eps:0.734  
Episode#5400 reward:-200.0 best\_reward:-200.0 eps:0.729  
Episode#5500 reward:-200.0 best\_reward:-200.0 eps:0.724  
Episode#5600 reward:-200.0 best\_reward:-200.0 eps:0.719  
Episode#5700 reward:-200.0 best\_reward:-200.0 eps:0.714  
Episode#5800 reward:-200.0 best\_reward:-200.0 eps:0.708  
Episode#5900 reward:-200.0 best\_reward:-200.0 eps:0.703  
Episode#6000 reward:-200.0 best\_reward:-200.0 eps:0.698  
Episode#6100 reward:-200.0 best\_reward:-200.0 eps:0.693  
Episode#6200 reward:-200.0 best\_reward:-200.0 eps:0.688  
Episode#6300 reward:-200.0 best\_reward:-200.0 eps:0.683  
Episode#6400 reward:-200.0 best\_reward:-200.0 eps:0.678  
Episode#6500 reward:-200.0 best\_reward:-200.0 eps:0.673  
Episode#6600 reward:-200.0 best\_reward:-200.0 eps:0.668  
Episode#6700 reward:-200.0 best\_reward:-200.0 eps:0.663  
Episode#6800 reward:-200.0 best\_reward:-200.0 eps:0.658  
Episode#6900 reward:-200.0 best\_reward:-200.0 eps:0.653  
Episode#7000 reward:-200.0 best\_reward:-200.0 eps:0.648  
Episode#7100 reward:-200.0 best\_reward:-200.0 eps:0.643  
Episode#7200 reward:-200.0 best\_reward:-200.0 eps:0.638  
Episode#7300 reward:-200.0 best\_reward:-200.0 eps:0.633  
Episode#7400 reward:-200.0 best\_reward:-200.0 eps:0.628  
Episode#7500 reward:-200.0 best\_reward:-200.0 eps:0.623  
Episode#7600 reward:-200.0 best\_reward:-200.0 eps:0.618  
Episode#7700 reward:-200.0 best\_reward:-200.0 eps:0.613

Episode#7800 reward:-200.0 best\_reward:-200.0 eps:0.608  
Episode#7900 reward:-200.0 best\_reward:-200.0 eps:0.603  
Episode#8000 reward:-200.0 best\_reward:-200.0 eps:0.598  
Episode#8100 reward:-200.0 best\_reward:-200.0 eps:0.593  
Episode#8200 reward:-200.0 best\_reward:-200.0 eps:0.588  
Episode#8300 reward:-200.0 best\_reward:-200.0 eps:0.583  
Episode#8400 reward:-200.0 best\_reward:-169.0 eps:0.578  
Episode#8500 reward:-200.0 best\_reward:-169.0 eps:0.573  
Episode#8600 reward:-200.0 best\_reward:-169.0 eps:0.568  
Episode#8700 reward:-200.0 best\_reward:-168.0 eps:0.563  
Episode#8800 reward:-200.0 best\_reward:-168.0 eps:0.558  
Episode#8900 reward:-200.0 best\_reward:-168.0 eps:0.553  
Episode#9000 reward:-200.0 best\_reward:-168.0 eps:0.548  
Episode#9100 reward:-200.0 best\_reward:-168.0 eps:0.543  
Episode#9200 reward:-200.0 best\_reward:-168.0 eps:0.538  
Episode#9300 reward:-200.0 best\_reward:-168.0 eps:0.533  
Episode#9400 reward:-200.0 best\_reward:-168.0 eps:0.528  
Episode#9500 reward:-200.0 best\_reward:-168.0 eps:0.523  
Episode#9600 reward:-200.0 best\_reward:-162.0 eps:0.518  
Episode#9700 reward:-200.0 best\_reward:-162.0 eps:0.513  
Episode#9800 reward:-200.0 best\_reward:-162.0 eps:0.508  
Episode#9900 reward:-200.0 best\_reward:-162.0 eps:0.503  
Episode#10000 reward:-200.0 best\_reward:-162.0 eps:0.498  
Episode#10100 reward:-200.0 best\_reward:-162.0 eps:0.493  
Episode#10200 reward:-200.0 best\_reward:-162.0 eps:0.488  
Episode#10300 reward:-200.0 best\_reward:-158.0 eps:0.483  
Episode#10400 reward:-200.0 best\_reward:-158.0 eps:0.478  
Episode#10500 reward:-200.0 best\_reward:-158.0 eps:0.473  
Episode#10600 reward:-200.0 best\_reward:-158.0 eps:0.468  
Episode#10700 reward:-200.0 best\_reward:-158.0 eps:0.463  
Episode#10800 reward:-200.0 best\_reward:-155.0 eps:0.458  
Episode#10900 reward:-200.0 best\_reward:-155.0 eps:0.453  
Episode#11000 reward:-167.0 best\_reward:-153.0 eps:0.448  
Episode#11100 reward:-200.0 best\_reward:-153.0 eps:0.443  
Episode#11200 reward:-200.0 best\_reward:-153.0 eps:0.438  
Episode#11300 reward:-196.0 best\_reward:-153.0 eps:0.433  
Episode#11400 reward:-166.0 best\_reward:-152.0 eps:0.428  
Episode#11500 reward:-200.0 best\_reward:-152.0 eps:0.424  
Episode#11600 reward:-167.0 best\_reward:-151.0 eps:0.419  
Episode#11700 reward:-195.0 best\_reward:-151.0 eps:0.414  
Episode#11800 reward:-200.0 best\_reward:-151.0 eps:0.409  
Episode#11900 reward:-200.0 best\_reward:-151.0 eps:0.404  
Episode#12000 reward:-174.0 best\_reward:-151.0 eps:0.400  
Episode#12100 reward:-194.0 best\_reward:-151.0 eps:0.395  
Episode#12200 reward:-200.0 best\_reward:-151.0 eps:0.390  
Episode#12300 reward:-200.0 best\_reward:-151.0 eps:0.385  
Episode#12400 reward:-159.0 best\_reward:-151.0 eps:0.380  
Episode#12500 reward:-200.0 best\_reward:-151.0 eps:0.375

