In [603…

```
import gym
import numpy as np
import matplotlib.pyplot as plt
from matplotlib.pyplot import figure
from gym import spaces
import math
```

In [ ]:

```
G = 6.6743e-11 # N m2 kg-2
EARTH_M_KG = 5.9722e24 # kg
EARTH_R_KM = 6.378137e3 # km WGS84 semi-major axis
EARTH_R_M = 6.378137e6 # m WGS84 semi-major axis
g = 9.802 # m s-2



# return the value of the acceleration due to gravity of EARTH
# at specified altitude in km
EARTH_R_KM_INVERSE = 1/EARTH_R_KM
def calc_gravity(height_above_earth_km):
    return g/(1 + height_above_earth_km*EARTH_R_KM_INVERSE)**2

def drag_deceleration(height_above_earth_km, orbital_velocity_kms, mass_kg, drag_coe
    return 0.5**drag_coeff*(orbital_velocity_kms**2)*drag_area / mass_kg

def physics_step(curr_pos, curr_vel):
    curr_alt = np.linalg.norm(curr_pos)
    curr_theta = np.arctan2(curr_pos[1], curr_pos[0])

    grav = calc_gravity(curr_alt)

    a = grav*np.array([np.cos(curr_theta), np.sin(curr_theta)]) # - dec*np.array([np.
    v = curr_vel + a*dt
    x = curr_pos + v*dt

    return v, x
```

In [641…

```
G = 6.6743e-11 # N m2 kg-2

class OrbitSimEnv(gym.Env):

    def __init__(self, random_distance=False, random_angle=False, random_distance_mi
        self._random_distance = random_distance
        self._random_angle = random_angle
        self._random_dist_minmax = random_distance_min_max

        # "Earth" object properties
        self.EARTH_MASS = 9.8e10
        self.EARTH_POS = np.array([0,0])

        # "Spacecraft, sc" object state
        self._SC_MASS = 1
        self._sc_init_pos = 2
        self._sc_init_vel = 1.808
        self._sc_init_delta_v = 10.0

        self._reset()
        self.calc_pos()
        self.calc_vel()

#         self._sc_pos_mag = np.linalg.norm(self._sc_pos)
#         self._sc_pos_angle = np.arctan2(self._sc_pos[1], self._sc_pos[0])
#         self._sc_vel_mag = np.linalg.norm(self._sc_vel)
#         self._sc_vel_angle = np.arctan2(self._sc_vel[1], self._sc_vel[0])

        self.dt = 0.01 # step size

        # actions: do nothing, fire 0.01, fire 0.1
        self.action_space = spaces.Discrete(2)
        self._action_to_impulse = {
            0: 0,
            1: 0.01,
            #2: 0.1
        }


        # end when distance is <1 or >5, when delta_v is 0

        # rewards +1 for living each time step
        # -1000 for hitting earth or leaving
        # -(abs(distance - 2))
        # + delta_v remaining

    def _get_obs(self):
        return np.array([self._sc_pos_mag, self._sc_vel_mag, self._sc_vel_angle]) # s

    def _get_info(self):
        return None

    def get_state(self):
        return self._sc_pos, self._sc_vel, self._sc_delta_v

    def reset(self, seed=None, options=None):
        # We need the following line to seed self.np_random
        super().reset(seed=seed)
```

```
        self._reset()

        observation = self._get_obs()
        info = self._get_info()

        return observation, info

    def _calc_init_pos_vel(self):
        self._sc_init_pos = 0.1*np.random.randint(self._random_dist_minmax[0],self._r
        self._sc_init_vel = np.sqrt(G*self.EARTH_MASS / self._sc_init_pos)

    def _reset(self):
        # random starting angle
        if self._random_angle:
            angle = np.random.rand()*np.pi*2
        else:
            angle = 0

        if self._random_distance:
            self._calc_init_pos_vel()

        #self._sc_init_delta_v = 1.0*np.random.randint(1, 10)
        self._sc_pos = self._sc_init_pos*np.array([np.cos(angle), np.sin(angle)])
        self._sc_vel = self._sc_init_vel*np.array([-np.sin(angle), np.cos(angle)])
        self._sc_delta_v = self._sc_init_delta_v

    def is_fuel_empty(self):
        return self._sc_delta_v <= 0.0

    def step(self, action):
        # impulse can only happen if we have fuel remaining
        if (not self.is_fuel_empty()):
            delta_v = self._action_to_impulse[action]
            self._sc_delta_v -= delta_v
            self._sc_vel += delta_v*(np.array(-self._sc_pos[1], self._sc_pos[0])/self

        grav = self.calc_grav()
        drag = self.calc_drag()
        a = grav + drag

        self._sc_vel += a*self.dt
        self._sc_pos += self._sc_vel*self.dt

        self.calc_pos()
        self.calc_vel()

        reward = 0.0
        done = False

        if self._sc_pos_mag < 1 or self._sc_pos_mag > 5:
            done = True
            reward = -657 # penalty for hitting earth/boundry
        else:
            reward += 1 # for living
            #reward -= abs(self._sc_pos_mag - self._sc_init_pos)**2 # pentaly for div
            #reward += self._sc_delta_v # reward for conserving delta_v

            #         elif self._sc_delta_v <= 0:
```

```
#             done = True
#             reward = -200 # pentaly for running out of fuel

        observation = self._get_obs()
        info = self._get_info()

        return observation, reward, done, False, info

    def calc_pos(self):
        self._sc_pos_mag = np.linalg.norm(self._sc_pos)
        self._sc_pos_angle = np.arctan2(self._sc_pos[1], self._sc_pos[0])

    def calc_vel(self):
        self._sc_vel_mag = np.linalg.norm(self._sc_vel)
        self._sc_vel_angle = np.arctan2(self._sc_vel[1], self._sc_vel[0])

    def calc_grav(self):
        return -(G*self._SC_MASS*self.EARTH_MASS*self._sc_pos)/(np.linalg.norm(self._

    def calc_drag(self):
        return -0.1*(self._sc_vel/self._sc_vel_mag)
```

In [626…

```
orbitEnv = OrbitSimEnv()
```

In [48]:

```
class DiscretObs():


    def __init__(self, bins_list):
        self._bins_list = bins_list

        self._bins_num = len(bins_list)
        self._state_num_list = [len(bins)+1 for bins in bins_list]
        self._state_num_total = np.prod(self._state_num_list)

    def get_state_num_total(self):

        return self._state_num_total

    def _state_num_list(self):

        return self._state_num_list

    def obs2state(self, obs):

        if not len(obs)==self._bins_num:
            raise ValueError("observation must have length {}".format(self._bins_num)
        else:
            return [np.digitize(obs[i], bins=self._bins_list[i]) for i in range(self.

    def obs2idx(self, obs):

        state = self.obs2state(obs)

        return self.state2idx(state)

    def state2idx(self, state):

        idx = 0
        for i in range(self._bins_num-1,-1,-1):
            idx = idx*self._state_num_list[i]+state[i]

        return idx

    def idx2state(self, idx):

        state = [None]*self._bins_num
        state_num_cumul = np.cumprod(self._state_num_list)
        for i in range(self._bins_num-1,0,-1):
            state[i] = idx//state_num_cumul[i-1]
            idx -=state[i]*state_num_cumul[i-1]
        state[0] = idx%state_num_cumul[0]

        return state
```

In [358…

```
import math
# Recommended epsilon and learning_rate update (Feel free to modify existing and add
def get_epsilon2(t):
    return max(0.05, min(1., 1. - math.log10((t + 1) / 200)))

def get_learning_rate(t):
    return max(0.6, min(1., 1. - math.log10((t + 1) / 200)))
```

In [407…

```
## Suggested functions (Feel free to modify existing and add new functions)
def randargmax(b,**kw):
  """ a random tie-breaking argmax"""
  return np.argmax(np.random.random(b.shape) * (b == np.amax(b,**kw, keepdims=True)),

def update_Q(Q, current_idx, next_idx, current_action, next_action, alpha, R, gamma):
    # Update Q at the each step
    #
    # input:  current Q,                    (array)
    #         current_idx, next_idx         (array)  states
    #         current_action, next_action   (array)  actions
    #         alpha, R, gamma               (floats) learning rate, reward, discount
    # output: Updated Q
    #

    current_Q = Q[current_idx, current_action]
    next_Q = Q[next_idx, next_action]

    Q[current_idx,current_action] = current_Q + alpha*(R + gamma*next_Q - current_Q)
    return Q

def get_action(current_idx, Q, epsilon, num_actions):

    # Choose optimal action based on current state and Q
    #
    # input:  current_idx     (array)
    #         Q,              (array)
    #         epsilon,        (float)
    # output: action
    p = np.random.random()
    if p < epsilon:
        action = np.random.randint(num_actions)
    else:
        action = randargmax(Q[current_idx])
    return action
```