Episode#12600 reward:-200.0 best\_reward:-151.0 eps:0.371  
Episode#12700 reward:-200.0 best\_reward:-151.0 eps:0.366  
Episode#12800 reward:-200.0 best\_reward:-120.0 eps:0.361  
Episode#12900 reward:-154.0 best\_reward:-120.0 eps:0.356  
Episode#13000 reward:-200.0 best\_reward:-120.0 eps:0.352  
Episode#13100 reward:-200.0 best\_reward:-120.0 eps:0.347  
Episode#13200 reward:-164.0 best\_reward:-120.0 eps:0.343  
Episode#13300 reward:-200.0 best\_reward:-120.0 eps:0.338  
Episode#13400 reward:-188.0 best\_reward:-120.0 eps:0.334  
Episode#13500 reward:-200.0 best\_reward:-120.0 eps:0.329  
Episode#13600 reward:-200.0 best\_reward:-120.0 eps:0.324  
Episode#13700 reward:-200.0 best\_reward:-120.0 eps:0.320  
Episode#13800 reward:-200.0 best\_reward:-120.0 eps:0.315  
Episode#13900 reward:-163.0 best\_reward:-120.0 eps:0.311  
Episode#14000 reward:-171.0 best\_reward:-120.0 eps:0.306  
Episode#14100 reward:-200.0 best\_reward:-120.0 eps:0.301  
Episode#14200 reward:-200.0 best\_reward:-120.0 eps:0.297  
Episode#14300 reward:-200.0 best\_reward:-120.0 eps:0.292  
Episode#14400 reward:-200.0 best\_reward:-120.0 eps:0.288  
Episode#14500 reward:-159.0 best\_reward:-120.0 eps:0.283  
Episode#14600 reward:-200.0 best\_reward:-120.0 eps:0.279  
Episode#14700 reward:-154.0 best\_reward:-120.0 eps:0.274  
Episode#14800 reward:-159.0 best\_reward:-120.0 eps:0.270  
Episode#14900 reward:-200.0 best\_reward:-120.0 eps:0.265  
Episode#15000 reward:-164.0 best\_reward:-120.0 eps:0.261  
Episode#15100 reward:-199.0 best\_reward:-120.0 eps:0.256  
Episode#15200 reward:-159.0 best\_reward:-120.0 eps:0.252  
Episode#15300 reward:-156.0 best\_reward:-120.0 eps:0.248  
Episode#15400 reward:-196.0 best\_reward:-120.0 eps:0.243  
Episode#15500 reward:-190.0 best\_reward:-120.0 eps:0.239  
Episode#15600 reward:-152.0 best\_reward:-117.0 eps:0.235  
Episode#15700 reward:-199.0 best\_reward:-117.0 eps:0.230  
Episode#15800 reward:-154.0 best\_reward:-117.0 eps:0.226  
Episode#15900 reward:-168.0 best\_reward:-115.0 eps:0.222  
Episode#16000 reward:-150.0 best\_reward:-115.0 eps:0.217  
Episode#16100 reward:-200.0 best\_reward:-115.0 eps:0.213  
Episode#16200 reward:-155.0 best\_reward:-103.0 eps:0.208  
Episode#16300 reward:-127.0 best\_reward:-103.0 eps:0.204  
Episode#16400 reward:-160.0 best\_reward:-103.0 eps:0.200  
Episode#16500 reward:-200.0 best\_reward:-103.0 eps:0.195  
Episode#16600 reward:-160.0 best\_reward:-103.0 eps:0.191  
Episode#16700 reward:-163.0 best\_reward:-103.0 eps:0.187  
Episode#16800 reward:-200.0 best\_reward:-103.0 eps:0.183  
Episode#16900 reward:-118.0 best\_reward:-103.0 eps:0.178  
Episode#17000 reward:-160.0 best\_reward:-103.0 eps:0.174  
Episode#17100 reward:-165.0 best\_reward:-103.0 eps:0.170  
Episode#17200 reward:-154.0 best\_reward:-103.0 eps:0.166  
Episode#17300 reward:-157.0 best\_reward:-103.0 eps:0.162

Episode#17400 reward:-153.0 best\_reward:-103.0 eps:0.158  
Episode#17500 reward:-148.0 best\_reward:-103.0 eps:0.154  
Episode#17600 reward:-169.0 best\_reward:-103.0 eps:0.150  
Episode#17700 reward:-157.0 best\_reward:-103.0 eps:0.146  
Episode#17800 reward:-200.0 best\_reward:-103.0 eps:0.141  
Episode#17900 reward:-200.0 best\_reward:-103.0 eps:0.137  
Episode#18000 reward:-200.0 best\_reward:-103.0 eps:0.132  
Episode#18100 reward:-160.0 best\_reward:-103.0 eps:0.128  
Episode#18200 reward:-172.0 best\_reward:-103.0 eps:0.124  
Episode#18300 reward:-189.0 best\_reward:-103.0 eps:0.120  
Episode#18400 reward:-154.0 best\_reward:-103.0 eps:0.115  
Episode#18500 reward:-152.0 best\_reward:-103.0 eps:0.111  
Episode#18600 reward:-200.0 best\_reward:-103.0 eps:0.107  
Episode#18700 reward:-158.0 best\_reward:-103.0 eps:0.103  
Episode#18800 reward:-160.0 best\_reward:-103.0 eps:0.099  
Episode#18900 reward:-200.0 best\_reward:-103.0 eps:0.095  
Episode#19000 reward:-155.0 best\_reward:-103.0 eps:0.091  
Episode#19100 reward:-164.0 best\_reward:-103.0 eps:0.087  
Episode#19200 reward:-156.0 best\_reward:-103.0 eps:0.083  
Episode#19300 reward:-114.0 best\_reward:-103.0 eps:0.079  
Episode#19400 reward:-200.0 best\_reward:-103.0 eps:0.075  
Episode#19500 reward:-178.0 best\_reward:-103.0 eps:0.071  
Episode#19600 reward:-159.0 best\_reward:-103.0 eps:0.067  
Episode#19700 reward:-199.0 best\_reward:-103.0 eps:0.062  
Episode#19800 reward:-155.0 best\_reward:-103.0 eps:0.058  
Episode#19900 reward:-135.0 best\_reward:-103.0 eps:0.054  
Episode#20000 reward:-170.0 best\_reward:-103.0 eps:0.050  
Episode#20100 reward:-153.0 best\_reward:-103.0 eps:0.046  
Episode#20200 reward:-151.0 best\_reward:-103.0 eps:0.042  
Episode#20300 reward:-122.0 best\_reward:-103.0 eps:0.039  
Episode#20400 reward:-164.0 best\_reward:-103.0 eps:0.035  
Episode#20500 reward:-152.0 best\_reward:-103.0 eps:0.031  
Episode#20600 reward:-119.0 best\_reward:-103.0 eps:0.028  
Episode#20700 reward:-150.0 best\_reward:-103.0 eps:0.024  
Episode#20800 reward:-121.0 best\_reward:-103.0 eps:0.020  
Episode#20900 reward:-158.0 best\_reward:-103.0 eps:0.016  
Episode#21000 reward:-149.0 best\_reward:-103.0 eps:0.012  
Episode#21100 reward:-200.0 best\_reward:-103.0 eps:0.008  
Episode#21200 reward:-200.0 best\_reward:-103.0 eps:0.005  
Episode#21300 reward:-145.0 best\_reward:-103.0 eps:0.005  
Episode#21400 reward:-200.0 best\_reward:-103.0 eps:0.005  
Episode#21500 reward:-176.0 best\_reward:-103.0 eps:0.005  
Episode#21600 reward:-193.0 best\_reward:-103.0 eps:0.005  
Episode#21700 reward:-161.0 best\_reward:-100.0 eps:0.005  
Episode#21800 reward:-152.0 best\_reward:-100.0 eps:0.005  
Episode#21900 reward:-200.0 best\_reward:-100.0 eps:0.005  
Episode#22000 reward:-169.0 best\_reward:-100.0 eps:0.005  
Episode#22100 reward:-186.0 best\_reward:-100.0 eps:0.005