In [ ]:

```
## SARSA implementation
total_reward = 0

bins_pos = []
bins_d_pos = []
bins_ang = np.linspace(-0.41887903,0.41887903,5)
bins_d_ang = np.linspace(-0.87266,0.87266,11)

dobs = DiscretObs([bins_pos,bins_d_pos,bins_ang,bins_d_ang])

env = gym.make('CartPole-v1')
observation = env.reset()

# Q defined by states
#Q = np.zeros((2,dobs._state_num_list[0],dobs._state_num_list[1],dobs._state_num_list
# Q defined by index
sarsa_cartpole_Q = np.zeros((dobs.get_state_num_total(), 2))

count = 0

gamma = 0.98
sarsa_result = np.zeros(50)
s = 0
for ep in range(1000):
    if np.mod(ep,20)==0:
        sarsa_result[s] = total_reward/20
        s+=1
        total_reward = 0

    observation, other = env.reset()

    current_state = dobs.obs2state(observation)
    current_idx = dobs.obs2idx(observation)

    alpha = 0.9 #get_learning_rate(ep)
    epsilon = get_epsilon(ep)


    done = False

    while not done:
        total_reward += 1
        action = get_action(current_idx, sarsa_cartpole_Q, epsilon)
        observation, reward, done, info, other = env.step(action)

        next_idx = dobs.obs2idx(observation)
        next_state = dobs.obs2state(observation)
        next_action = get_action(next_idx, sarsa_cartpole_Q, epsilon)

        sarsa_cartpole_Q = update_Q(sarsa_cartpole_Q, current_idx, next_idx, action,
        current_idx = next_idx
```

In [392…

```
# Q defined by states
#Q = np.zeros((2,dobs._state_num_list[0],dobs._state_num_list[1],dobs._state_num_list
# Q defined by index
Q_qlearning = np.zeros((0,0))
qlearning_result = np.zeros(1)

def set_Q_qlearning(state_num_total, num_actions):
    return np.zeros((state_num_total, num_actions))
```

In [619…

```
## Suggested flow (Feel free to modify and add)
## Q-Learning
# state: [self._sc_pos_mag, self._sc_pos_angle, self._sc_vel_mag, self._sc_vel_angle,

def q_learning_algo(dobs_q, orbitEnv, gamma=0.98, alpha=None, epsilon=None, episode_
    max_steps = int(episode_count/report_step)

    print(f"Running Q-Learning Algorithm with gamma={gamma}, alpha={alpha}, epsilon=

    if alpha == None:
        def get_alpha(t):
            return max(0.7, min(1., 1. - math.log10((t + 1) / 200)))
    else:
        def get_alpha(t):
            return alpha

    if epsilon == None:
        def get_epsilon(t):
            return max(0.05, min(1., 1. - math.log10((t + 1) / 300)))
    else:
        def get_epsilon(t):
            return epsilon

#     bins_distance = np.linspace(0,5,50) # state for each 0.1 distance
#     bins_theta = np.linspace(0,360, 180) # state for every 2 degrees
#     bins_vel_mag = np.linspace(0,4,40) # 0.01
#     bins_vel_angle = np.linspace(0,360, 360) # doesn't need as fine a line, 6 deg
#     bins_delta_v = np.linspace(0, 10, 50) # 0.1

#     dobs_q = DiscretObs([bins_distance,bins_theta,bins_delta_v]) #bins_delta_v, bir

    observation, info = orbitEnv.reset()

    step = 0
    total_reward = 0
    global qlearning_result
    qlearning_result = np.zeros(max_steps)

    global Q_qlearning

    for ep in range(episode_count):
        alpha = get_alpha(ep)
        epsilon = get_epsilon(ep)

        if np.mod(ep,report_step)==0:
            print(f"Step {step}/{max_steps}\tReward: {total_reward/report_step}\tAlpl
            qlearning_result[step] = total_reward/report_step
            step += 1
            total_reward = 0

        observation, other = orbitEnv.reset()

        current_state = dobs_q.obs2state(observation)
        current_idx = dobs_q.obs2idx(observation)

        done = False

        ep_reward = 0
```

```
        while not done:
            action = get_action(current_idx, Q_qlearning, epsilon, orbitEnv.action_sp

            # if we are out of delta-v, only action is do nothing
            if (orbitEnv.is_fuel_empty()):
                action = 0

            observation, reward, done, info, other = orbitEnv.step(action)

            next_idx = dobs_q.obs2idx(observation)
            next_state = dobs_q.obs2state(observation)
            next_action = get_action(next_idx, Q_qlearning, 0, orbitEnv.action_space.

            Q_qlearning = update_Q(Q_qlearning, current_idx, next_idx, action, next_a
            current_idx = next_idx

            ep_reward += reward

        total_reward += ep_reward

    #return Q_qlearning, qlearning_result
```