Episode#22200 reward:-152.0 best\_reward:-100.0 eps:0.005  
Episode#22300 reward:-145.0 best\_reward:-100.0 eps:0.005  
Episode#22400 reward:-145.0 best\_reward:-100.0 eps:0.005  
Episode#22500 reward:-143.0 best\_reward:-100.0 eps:0.005  
Episode#22600 reward:-150.0 best\_reward:-100.0 eps:0.005  
Episode#22700 reward:-158.0 best\_reward:-100.0 eps:0.005  
Episode#22800 reward:-200.0 best\_reward:-100.0 eps:0.005  
Episode#22900 reward:-135.0 best\_reward:-100.0 eps:0.005  
Episode#23000 reward:-119.0 best\_reward:-100.0 eps:0.005  
Episode#23100 reward:-153.0 best\_reward:-100.0 eps:0.005  
Episode#23200 reward:-152.0 best\_reward:-100.0 eps:0.005  
Episode#23300 reward:-138.0 best\_reward:-100.0 eps:0.005  
Episode#23400 reward:-120.0 best\_reward:-100.0 eps:0.005  
Episode#23500 reward:-114.0 best\_reward:-100.0 eps:0.005  
Episode#23600 reward:-153.0 best\_reward:-100.0 eps:0.005  
Episode#23700 reward:-154.0 best\_reward:-100.0 eps:0.005  
Episode#23800 reward:-150.0 best\_reward:-100.0 eps:0.005  
Episode#23900 reward:-156.0 best\_reward:-100.0 eps:0.005  
Episode#24000 reward:-153.0 best\_reward:-100.0 eps:0.005  
Episode#24100 reward:-151.0 best\_reward:-100.0 eps:0.005  
Episode#24200 reward:-155.0 best\_reward:-100.0 eps:0.005  
Episode#24300 reward:-189.0 best\_reward:-100.0 eps:0.005  
Episode#24400 reward:-190.0 best\_reward:-100.0 eps:0.005  
Episode#24500 reward:-191.0 best\_reward:-100.0 eps:0.005  
Episode#24600 reward:-161.0 best\_reward:-100.0 eps:0.005  
Episode#24700 reward:-159.0 best\_reward:-100.0 eps:0.005  
Episode#24800 reward:-156.0 best\_reward:-100.0 eps:0.005  
Episode#24900 reward:-147.0 best\_reward:-100.0 eps:0.005  
Episode#25000 reward:-151.0 best\_reward:-100.0 eps:0.005  
Episode#25100 reward:-147.0 best\_reward:-100.0 eps:0.005  
Episode#25200 reward:-112.0 best\_reward:-100.0 eps:0.005  
Episode#25300 reward:-154.0 best\_reward:-100.0 eps:0.005  
Episode#25400 reward:-114.0 best\_reward:-100.0 eps:0.005  
Episode#25500 reward:-144.0 best\_reward:-100.0 eps:0.005  
Episode#25600 reward:-131.0 best\_reward:-100.0 eps:0.005  
Episode#25700 reward:-200.0 best\_reward:-100.0 eps:0.005  
Episode#25800 reward:-145.0 best\_reward:-100.0 eps:0.005  
Episode#25900 reward:-151.0 best\_reward:-100.0 eps:0.005  
Episode#26000 reward:-149.0 best\_reward:-100.0 eps:0.005  
Episode#26100 reward:-143.0 best\_reward:-100.0 eps:0.005  
Episode#26200 reward:-156.0 best\_reward:-100.0 eps:0.005  
Episode#26300 reward:-113.0 best\_reward:-100.0 eps:0.005  
Episode#26400 reward:-139.0 best\_reward:-100.0 eps:0.005  
Episode#26500 reward:-117.0 best\_reward:-100.0 eps:0.005  
Episode#26600 reward:-139.0 best\_reward:-100.0 eps:0.005  
Episode#26700 reward:-184.0 best\_reward:-100.0 eps:0.005  
Episode#26800 reward:-196.0 best\_reward:-100.0 eps:0.005  
Episode#26900 reward:-149.0 best\_reward:-100.0 eps:0.005