In [627…

```
bins_distance = np.linspace(0,5,50) # state for each 0.1 distance
bins_theta = np.linspace(-np.pi,np.pi, 90) # state for every 4 degrees
bins_vel_mag = np.linspace(1,4,60) # 0.05
bins_vel_angle = np.linspace(-np.pi,np.pi, 90) # +/- 45 degrees from tangent to accel
bins_delta_v = np.linspace(0, 10, 50) # 0.1

bins_vel_x = np.linspace(-3,3,30)
bins_vel_y = np.linspace(-3,3,30)


dobs_q = DiscretObs([bins_distance,bins_vel_mag,bins_vel_angle]) #bins_delta_v, bins_

orbitEnv = OrbitSimEnv(random_angle=True, random_distance=True)

Q_qlearning = set_Q_qlearning(dobs_q.get_state_num_total(), 2)
q_learning_algo(dobs_q, orbitEnv, gamma=0.98)
ql_Q_98_60_20 = np.copy(Q_qlearning)
ql_result_98_60_20 = np.copy(qlearning_result)
```

```
Running Q-Learning Algorithm with gamma=0.98, alpha=None, epsilon=None, ep_count=1000
0, report_step=50
Step 0/200      Reward: 0.0     Alpha:1.0       Epsilon:1.0
Step 1/200      Reward: 363.36  Alpha:1.0       Epsilon:1.0
Step 2/200      Reward: 386.04  Alpha:1.0       Epsilon:1.0
Step 3/200      Reward: 395.06  Alpha:1.0       Epsilon:1.0
Step 4/200      Reward: 367.64  Alpha:0.9978339382434924        Epsilon:1.0
Step 5/200      Reward: 385.26  Alpha:0.9013562741829431        Epsilon:1.0
Step 6/200      Reward: 358.52  Alpha:0.8224635000701379        Epsilon:0.99855475912
5819
Step 7/200      Reward: 357.98  Alpha:0.7557228791981572        Epsilon:0.93181413825
38384
Step 8/200      Reward: 359.64  Alpha:0.7        Epsilon:0.8739768820994801
Step 9/200      Reward: 379.9   Alpha:0.7       Epsilon:0.8229447128417019
Step 10/200     Reward: 324.2   Alpha:0.7       Epsilon:0.7772835288524167
Step 11/200     Reward: 364.56  Alpha:0.7       Epsilon:0.7359696558678774
```

```
Step 12/200     Reward: 435.78  Alpha:0.7       Epsilon:0.698246782716923
Step 13/200     Reward: 435.24  Alpha:0.7       Epsilon:0.6635402661514704
Step 14/200     Reward: 453.02  Alpha:0.7       Epsilon:0.6314032367530038
Step 15/200     Reward: 452.94  Alpha:0.7       Epsilon:0.6014813177154941
Step 16/200     Reward: 401.9   Alpha:0.7       Epsilon:0.5734887386354248
Step 17/200     Reward: 454.86  Alpha:0.7       Epsilon:0.5471916946350746
Step 18/200     Reward: 487.4   Alpha:0.7       Epsilon:0.5223964637405994
Step 19/200     Reward: 451.58  Alpha:0.7       Epsilon:0.49894073778224857
Step 20/200     Reward: 395.36  Alpha:0.7       Epsilon:0.4766871772403438
Step 21/200     Reward: 448.24  Alpha:0.7       Epsilon:0.4555185386914202
Step 22/200     Reward: 395.84  Alpha:0.7       Epsilon:0.4353339357479107
Step 23/200     Reward: 466.66  Alpha:0.7       Epsilon:0.41604593108987076
Step 24/200     Reward: 435.86  Alpha:0.7       Epsilon:0.3975782473167564
Step 25/200     Reward: 406.34  Alpha:0.7       Epsilon:0.37986394502624243
Step 26/200     Reward: 434.48  Alpha:0.7       Epsilon:0.36284395815807613
Step 27/200     Reward: 417.88  Alpha:0.7       Epsilon:0.3464659056976319
Step 28/200     Reward: 460.16  Alpha:0.7       Epsilon:0.33068311943388784
Step 29/200     Reward: 450.88  Alpha:0.7       Epsilon:0.3154538422819265
Step 30/200     Reward: 453.18  Alpha:0.7       Epsilon:0.300740562476392
Step 31/200     Reward: 461.56  Alpha:0.7       Epsilon:0.28650945690605756
Step 32/200     Reward: 443.08  Alpha:0.7       Epsilon:0.2727299228003627
Step 33/200     Reward: 461.76  Alpha:0.7       Epsilon:0.25937418145686886
Step 34/200     Reward: 418.06  Alpha:0.7       Epsilon:0.24641694110709345
Step 35/200     Reward: 429.26  Alpha:0.7       Epsilon:0.23383510863621626
Step 36/200     Reward: 447.84  Alpha:0.7       Epsilon:0.2216075419001291
Step 37/200     Reward: 490.12  Alpha:0.7       Epsilon:0.20971483596675833
Step 38/200     Reward: 546.9   Alpha:0.7       Epsilon:0.19813913785421933
Step 39/200     Reward: 543.98  Alpha:0.7       Epsilon:0.18686398532514437
Step 40/200     Reward: 504.1   Alpha:0.7       Epsilon:0.17587416608345108
Step 41/200     Reward: 492.96  Alpha:0.7       Epsilon:0.16515559435129612
Step 42/200     Reward: 492.48  Alpha:0.7       Epsilon:0.1546952023137098
Step 43/200     Reward: 516.5   Alpha:0.7       Epsilon:0.14448084433219988
Step 44/200     Reward: 490.68  Alpha:0.7       Epsilon:0.1345012121663145
Step 45/200     Reward: 496.88  Alpha:0.7       Epsilon:0.12474575971914248
Step 46/200     Reward: 530.84  Alpha:0.7       Epsilon:0.11520463605101905
Step 47/200     Reward: 501.3   Alpha:0.7       Epsilon:0.10586862559472299
Step 48/200     Reward: 445.38  Alpha:0.7       Epsilon:0.09672909466263513
Step 49/200     Reward: 574.84  Alpha:0.7       Epsilon:0.0877779434675845
Step 50/200     Reward: 465.64  Alpha:0.7       Epsilon:0.0790075629891599
Step 51/200     Reward: 534.02  Alpha:0.7       Epsilon:0.07041079610987233
Step 52/200     Reward: 555.08  Alpha:0.7       Epsilon:0.06198090252378974
Step 53/200     Reward: 496.44  Alpha:0.7       Epsilon:0.053711526986569
Step 54/200     Reward: 591.88  Alpha:0.7       Epsilon:0.05
Step 55/200     Reward: 524.7   Alpha:0.7       Epsilon:0.05
Step 56/200     Reward: 678.4   Alpha:0.7       Epsilon:0.05
Step 57/200     Reward: 612.36  Alpha:0.7       Epsilon:0.05
Step 58/200     Reward: 609.02  Alpha:0.7       Epsilon:0.05
Step 59/200     Reward: 615.2   Alpha:0.7       Epsilon:0.05
Step 60/200     Reward: 589.84  Alpha:0.7       Epsilon:0.05
Step 61/200     Reward: 601.96  Alpha:0.7       Epsilon:0.05
Step 62/200     Reward: 553.72  Alpha:0.7       Epsilon:0.05
Step 63/200     Reward: 452.12  Alpha:0.7       Epsilon:0.05
Step 64/200     Reward: 569.62  Alpha:0.7       Epsilon:0.05
Step 65/200     Reward: 487.12  Alpha:0.7       Epsilon:0.05
Step 66/200     Reward: 563.92  Alpha:0.7       Epsilon:0.05
Step 67/200     Reward: 585.52  Alpha:0.7       Epsilon:0.05
Step 68/200     Reward: 562.66  Alpha:0.7       Epsilon:0.05
Step 69/200     Reward: 672.8   Alpha:0.7       Epsilon:0.05
Step 70/200     Reward: 625.48  Alpha:0.7       Epsilon:0.05
```

```
Step 71/200     Reward: 595.12  Alpha:0.7      Epsilon:0.05
Step 72/200     Reward: 613.62  Alpha:0.7      Epsilon:0.05
Step 73/200     Reward: 539.84  Alpha:0.7      Epsilon:0.05
Step 74/200     Reward: 551.48  Alpha:0.7      Epsilon:0.05
Step 75/200     Reward: 570.18  Alpha:0.7      Epsilon:0.05
Step 76/200     Reward: 649.7   Alpha:0.7      Epsilon:0.05
Step 77/200     Reward: 630.26  Alpha:0.7      Epsilon:0.05
Step 78/200     Reward: 538.84  Alpha:0.7      Epsilon:0.05
Step 79/200     Reward: 581.82  Alpha:0.7      Epsilon:0.05
Step 80/200     Reward: 537.98  Alpha:0.7      Epsilon:0.05
Step 81/200     Reward: 627.6   Alpha:0.7      Epsilon:0.05
Step 82/200     Reward: 564.34  Alpha:0.7      Epsilon:0.05
Step 83/200     Reward: 612.84  Alpha:0.7      Epsilon:0.05
Step 84/200     Reward: 610.56  Alpha:0.7      Epsilon:0.05
Step 85/200     Reward: 600.8   Alpha:0.7      Epsilon:0.05