Episode#27000 reward:-158.0 best\_reward:-100.0 eps:0.005  
Episode#27100 reward:-162.0 best\_reward:-100.0 eps:0.005  
Episode#27200 reward:-140.0 best\_reward:-100.0 eps:0.005  
Episode#27300 reward:-154.0 best\_reward:-100.0 eps:0.005  
Episode#27400 reward:-159.0 best\_reward:-100.0 eps:0.005  
Episode#27500 reward:-112.0 best\_reward:-100.0 eps:0.005  
Episode#27600 reward:-155.0 best\_reward:-100.0 eps:0.005  
Episode#27700 reward:-161.0 best\_reward:-100.0 eps:0.005  
Episode#27800 reward:-147.0 best\_reward:-100.0 eps:0.005  
Episode#27900 reward:-153.0 best\_reward:-100.0 eps:0.005  
Episode#28000 reward:-159.0 best\_reward:-100.0 eps:0.005  
Episode#28100 reward:-143.0 best\_reward:-100.0 eps:0.005  
Episode#28200 reward:-144.0 best\_reward:-100.0 eps:0.005  
Episode#28300 reward:-155.0 best\_reward:-100.0 eps:0.005  
Episode#28400 reward:-193.0 best\_reward:-100.0 eps:0.005  
Episode#28500 reward:-144.0 best\_reward:-100.0 eps:0.005  
Episode#28600 reward:-153.0 best\_reward:-100.0 eps:0.005  
Episode#28700 reward:-117.0 best\_reward:-95.0 eps:0.005  
Episode#28800 reward:-143.0 best\_reward:-95.0 eps:0.005  
Episode#28900 reward:-144.0 best\_reward:-95.0 eps:0.005  
Episode#29000 reward:-159.0 best\_reward:-95.0 eps:0.005  
Episode#29100 reward:-140.0 best\_reward:-95.0 eps:0.005  
Episode#29200 reward:-124.0 best\_reward:-95.0 eps:0.005  
Episode#29300 reward:-153.0 best\_reward:-95.0 eps:0.005  
Episode#29400 reward:-151.0 best\_reward:-95.0 eps:0.005  
Episode#29500 reward:-134.0 best\_reward:-95.0 eps:0.005  
Episode#29600 reward:-143.0 best\_reward:-95.0 eps:0.005  
Episode#29700 reward:-109.0 best\_reward:-95.0 eps:0.005  
Episode#29800 reward:-112.0 best\_reward:-95.0 eps:0.005  
Episode#29900 reward:-109.0 best\_reward:-95.0 eps:0.005  
Episode#30000 reward:-152.0 best\_reward:-95.0 eps:0.005  
Episode#30100 reward:-106.0 best\_reward:-95.0 eps:0.005  
Episode#30200 reward:-143.0 best\_reward:-95.0 eps:0.005  
Episode#30300 reward:-153.0 best\_reward:-95.0 eps:0.005  
Episode#30400 reward:-143.0 best\_reward:-95.0 eps:0.005  
Episode#30500 reward:-150.0 best\_reward:-95.0 eps:0.005  
Episode#30600 reward:-116.0 best\_reward:-95.0 eps:0.005  
Episode#30700 reward:-164.0 best\_reward:-95.0 eps:0.005  
Episode#30800 reward:-177.0 best\_reward:-94.0 eps:0.005  
Episode#30900 reward:-152.0 best\_reward:-94.0 eps:0.005  
Episode#31000 reward:-164.0 best\_reward:-94.0 eps:0.005  
Episode#31100 reward:-200.0 best\_reward:-94.0 eps:0.005  
Episode#31200 reward:-154.0 best\_reward:-94.0 eps:0.005  
Episode#31300 reward:-138.0 best\_reward:-94.0 eps:0.005  
Episode#31400 reward:-160.0 best\_reward:-90.0 eps:0.005  
Episode#31500 reward:-165.0 best\_reward:-90.0 eps:0.005  
Episode#31600 reward:-172.0 best\_reward:-90.0 eps:0.005  
Episode#31700 reward:-115.0 best\_reward:-90.0 eps:0.005

Episode#31800 reward:-144.0 best\_reward:-90.0 eps:0.005  
Episode#31900 reward:-143.0 best\_reward:-90.0 eps:0.005  
Episode#32000 reward:-145.0 best\_reward:-90.0 eps:0.005  
Episode#32100 reward:-145.0 best\_reward:-90.0 eps:0.005  
Episode#32200 reward:-150.0 best\_reward:-90.0 eps:0.005  
Episode#32300 reward:-128.0 best\_reward:-90.0 eps:0.005  
Episode#32400 reward:-135.0 best\_reward:-90.0 eps:0.005  
Episode#32500 reward:-150.0 best\_reward:-90.0 eps:0.005  
Episode#32600 reward:-149.0 best\_reward:-90.0 eps:0.005  
Episode#32700 reward:-149.0 best\_reward:-90.0 eps:0.005  
Episode#32800 reward:-114.0 best\_reward:-90.0 eps:0.005  
Episode#32900 reward:-146.0 best\_reward:-90.0 eps:0.005  
Episode#33000 reward:-119.0 best\_reward:-90.0 eps:0.005  
Episode#33100 reward:-142.0 best\_reward:-90.0 eps:0.005  
Episode#33200 reward:-113.0 best\_reward:-90.0 eps:0.005  
Episode#33300 reward:-150.0 best\_reward:-90.0 eps:0.005  
Episode#33400 reward:-154.0 best\_reward:-90.0 eps:0.005  
Episode#33500 reward:-111.0 best\_reward:-90.0 eps:0.005  
Episode#33600 reward:-145.0 best\_reward:-90.0 eps:0.005  
Episode#33700 reward:-113.0 best\_reward:-90.0 eps:0.005  
Episode#33800 reward:-141.0 best\_reward:-90.0 eps:0.005  
Episode#33900 reward:-141.0 best\_reward:-90.0 eps:0.005  
Episode#34000 reward:-149.0 best\_reward:-90.0 eps:0.005  
Episode#34100 reward:-113.0 best\_reward:-90.0 eps:0.005  
Episode#34200 reward:-112.0 best\_reward:-90.0 eps:0.005  
Episode#34300 reward:-143.0 best\_reward:-90.0 eps:0.005  
Episode#34400 reward:-119.0 best\_reward:-90.0 eps:0.005  
Episode#34500 reward:-118.0 best\_reward:-90.0 eps:0.005  
Episode#34600 reward:-160.0 best\_reward:-90.0 eps:0.005  
Episode#34700 reward:-164.0 best\_reward:-87.0 eps:0.005  
Episode#34800 reward:-129.0 best\_reward:-87.0 eps:0.005  
Episode#34900 reward:-119.0 best\_reward:-87.0 eps:0.005  
Episode#35000 reward:-162.0 best\_reward:-87.0 eps:0.005  
Episode#35100 reward:-169.0 best\_reward:-87.0 eps:0.005  
Episode#35200 reward:-111.0 best\_reward:-87.0 eps:0.005  
Episode#35300 reward:-156.0 best\_reward:-85.0 eps:0.005  
Episode#35400 reward:-139.0 best\_reward:-85.0 eps:0.005  
Episode#35500 reward:-164.0 best\_reward:-85.0 eps:0.005  
Episode#35600 reward:-87.0 best\_reward:-85.0 eps:0.005  
Episode#35700 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#35800 reward:-150.0 best\_reward:-85.0 eps:0.005  
Episode#35900 reward:-87.0 best\_reward:-85.0 eps:0.005  
Episode#36000 reward:-161.0 best\_reward:-85.0 eps:0.005  
Episode#36100 reward:-163.0 best\_reward:-85.0 eps:0.005  
Episode#36200 reward:-146.0 best\_reward:-85.0 eps:0.005  
Episode#36300 reward:-200.0 best\_reward:-85.0 eps:0.005  
Episode#36400 reward:-174.0 best\_reward:-85.0 eps:0.005  
Episode#36500 reward:-136.0 best\_reward:-85.0 eps:0.005

Episode#36600 reward:-152.0 best\_reward:-85.0 eps:0.005  
Episode#36700 reward:-153.0 best\_reward:-85.0 eps:0.005  
Episode#36800 reward:-141.0 best\_reward:-85.0 eps:0.005  
Episode#36900 reward:-157.0 best\_reward:-85.0 eps:0.005  
Episode#37000 reward:-197.0 best\_reward:-85.0 eps:0.005  
Episode#37100 reward:-172.0 best\_reward:-85.0 eps:0.005  
Episode#37200 reward:-147.0 best\_reward:-85.0 eps:0.005  
Episode#37300 reward:-135.0 best\_reward:-85.0 eps:0.005  
Episode#37400 reward:-117.0 best\_reward:-85.0 eps:0.005  
Episode#37500 reward:-152.0 best\_reward:-85.0 eps:0.005  
Episode#37600 reward:-120.0 best\_reward:-85.0 eps:0.005  
Episode#37700 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#37800 reward:-148.0 best\_reward:-85.0 eps:0.005  
Episode#37900 reward:-149.0 best\_reward:-85.0 eps:0.005  
Episode#38000 reward:-140.0 best\_reward:-85.0 eps:0.005  
Episode#38100 reward:-142.0 best\_reward:-85.0 eps:0.005  
Episode#38200 reward:-140.0 best\_reward:-85.0 eps:0.005  
Episode#38300 reward:-139.0 best\_reward:-85.0 eps:0.005  
Episode#38400 reward:-141.0 best\_reward:-85.0 eps:0.005  
Episode#38500 reward:-109.0 best\_reward:-85.0 eps:0.005  
Episode#38600 reward:-186.0 best\_reward:-85.0 eps:0.005  
Episode#38700 reward:-190.0 best\_reward:-85.0 eps:0.005  
Episode#38800 reward:-139.0 best\_reward:-85.0 eps:0.005  
Episode#38900 reward:-200.0 best\_reward:-85.0 eps:0.005  
Episode#39000 reward:-200.0 best\_reward:-85.0 eps:0.005  
Episode#39100 reward:-162.0 best\_reward:-85.0 eps:0.005  
Episode#39200 reward:-193.0 best\_reward:-85.0 eps:0.005  
Episode#39300 reward:-110.0 best\_reward:-85.0 eps:0.005  
Episode#39400 reward:-118.0 best\_reward:-85.0 eps:0.005  
Episode#39500 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#39600 reward:-125.0 best\_reward:-85.0 eps:0.005  
Episode#39700 reward:-166.0 best\_reward:-85.0 eps:0.005  
Episode#39800 reward:-137.0 best\_reward:-85.0 eps:0.005  
Episode#39900 reward:-138.0 best\_reward:-85.0 eps:0.005  
Episode#40000 reward:-160.0 best\_reward:-85.0 eps:0.005  
Episode#40100 reward:-149.0 best\_reward:-85.0 eps:0.005  
Episode#40200 reward:-109.0 best\_reward:-85.0 eps:0.005  
Episode#40300 reward:-142.0 best\_reward:-85.0 eps:0.005  
Episode#40400 reward:-133.0 best\_reward:-85.0 eps:0.005  
Episode#40500 reward:-149.0 best\_reward:-85.0 eps:0.005  
Episode#40600 reward:-108.0 best\_reward:-85.0 eps:0.005  
Episode#40700 reward:-142.0 best\_reward:-85.0 eps:0.005  
Episode#40800 reward:-143.0 best\_reward:-85.0 eps:0.005  
Episode#40900 reward:-173.0 best\_reward:-85.0 eps:0.005  
Episode#41000 reward:-144.0 best\_reward:-85.0 eps:0.005  
Episode#41100 reward:-144.0 best\_reward:-85.0 eps:0.005  
Episode#41200 reward:-118.0 best\_reward:-85.0 eps:0.005  
Episode#41300 reward:-135.0 best\_reward:-85.0 eps:0.005