Step 86/200     Reward: 781.14  Alpha:0.7      Epsilon:0.05
Step 87/200     Reward: 570.38  Alpha:0.7      Epsilon:0.05
Step 88/200     Reward: 705.02  Alpha:0.7      Epsilon:0.05
Step 89/200     Reward: 548.86  Alpha:0.7      Epsilon:0.05
Step 90/200     Reward: 720.78  Alpha:0.7      Epsilon:0.05
Step 91/200     Reward: 695.08  Alpha:0.7      Epsilon:0.05
Step 92/200     Reward: 437.34  Alpha:0.7      Epsilon:0.05
Step 93/200     Reward: 671.02  Alpha:0.7      Epsilon:0.05
Step 94/200     Reward: 503.44  Alpha:0.7      Epsilon:0.05
Step 95/200     Reward: 699.56  Alpha:0.7      Epsilon:0.05
Step 96/200     Reward: 517.08  Alpha:0.7      Epsilon:0.05
Step 97/200     Reward: 670.22  Alpha:0.7      Epsilon:0.05
Step 98/200     Reward: 669.74  Alpha:0.7      Epsilon:0.05
Step 99/200     Reward: 520.18  Alpha:0.7      Epsilon:0.05
Step 100/200    Reward: 601.06  Alpha:0.7      Epsilon:0.05
Step 101/200    Reward: 683.72  Alpha:0.7      Epsilon:0.05
Step 102/200    Reward: 770.08  Alpha:0.7      Epsilon:0.05
Step 103/200    Reward: 563.32  Alpha:0.7      Epsilon:0.05
Step 104/200    Reward: 671.38  Alpha:0.7      Epsilon:0.05
Step 105/200    Reward: 593.48  Alpha:0.7      Epsilon:0.05
Step 106/200    Reward: 664.56  Alpha:0.7      Epsilon:0.05
Step 107/200    Reward: 665.06  Alpha:0.7      Epsilon:0.05
Step 108/200    Reward: 516.72  Alpha:0.7      Epsilon:0.05
Step 109/200    Reward: 650.12  Alpha:0.7      Epsilon:0.05
Step 110/200    Reward: 778.56  Alpha:0.7      Epsilon:0.05
Step 111/200    Reward: 774.54  Alpha:0.7      Epsilon:0.05
Step 112/200    Reward: 656.44  Alpha:0.7      Epsilon:0.05
Step 113/200    Reward: 723.28  Alpha:0.7      Epsilon:0.05
Step 114/200    Reward: 818.34  Alpha:0.7      Epsilon:0.05
Step 115/200    Reward: 751.66  Alpha:0.7      Epsilon:0.05
Step 116/200    Reward: 688.28  Alpha:0.7      Epsilon:0.05
Step 117/200    Reward: 630.74  Alpha:0.7      Epsilon:0.05
Step 118/200    Reward: 675.62  Alpha:0.7      Epsilon:0.05
Step 119/200    Reward: 604.66  Alpha:0.7      Epsilon:0.05
Step 120/200    Reward: 597.6   Alpha:0.7      Epsilon:0.05
Step 121/200    Reward: 654.94  Alpha:0.7      Epsilon:0.05
Step 122/200    Reward: 652.56  Alpha:0.7      Epsilon:0.05
Step 123/200    Reward: 609.92  Alpha:0.7      Epsilon:0.05
Step 124/200    Reward: 712.44  Alpha:0.7      Epsilon:0.05
Step 125/200    Reward: 619.74  Alpha:0.7      Epsilon:0.05
Step 126/200    Reward: 541.84  Alpha:0.7      Epsilon:0.05
Step 127/200    Reward: 599.0   Alpha:0.7      Epsilon:0.05
Step 128/200    Reward: 724.56  Alpha:0.7      Epsilon:0.05
Step 129/200    Reward: 831.22  Alpha:0.7      Epsilon:0.05
```

```
Step 130/200    Reward: 779.64  Alpha:0.7       Epsilon:0.05
Step 131/200    Reward: 749.06  Alpha:0.7       Epsilon:0.05
Step 132/200    Reward: 663.32  Alpha:0.7       Epsilon:0.05
Step 133/200    Reward: 640.0   Alpha:0.7       Epsilon:0.05
Step 134/200    Reward: 660.42  Alpha:0.7       Epsilon:0.05
Step 135/200    Reward: 632.4   Alpha:0.7       Epsilon:0.05
Step 136/200    Reward: 677.54  Alpha:0.7       Epsilon:0.05
Step 137/200    Reward: 784.18  Alpha:0.7       Epsilon:0.05
Step 138/200    Reward: 667.54  Alpha:0.7       Epsilon:0.05
Step 139/200    Reward: 637.5   Alpha:0.7       Epsilon:0.05
Step 140/200    Reward: 601.88  Alpha:0.7       Epsilon:0.05
Step 141/200    Reward: 660.74  Alpha:0.7       Epsilon:0.05
Step 142/200    Reward: 765.36  Alpha:0.7       Epsilon:0.05
Step 143/200    Reward: 727.58  Alpha:0.7       Epsilon:0.05
Step 144/200    Reward: 816.68  Alpha:0.7       Epsilon:0.05
Step 145/200    Reward: 669.94  Alpha:0.7       Epsilon:0.05
Step 146/200    Reward: 762.64  Alpha:0.7       Epsilon:0.05
Step 147/200    Reward: 826.06  Alpha:0.7       Epsilon:0.05
Step 148/200    Reward: 781.4   Alpha:0.7       Epsilon:0.05
Step 149/200    Reward: 789.72  Alpha:0.7       Epsilon:0.05
Step 150/200    Reward: 729.2   Alpha:0.7       Epsilon:0.05
Step 151/200    Reward: 715.16  Alpha:0.7       Epsilon:0.05
Step 152/200    Reward: 808.12  Alpha:0.7       Epsilon:0.05
Step 153/200    Reward: 707.06  Alpha:0.7       Epsilon:0.05
Step 154/200    Reward: 576.82  Alpha:0.7       Epsilon:0.05
Step 155/200    Reward: 732.62  Alpha:0.7       Epsilon:0.05
Step 156/200    Reward: 744.9   Alpha:0.7       Epsilon:0.05
Step 157/200    Reward: 678.28  Alpha:0.7       Epsilon:0.05
Step 158/200    Reward: 671.98  Alpha:0.7       Epsilon:0.05
Step 159/200    Reward: 827.08  Alpha:0.7       Epsilon:0.05
Step 160/200    Reward: 832.44  Alpha:0.7       Epsilon:0.05
Step 161/200    Reward: 808.0   Alpha:0.7       Epsilon:0.05
Step 162/200    Reward: 809.16  Alpha:0.7       Epsilon:0.05
Step 163/200    Reward: 820.42  Alpha:0.7       Epsilon:0.05
Step 164/200    Reward: 719.92  Alpha:0.7       Epsilon:0.05
Step 165/200    Reward: 627.62  Alpha:0.7       Epsilon:0.05
Step 166/200    Reward: 666.34  Alpha:0.7       Epsilon:0.05
Step 167/200    Reward: 745.48  Alpha:0.7       Epsilon:0.05
Step 168/200    Reward: 696.66  Alpha:0.7       Epsilon:0.05
Step 169/200    Reward: 774.46  Alpha:0.7       Epsilon:0.05
Step 170/200    Reward: 751.66  Alpha:0.7       Epsilon:0.05
Step 171/200    Reward: 642.96  Alpha:0.7       Epsilon:0.05
Step 172/200    Reward: 659.12  Alpha:0.7       Epsilon:0.05
Step 173/200    Reward: 759.24  Alpha:0.7       Epsilon:0.05
Step 174/200    Reward: 782.36  Alpha:0.7       Epsilon:0.05
Step 175/200    Reward: 660.74  Alpha:0.7       Epsilon:0.05
Step 176/200    Reward: 787.58  Alpha:0.7       Epsilon:0.05
Step 177/200    Reward: 698.96  Alpha:0.7       Epsilon:0.05
Step 178/200    Reward: 723.66  Alpha:0.7       Epsilon:0.05
Step 179/200    Reward: 890.38  Alpha:0.7       Epsilon:0.05
Step 180/200    Reward: 835.82  Alpha:0.7       Epsilon:0.05
Step 181/200    Reward: 653.46  Alpha:0.7       Epsilon:0.05
Step 182/200    Reward: 671.58  Alpha:0.7       Epsilon:0.05
Step 183/200    Reward: 719.86  Alpha:0.7       Epsilon:0.05
Step 184/200    Reward: 771.26  Alpha:0.7       Epsilon:0.05
Step 185/200    Reward: 676.96  Alpha:0.7       Epsilon:0.05
Step 186/200    Reward: 611.38  Alpha:0.7       Epsilon:0.05
Step 187/200    Reward: 834.3   Alpha:0.7       Epsilon:0.05
Step 188/200    Reward: 880.38  Alpha:0.7       Epsilon:0.05
```

```
Step 189/200    Reward: 678.92  Alpha:0.7       Epsilon:0.05
Step 190/200    Reward: 840.32  Alpha:0.7       Epsilon:0.05
Step 191/200    Reward: 887.26  Alpha:0.7       Epsilon:0.05
Step 192/200    Reward: 674.78  Alpha:0.7       Epsilon:0.05
Step 193/200    Reward: 744.6   Alpha:0.7       Epsilon:0.05
Step 194/200    Reward: 886.14  Alpha:0.7       Epsilon:0.05
Step 195/200    Reward: 818.44  Alpha:0.7       Epsilon:0.05
Step 196/200    Reward: 824.06  Alpha:0.7       Epsilon:0.05
Step 197/200    Reward: 727.66  Alpha:0.7       Epsilon:0.05
Step 198/200    Reward: 820.04  Alpha:0.7       Epsilon:0.05
```