Episode#41400 reward:-134.0 best\_reward:-85.0 eps:0.005  
Episode#41500 reward:-135.0 best\_reward:-85.0 eps:0.005  
Episode#41600 reward:-137.0 best\_reward:-85.0 eps:0.005  
Episode#41700 reward:-143.0 best\_reward:-85.0 eps:0.005  
Episode#41800 reward:-154.0 best\_reward:-85.0 eps:0.005  
Episode#41900 reward:-147.0 best\_reward:-85.0 eps:0.005  
Episode#42000 reward:-112.0 best\_reward:-85.0 eps:0.005  
Episode#42100 reward:-143.0 best\_reward:-85.0 eps:0.005  
Episode#42200 reward:-143.0 best\_reward:-85.0 eps:0.005  
Episode#42300 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#42400 reward:-111.0 best\_reward:-85.0 eps:0.005  
Episode#42500 reward:-152.0 best\_reward:-85.0 eps:0.005  
Episode#42600 reward:-148.0 best\_reward:-85.0 eps:0.005  
Episode#42700 reward:-111.0 best\_reward:-85.0 eps:0.005  
Episode#42800 reward:-152.0 best\_reward:-85.0 eps:0.005  
Episode#42900 reward:-113.0 best\_reward:-85.0 eps:0.005  
Episode#43000 reward:-143.0 best\_reward:-85.0 eps:0.005  
Episode#43100 reward:-118.0 best\_reward:-85.0 eps:0.005  
Episode#43200 reward:-115.0 best\_reward:-85.0 eps:0.005  
Episode#43300 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#43400 reward:-117.0 best\_reward:-85.0 eps:0.005  
Episode#43500 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#43600 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#43700 reward:-115.0 best\_reward:-85.0 eps:0.005  
Episode#43800 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#43900 reward:-154.0 best\_reward:-85.0 eps:0.005  
Episode#44000 reward:-153.0 best\_reward:-85.0 eps:0.005  
Episode#44100 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#44200 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#44300 reward:-136.0 best\_reward:-85.0 eps:0.005  
Episode#44400 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#44500 reward:-139.0 best\_reward:-85.0 eps:0.005  
Episode#44600 reward:-151.0 best\_reward:-85.0 eps:0.005  
Episode#44700 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#44800 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#44900 reward:-145.0 best\_reward:-85.0 eps:0.005  
Episode#45000 reward:-111.0 best\_reward:-85.0 eps:0.005  
Episode#45100 reward:-112.0 best\_reward:-85.0 eps:0.005  
Episode#45200 reward:-144.0 best\_reward:-85.0 eps:0.005  
Episode#45300 reward:-111.0 best\_reward:-85.0 eps:0.005  
Episode#45400 reward:-137.0 best\_reward:-85.0 eps:0.005  
Episode#45500 reward:-112.0 best\_reward:-85.0 eps:0.005  
Episode#45600 reward:-113.0 best\_reward:-85.0 eps:0.005  
Episode#45700 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#45800 reward:-158.0 best\_reward:-85.0 eps:0.005  
Episode#45900 reward:-154.0 best\_reward:-85.0 eps:0.005  
Episode#46000 reward:-113.0 best\_reward:-85.0 eps:0.005  
Episode#46100 reward:-146.0 best\_reward:-85.0 eps:0.005