In [647…

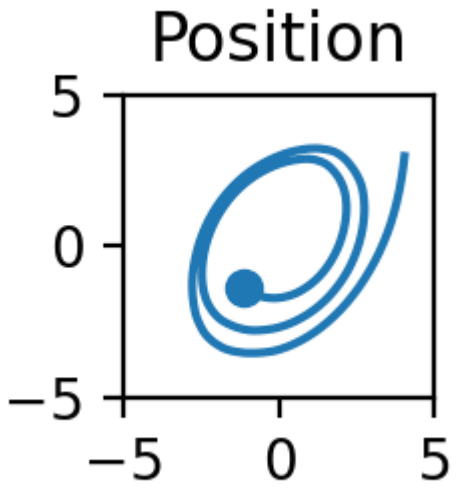
```
def plot_policy(Q):

    done = False
    observation, other = orbitEnv.reset()

    current_state = dobs_q.obs2state(observation)
    current_idx = dobs_q.obs2idx(observation)


    time = -1
    positions = [[], []]
    velocities = [[], []]
    vel_norms = []
    pos_norms = []
    delta_v = []
    actions = []
    observations = []
    total_reward_here = 0
    while not done:

        time += 1
        action = get_action(current_idx, Q, 0, orbitEnv.action_space.n) # take full g
        observation, reward, done, info, other = orbitEnv.step(action)
        total_reward_here += reward

        current_idx = dobs_q.obs2idx(observation)

        observations.append(observation)

        state = orbitEnv.get_state()
        pos_norms.append(np.linalg.norm(state[0]))
        positions[0].append(state[0][0])
        positions[1].append(state[0][1])

        vel_norms.append(np.linalg.norm(state[1]))
        velocities[0].append(state[1][0])
        velocities[1].append(state[1][1])

        delta_v.append(state[2])
        actions.append(action)

    print(total_reward_here)
    print(np.linalg.norm([positions[0][0], positions[1][0]]))

    figure(figsize=(1, 1), dpi=200)
    plt.scatter(positions[0][0], positions[1][0])
    plt.plot(positions[0], positions[1])
    plt.xlim(-5.0, 5.0)
    plt.ylim(-5.0, 5.0)
    plt.title("Position")
    plt.show()

    plt.plot(np.arange(0,time+1), delta_v)
    plt.title("Delta-V over time")
    plt.show()

    plt.plot(np.arange(0, time+1), vel_norms)
    plt.title("Velocity Magnitude")
```

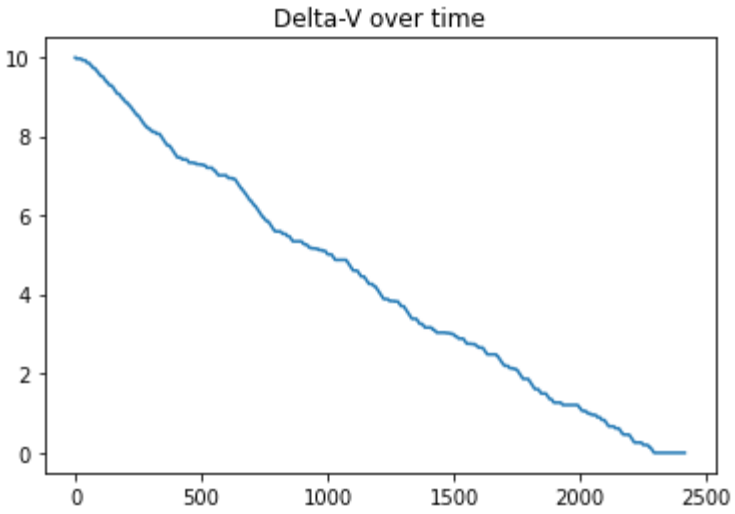
```
    plt.show()

    plt.plot(np.arange(0, time+1), pos_norms)
    plt.title("Position Magnitude")
    plt.show()


    plt.plot(np.arange(0, time+1), [ob[1] for ob in observations])
    plt.show()

    plt.plot(np.arange(0, time+1), [ob[2] for ob in observations])
    plt.show()

    #plt.plot(velocities[0] velocities[1])


def plot_result():
    plt.plot(np.arange(1, len(qlearning_result)), qlearning_result[1:])
    plt.show()
```

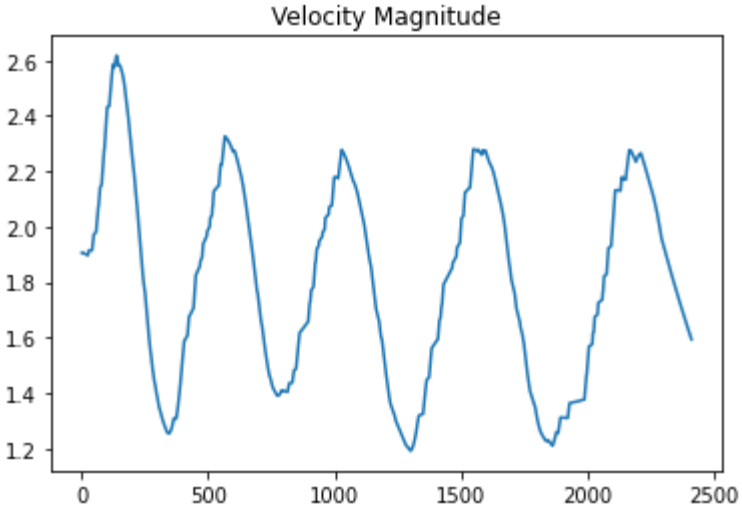
In [630…

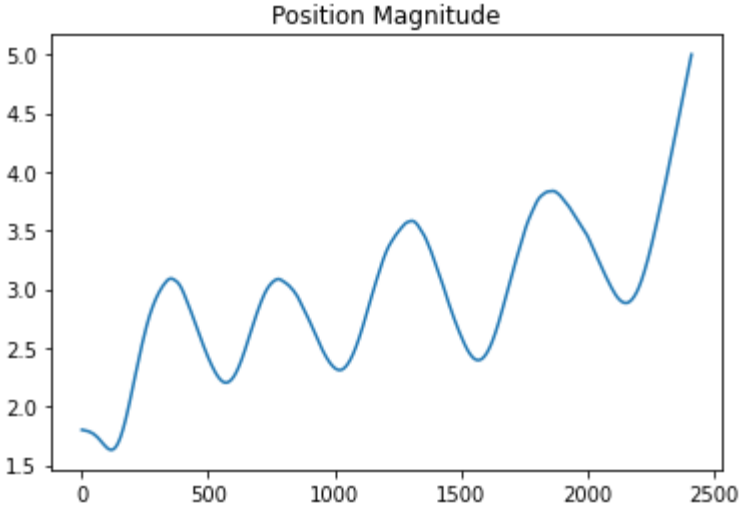
```
plot_policy(Q_qlearning)
```

```
1757.0
1.7998610260529673
```



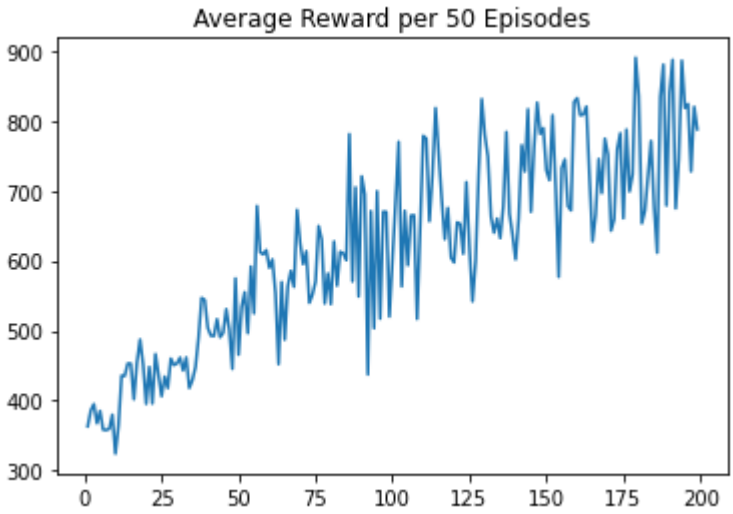

Delta-V over time
10
8
6
4
2
0
0
500
1000
1500
2000
2500


Velocity Magnitude
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
0
500
1000
1500
2000
2500


Position Magnitude
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
0
500
1000
1500
2000
2500

In [632…

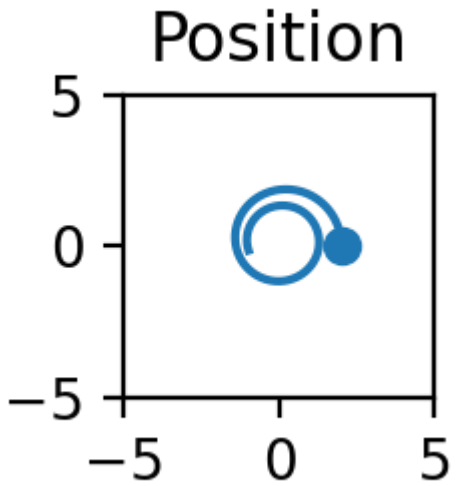
```
def plot_result():
    plt.plot(np.arange(1, len(qlearning_result)), qlearning_result[1:])
    plt.title("Average Reward per 50 Episodes")
    plt.show()

plot_result()
```

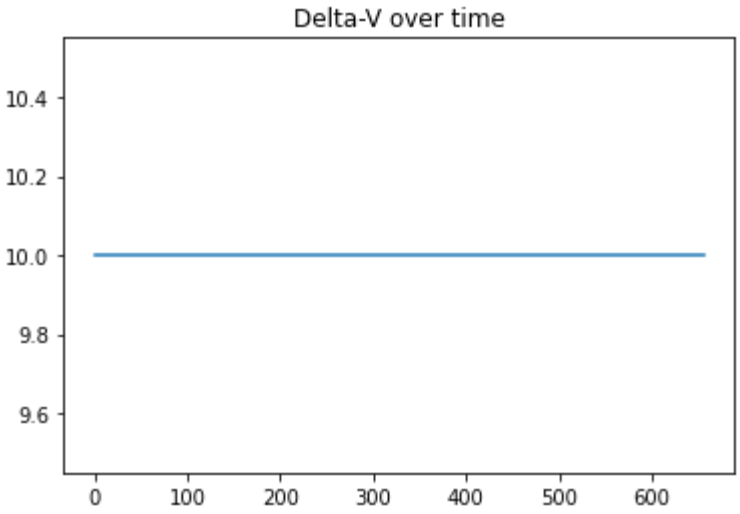


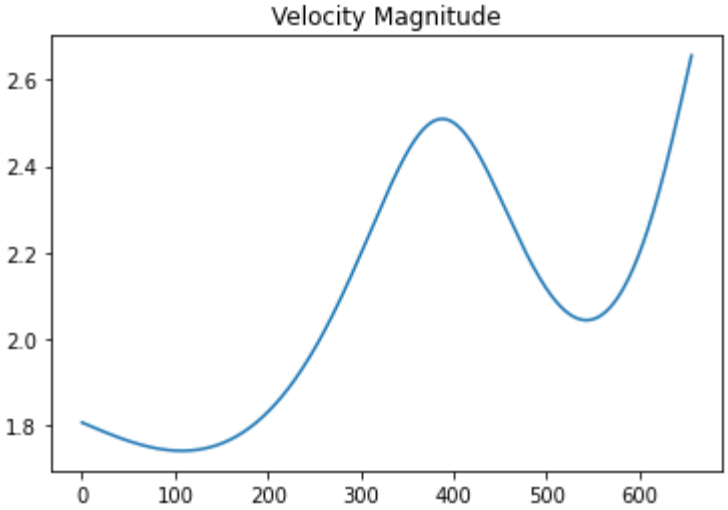
In [639…

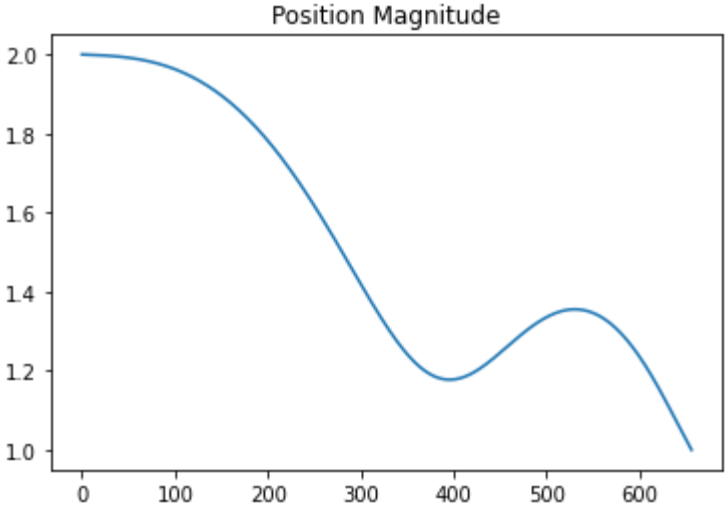
```
# plot baseline
orbitEnv = OrbitSimEnv(random_angle=False, random_distance=False)
plot_policy(Q_qlearning)
```

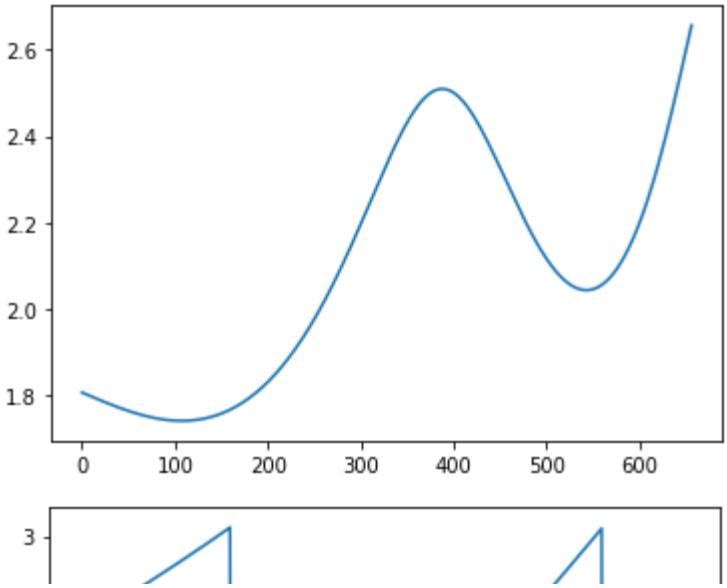
```
-1.0
1.9999181158834742
```




Delta-V over time
10.4
10.2
10.0
9.8
9.6
0
100
200
300
400
500
600


Velocity Magnitude
2.6
2.4
2.2
2.0
1.8
0
100
200
300
400
500
600


Position Magnitude
2.0
1.8
1.6
1.4
1.2
1.0
0
100
200
300
400
500
600



In [640…

```
First_pass_q_learning = np.copy(Q_qlearning)
First_pass_q_result = np.copy(qlearning_result)
```