Episode#46200 reward:-116.0 best\_reward:-85.0 eps:0.005  
Episode#46300 reward:-114.0 best\_reward:-85.0 eps:0.005  
Episode#46400 reward:-117.0 best\_reward:-85.0 eps:0.005  
Episode#46500 reward:-115.0 best\_reward:-85.0 eps:0.005  
Episode#46600 reward:-117.0 best\_reward:-85.0 eps:0.005  
Episode#46700 reward:-168.0 best\_reward:-85.0 eps:0.005  
Episode#46800 reward:-152.0 best\_reward:-85.0 eps:0.005  
Episode#46900 reward:-108.0 best\_reward:-85.0 eps:0.005  
Episode#47000 reward:-110.0 best\_reward:-85.0 eps:0.005  
Episode#47100 reward:-115.0 best\_reward:-85.0 eps:0.005  
Episode#47200 reward:-148.0 best\_reward:-85.0 eps:0.005  
Episode#47300 reward:-142.0 best\_reward:-85.0 eps:0.005  
Episode#47400 reward:-109.0 best\_reward:-85.0 eps:0.005  
Episode#47500 reward:-160.0 best\_reward:-85.0 eps:0.005  
Episode#47600 reward:-152.0 best\_reward:-85.0 eps:0.005  
Episode#47700 reward:-151.0 best\_reward:-85.0 eps:0.005  
Episode#47800 reward:-150.0 best\_reward:-85.0 eps:0.005  
Episode#47900 reward:-109.0 best\_reward:-85.0 eps:0.005  
Episode#48000 reward:-109.0 best\_reward:-85.0 eps:0.005  
Episode#48100 reward:-113.0 best\_reward:-85.0 eps:0.005  
Episode#48200 reward:-112.0 best\_reward:-85.0 eps:0.005  
Episode#48300 reward:-109.0 best\_reward:-85.0 eps:0.005  
Episode#48400 reward:-168.0 best\_reward:-85.0 eps:0.005  
Episode#48500 reward:-108.0 best\_reward:-85.0 eps:0.005  
Episode#48600 reward:-197.0 best\_reward:-85.0 eps:0.005  
Episode#48700 reward:-106.0 best\_reward:-85.0 eps:0.005  
Episode#48800 reward:-108.0 best\_reward:-85.0 eps:0.005  
Episode#48900 reward:-108.0 best\_reward:-85.0 eps:0.005  
Episode#49000 reward:-118.0 best\_reward:-85.0 eps:0.005  
Episode#49100 reward:-118.0 best\_reward:-85.0 eps:0.005  
Episode#49200 reward:-108.0 best\_reward:-85.0 eps:0.005  
Episode#49300 reward:-120.0 best\_reward:-85.0 eps:0.005  
Episode#49400 reward:-129.0 best\_reward:-85.0 eps:0.005  
Episode#49500 reward:-118.0 best\_reward:-85.0 eps:0.005  
Episode#49600 reward:-155.0 best\_reward:-85.0 eps:0.005  
Episode#49700 reward:-115.0 best\_reward:-85.0 eps:0.005  
Episode#49800 reward:-151.0 best\_reward:-85.0 eps:0.005  
Episode#49900 reward:-118.0 best\_reward:-85.0 eps:0.005  
Moviepy - Building video /content/videos/rl-video-episode-0.mp4.  
Moviepy - Writing video /content/videos/rl-video-episode-0.mp4

Moviepy - Done !  
Moviepy - video ready /content/videos/rl-video-episode-0.mp4  
Test episode reward: -111.0  
Moviepy - Building video /content/videos/rl-video-episode-1.mp4.  
Moviepy - Writing video /content/videos/rl-video-episode-1.mp4

```
Moviepy - Done !
Moviepy - video ready /content/videos/rl-video-episode-1.mp4
Test episode reward: -110.0
Test episode reward: -100.0
Test episode reward: -148.0
Test episode reward: -147.0
Test episode reward: -111.0
Test episode reward: -109.0
Test episode reward: -111.0
Moviepy - Building video /content/videos/rl-video-episode-8.mp4.
Moviepy - Writing video /content/videos/rl-video-episode-8.mp4
```

```
Moviepy - Done !
Moviepy - video ready /content/videos/rl-video-episode-8.mp4
Test episode reward: -111.0
Test episode reward: -111.0
Test episode reward: -97.0
Test episode reward: -100.0
Test episode reward: -111.0
Test episode reward: -112.0
Test episode reward: -111.0
Test episode reward: -156.0
Test episode reward: -143.0
Test episode reward: -100.0
Test episode reward: -153.0
Test episode reward: -112.0
```

## 2 Results

In this project, two reinforcement learning algorithms, Q-Learning and SARSA, were applied to solve the Mountain Car problem. The results demonstrate distinct learning patterns and final performances for each method, illustrating their unique strengths and weaknesses.

### Q-Learning vs. SARSA Performance

#### 1. Learning Behavior:

- **Q-Learning:** In the early stages, Q-Learning exhibited faster learning due to its off-policy nature, which allows the agent to maximize long-term rewards by considering actions outside the current policy. This led to more aggressive exploration, yielding higher episodic rewards earlier in the training.

- **SARSA:** On the other hand, SARSA, being an on-policy method, tended to be more conservative in its exploration. This approach made the learning process more stable but resulted in slower initial progress compared to Q-Learning. SARSA gradually improved its performance as the number of episodes increased.
2. **Convergence and Stability:**
- **Q-Learning** displayed more fluctuation in rewards as it focused on maximizing future rewards even at the risk of taking suboptimal actions during training. Despite these fluctuations, Q-Learning converged faster in terms of finding a solution that allowed the car to reach the top of the hill.
  - **SARSA** had a smoother learning curve with fewer fluctuations, showing a more stable convergence pattern. This is attributed to SARSA's conservative nature, which ensures it sticks closely to the policy being evaluated, reducing the chances of drastic changes in learned behavior.
3. **Exploration and Exploitation:**
- The exploration rate (epsilon) decreased steadily in both methods, indicating that both algorithms reduced exploration over time as the policy became more stable. By the end of the training, both Q-Learning and SARSA were operating with minimal exploration, focusing on exploiting the learned policy.

**Conclusion** The results show that Q-Learning outperforms SARSA in terms of speed of convergence due to its off-policy nature, which allows for more aggressive learning. However, SARSA's on-policy approach results in a more stable learning process, which may be preferable in environments where stability is critical. Overall, the choice between Q-Learning and SARSA should be guided by the specific requirements of the problem at hand, such as the need for speed or stability in learning.

This analysis highlights the trade-offs between these two algorithms, providing insights into their behavior and suitability for solving reinforcement learning problems.