In [642…

```
bins_distance = np.linspace(0,5,50) # state for each 0.1 distance
bins_theta = np.linspace(-np.pi,np.pi, 90) # state for every 4 degrees
bins_vel_mag = np.linspace(1,4,60) # 0.05
bins_vel_angle = np.linspace(-np.pi,np.pi, 90) # +/- 45 degrees from tangent to accel
bins_delta_v = np.linspace(0, 10, 50) # 0.1

bins_vel_x = np.linspace(-3,3,30)
bins_vel_y = np.linspace(-3,3,30)

dobs_q = DiscretObs([bins_distance,bins_vel_mag,bins_vel_angle]) #bins_delta_v, bins_

orbitEnv = OrbitSimEnv(random_angle=True, random_distance=True, random_distance_min_r

Q_qlearning = set_Q_qlearning(dobs_q.get_state_num_total(), 2)
q_learning_algo(dobs_q, orbitEnv, gamma=0.98, episode_count=50000)
ql_Q_98_50000 = np.copy(Q_qlearning)
ql_result_98_50000 = np.copy(qlearning_result)
```

```
Running Q-Learning Algorithm with gamma=0.98, alpha=None, epsilon=None, ep_count=5000
0, report_step=50
Step 0/1000     Reward: 0.0     Alpha:1.0       Epsilon:1.0
Step 1/1000     Reward: 158.36  Alpha:1.0       Epsilon:1.0
Step 2/1000     Reward: 212.24  Alpha:1.0       Epsilon:1.0
Step 3/1000     Reward: 126.64  Alpha:1.0       Epsilon:1.0
Step 4/1000     Reward: 175.5   Alpha:0.9978339382434924        Epsilon:1.0
Step 5/1000     Reward: 220.2   Alpha:0.9013562741829431        Epsilon:1.0
Step 6/1000     Reward: 179.5   Alpha:0.8224635000701379        Epsilon:0.99855475912
5819
Step 7/1000     Reward: 162.78  Alpha:0.7557228791981572        Epsilon:0.93181413825
38384
Step 8/1000     Reward: 196.22  Alpha:0.7       Epsilon:0.8739768820994801
Step 9/1000     Reward: 139.52  Alpha:0.7       Epsilon:0.8229447128417019
Step 34/1000    Reward: 257.6   Alpha:0.7       Epsilon:0.24641694110709345
Step 35/1000    Reward: 202.42  Alpha:0.7       Epsilon:0.23383510863621626
Step 36/1000    Reward: 152.76  Alpha:0.7       Epsilon:0.2216075419001291
```

```
Step 37/1000    Reward: 166.06  Alpha:0.7       Epsilon:0.20971483596675833
Step 38/1000    Reward: 195.28  Alpha:0.7       Epsilon:0.19813913785421933
Step 39/1000    Reward: 307.32  Alpha:0.7       Epsilon:0.18686398532514437
Step 40/1000    Reward: 232.8   Alpha:0.7       Epsilon:0.17587416608345108
Step 41/1000    Reward: 187.0   Alpha:0.7       Epsilon:0.16515559435129612
Step 42/1000    Reward: 242.46  Alpha:0.7       Epsilon:0.1546952023137098
Step 43/1000    Reward: 258.58  Alpha:0.7       Epsilon:0.14448084433219988
Step 44/1000    Reward: 194.12  Alpha:0.7       Epsilon:0.1345012121663145
Step 45/1000    Reward: 227.78  Alpha:0.7       Epsilon:0.12474575971914248
Step 46/1000    Reward: 270.18  Alpha:0.7       Epsilon:0.11520463605101905
Step 47/1000    Reward: 158.46  Alpha:0.7       Epsilon:0.10586862559472299
Step 48/1000    Reward: 166.42  Alpha:0.7       Epsilon:0.09672909466263513
Step 49/1000    Reward: 200.2   Alpha:0.7       Epsilon:0.0877779434675845
Step 50/1000    Reward: 352.06  Alpha:0.7       Epsilon:0.0790075629891599
Step 51/1000    Reward: 233.96  Alpha:0.7       Epsilon:0.07041079610987233
Step 52/1000    Reward: 394.46  Alpha:0.7       Epsilon:0.06198090252378974
Step 53/1000    Reward: 253.98  Alpha:0.7       Epsilon:0.053711526986569
Step 54/1000    Reward: 169.36  Alpha:0.7       Epsilon:0.05
Step 55/1000    Reward: 384.36  Alpha:0.7       Epsilon:0.05
Step 56/1000    Reward: 324.3   Alpha:0.7       Epsilon:0.05
Step 57/1000    Reward: 235.7   Alpha:0.7       Epsilon:0.05
Step 58/1000    Reward: 293.2   Alpha:0.7       Epsilon:0.05
Step 59/1000    Reward: 333.68  Alpha:0.7       Epsilon:0.05
Step 60/1000    Reward: 247.12  Alpha:0.7       Epsilon:0.05
Step 61/1000    Reward: 295.08  Alpha:0.7       Epsilon:0.05
Step 62/1000    Reward: 319.7   Alpha:0.7       Epsilon:0.05
Step 63/1000    Reward: 298.78  Alpha:0.7       Epsilon:0.05
Step 64/1000    Reward: 206.9   Alpha:0.7       Epsilon:0.05
Step 65/1000    Reward: 298.06  Alpha:0.7       Epsilon:0.05
Step 66/1000    Reward: 168.04  Alpha:0.7       Epsilon:0.05
Step 67/1000    Reward: 277.52  Alpha:0.7       Epsilon:0.05
Step 68/1000    Reward: 177.18  Alpha:0.7       Epsilon:0.05
Step 69/1000    Reward: 216.4   Alpha:0.7       Epsilon:0.05
Step 70/1000    Reward: 359.78  Alpha:0.7       Epsilon:0.05
Step 71/1000    Reward: 258.36  Alpha:0.7       Epsilon:0.05
Step 72/1000    Reward: 328.12  Alpha:0.7       Epsilon:0.05
Step 73/1000    Reward: 175.48  Alpha:0.7       Epsilon:0.05
Step 74/1000    Reward: 312.2   Alpha:0.7       Epsilon:0.05
Step 75/1000    Reward: 230.28  Alpha:0.7       Epsilon:0.05
Step 76/1000    Reward: 378.56  Alpha:0.7       Epsilon:0.05
Step 77/1000    Reward: 397.1   Alpha:0.7       Epsilon:0.05
Step 78/1000    Reward: 209.76  Alpha:0.7       Epsilon:0.05
Step 79/1000    Reward: 274.0   Alpha:0.7       Epsilon:0.05
Step 80/1000    Reward: 379.56  Alpha:0.7       Epsilon:0.05
Step 81/1000    Reward: 220.1   Alpha:0.7       Epsilon:0.05
Step 82/1000    Reward: 268.52  Alpha:0.7       Epsilon:0.05
Step 83/1000    Reward: 281.26  Alpha:0.7       Epsilon:0.05
Step 84/1000    Reward: 359.4   Alpha:0.7       Epsilon:0.05
Step 85/1000    Reward: 309.04  Alpha:0.7       Epsilon:0.05
Step 86/1000    Reward: 293.24  Alpha:0.7       Epsilon:0.05
Step 87/1000    Reward: 365.66  Alpha:0.7       Epsilon:0.05
Step 88/1000    Reward: 279.98  Alpha:0.7       Epsilon:0.05
Step 89/1000    Reward: 354.56  Alpha:0.7       Epsilon:0.05
Step 90/1000    Reward: 182.4   Alpha:0.7       Epsilon:0.05
Step 91/1000    Reward: 286.26  Alpha:0.7       Epsilon:0.05
Step 92/1000    Reward: 303.74  Alpha:0.7       Epsilon:0.05
Step 93/1000    Reward: 373.64  Alpha:0.7       Epsilon:0.05
Step 94/1000    Reward: 189.8   Alpha:0.7       Epsilon:0.05
Step 95/1000    Reward: 254.36  Alpha:0.7       Epsilon:0.05
```

```
Step 96/1000    Reward: 291.02  Alpha:0.7       Epsilon:0.05
Step 97/1000    Reward: 230.84  Alpha:0.7       Epsilon:0.05
Step 98/1000    Reward: 283.78  Alpha:0.7       Epsilon:0.05
Step 99/1000    Reward: 344.58  Alpha:0.7       Epsilon:0.05
Step 100/1000   Reward: 117.26  Alpha:0.7       Epsilon:0.05
Step 101/1000   Reward: 260.7   Alpha:0.7       Epsilon:0.05
Step 102/1000   Reward: 367.86  Alpha:0.7       Epsilon:0.05
Step 103/1000   Reward: 368.48  Alpha:0.7       Epsilon:0.05
Step 104/1000   Reward: 191.14  Alpha:0.7       Epsilon:0.05
Step 105/1000   Reward: 359.44  Alpha:0.7       Epsilon:0.05
Step 106/1000   Reward: 368.84  Alpha:0.7       Epsilon:0.05
Step 107/1000   Reward: 261.5   Alpha:0.7       Epsilon:0.05
Step 108/1000   Reward: 430.84  Alpha:0.7       Epsilon:0.05
Step 109/1000   Reward: 412.7   Alpha:0.7       Epsilon:0.05
Step 110/1000   Reward: 334.8   Alpha:0.7       Epsilon:0.05
Step 111/1000   Reward: 293.54  Alpha:0.7       Epsilon:0.05
Step 112/1000   Reward: 209.2   Alpha:0.7       Epsilon:0.05
Step 113/1000   Reward: 190.92  Alpha:0.7       Epsilon:0.05
Step 114/1000   Reward: 129.1   Alpha:0.7       Epsilon:0.05
Step 115/1000   Reward: 334.2   Alpha:0.7       Epsilon:0.05
Step 116/1000   Reward: 352.56  Alpha:0.7       Epsilon:0.05
Step 117/1000   Reward: 301.02  Alpha:0.7       Epsilon:0.05
Step 118/1000   Reward: 327.96  Alpha:0.7       Epsilon:0.05
Step 119/1000   Reward: 321.64  Alpha:0.7       Epsilon:0.05
Step 120/1000   Reward: 333.54  Alpha:0.7       Epsilon:0.05
Step 121/1000   Reward: 271.44  Alpha:0.7       Epsilon:0.05
Step 122/1000   Reward: 315.12  Alpha:0.7       Epsilon:0.05
Step 123/1000   Reward: 201.36  Alpha:0.7       Epsilon:0.05
Step 124/1000   Reward: 445.74  Alpha:0.7       Epsilon:0.05
Step 125/1000   Reward: 227.22  Alpha:0.7       Epsilon:0.05
Step 126/1000   Reward: 266.36  Alpha:0.7       Epsilon:0.05
Step 127/1000   Reward: 349.1   Alpha:0.7       Epsilon:0.05
Step 128/1000   Reward: 245.44  Alpha:0.7       Epsilon:0.05
Step 129/1000   Reward: 204.92  Alpha:0.7       Epsilon:0.05
Step 130/1000   Reward: 321.98  Alpha:0.7       Epsilon:0.05
Step 131/1000   Reward: 375.18  Alpha:0.7       Epsilon:0.05
Step 132/1000   Reward: 332.58  Alpha:0.7       Epsilon:0.05
Step 133/1000   Reward: 216.24  Alpha:0.7       Epsilon:0.05
Step 134/1000   Reward: 240.38  Alpha:0.7       Epsilon:0.05
Step 135/1000   Reward: 397.12  Alpha:0.7       Epsilon:0.05
Step 136/1000   Reward: 180.92  Alpha:0.7       Epsilon:0.05
Step 137/1000   Reward: 185.16  Alpha:0.7       Epsilon:0.05
Step 138/1000   Reward: 364.48  Alpha:0.7       Epsilon:0.05
Step 139/1000   Reward: 279.8   Alpha:0.7       Epsilon:0.05
Step 140/1000   Reward: 184.38  Alpha:0.7       Epsilon:0.05
Step 141/1000   Reward: 365.06  Alpha:0.7       Epsilon:0.05
Step 142/1000   Reward: 212.78  Alpha:0.7       Epsilon:0.05
Step 143/1000   Reward: 314.78  Alpha:0.7       Epsilon:0.05
Step 144/1000   Reward: 251.16  Alpha:0.7       Epsilon:0.05
Step 145/1000   Reward: 280.8   Alpha:0.7       Epsilon:0.05
Step 146/1000   Reward: 416.72  Alpha:0.7       Epsilon:0.05
Step 147/1000   Reward: 288.52  Alpha:0.7       Epsilon:0.05
Step 148/1000   Reward: 352.9   Alpha:0.7       Epsilon:0.05
Step 149/1000   Reward: 275.02  Alpha:0.7       Epsilon:0.05
Step 150/1000   Reward: 275.52  Alpha:0.7       Epsilon:0.05
Step 151/1000   Reward: 352.78  Alpha:0.7       Epsilon:0.05
Step 152/1000   Reward: 231.14  Alpha:0.7       Epsilon:0.05
Step 153/1000   Reward: 432.8   Alpha:0.7       Epsilon:0.05
Step 154/1000   Reward: 305.86  Alpha:0.7       Epsilon:0.05
```

```
Step 155/1000   Reward: 268.44  Alpha:0.7       Epsilon:0.05
Step 156/1000   Reward: 418.3   Alpha:0.7       Epsilon:0.05
Step 157/1000   Reward: 152.64  Alpha:0.7       Epsilon:0.05
Step 158/1000   Reward: 317.46  Alpha:0.7       Epsilon:0.05
Step 159/1000   Reward: 181.46  Alpha:0.7       Epsilon:0.05
Step 160/1000   Reward: 415.46  Alpha:0.7       Epsilon:0.05
Step 161/1000   Reward: 339.9   Alpha:0.7       Epsilon:0.05
Step 162/1000   Reward: 295.16  Alpha:0.7       Epsilon:0.05
Step 163/1000   Reward: 241.7   Alpha:0.7       Epsilon:0.05
Step 164/1000   Reward: 302.54  Alpha:0.7       Epsilon:0.05
Step 165/1000   Reward: 296.16  Alpha:0.7       Epsilon:0.05
Step 166/1000   Reward: 362.32  Alpha:0.7       Epsilon:0.05
Step 167/1000   Reward: 96.82   Alpha:0.7       Epsilon:0.05
Step 168/1000   Reward: 285.32  Alpha:0.7       Epsilon:0.05
Step 169/1000   Reward: 275.76  Alpha:0.7       Epsilon:0.05
Step 170/1000   Reward: 248.28  Alpha:0.7       Epsilon:0.05
Step 171/1000   Reward: 255.46  Alpha:0.7       Epsilon:0.05
Step 172/1000   Reward: 309.76  Alpha:0.7       Epsilon:0.05
Step 173/1000   Reward: 196.62  Alpha:0.7       Epsilon:0.05
Step 174/1000   Reward: 330.44  Alpha:0.7       Epsilon:0.05
Step 175/1000   Reward: 144.52  Alpha:0.7       Epsilon:0.05
Step 176/1000   Reward: 407.28  Alpha:0.7       Epsilon:0.05
Step 177/1000   Reward: 290.3   Alpha:0.7       Epsilon:0.05
Step 178/1000   Reward: 293.96  Alpha:0.7       Epsilon:0.05
Step 179/1000   Reward: 203.22  Alpha:0.7       Epsilon:0.05
Step 180/1000   Reward: 262.46  Alpha:0.7       Epsilon:0.05
Step 181/1000   Reward: 294.94  Alpha:0.7       Epsilon:0.05
Step 182/1000   Reward: 270.58  Alpha:0.7       Epsilon:0.05
Step 183/1000   Reward: 322.78  Alpha:0.7       Epsilon:0.05
Step 184/1000   Reward: 288.08  Alpha:0.7       Epsilon:0.05
Step 185/1000   Reward: 323.4   Alpha:0.7       Epsilon:0.05
Step 186/1000   Reward: 310.02  Alpha:0.7       Epsilon:0.05
Step 187/1000   Reward: 329.92  Alpha:0.7       Epsilon:0.05
Step 188/1000   Reward: 350.86  Alpha:0.7       Epsilon:0.05
Step 189/1000   Reward: 302.02  Alpha:0.7       Epsilon:0.05
Step 190/1000   Reward: 321.5   Alpha:0.7       Epsilon:0.05
Step 191/1000   Reward: 310.1   Alpha:0.7       Epsilon:0.05
Step 192/1000   Reward: 266.84  Alpha:0.7       Epsilon:0.05
Step 193/1000   Reward: 327.54  Alpha:0.7       Epsilon:0.05
Step 194/1000   Reward: 367.2   Alpha:0.7       Epsilon:0.05
Step 195/1000   Reward: 248.0   Alpha:0.7       Epsilon:0.05
Step 196/1000   Reward: 227.86  Alpha:0.7       Epsilon:0.05
Step 197/1000   Reward: 248.44  Alpha:0.7       Epsilon:0.05
Step 198/1000   Reward: 386.7   Alpha:0.7       Epsilon:0.05
Step 199/1000   Reward: 278.24  Alpha:0.7       Epsilon:0.05
Step 200/1000   Reward: 304.8   Alpha:0.7       Epsilon:0.05
Step 201/1000   Reward: 242.48  Alpha:0.7       Epsilon:0.05
Step 202/1000   Reward: 314.56  Alpha:0.7       Epsilon:0.05
Step 203/1000   Reward: 340.68  Alpha:0.7       Epsilon:0.05
Step 204/1000   Reward: 301.66  Alpha:0.7       Epsilon:0.05
Step 205/1000   Reward: 277.24  Alpha:0.7       Epsilon:0.05
Step 206/1000   Reward: 319.12  Alpha:0.7       Epsilon:0.05
Step 207/1000   Reward: 288.52  Alpha:0.7       Epsilon:0.05
Step 208/1000   Reward: 253.22  Alpha:0.7       Epsilon:0.05
Step 209/1000   Reward: 398.76  Alpha:0.7       Epsilon:0.05
Step 210/1000   Reward: 311.16  Alpha:0.7       Epsilon:0.05
Step 211/1000   Reward: 273.34  Alpha:0.7       Epsilon:0.05
Step 212/1000   Reward: 286.04  Alpha:0.7       Epsilon:0.05
Step 213/1000   Reward: 379.36  Alpha:0.7       Epsilon:0.05
```

```
Step 214/1000   Reward: 297.16  Alpha:0.7       Epsilon:0.05
Step 215/1000   Reward: 302.32  Alpha:0.7       Epsilon:0.05
Step 216/1000   Reward: 272.32  Alpha:0.7       Epsilon:0.05
Step 217/1000   Reward: 318.5   Alpha:0.7       Epsilon:0.05
Step 218/1000   Reward: 400.42  Alpha:0.7       Epsilon:0.05
Step 219/1000   Reward: 360.56  Alpha:0.7       Epsilon:0.05
Step 220/1000   Reward: 411.88  Alpha:0.7       Epsilon:0.05
Step 221/1000   Reward: 438.98  Alpha:0.7       Epsilon:0.05
Step 222/1000   Reward: 497.04  Alpha:0.7       Epsilon:0.05
Step 223/1000   Reward: 176.36  Alpha:0.7       Epsilon:0.05
Step 224/1000   Reward: 420.84  Alpha:0.7       Epsilon:0.05
Step 225/1000   Reward: 313.88  Alpha:0.7       Epsilon:0.05
Step 226/1000   Reward: 405.84  Alpha:0.7       Epsilon:0.05
Step 227/1000   Reward: 356.9   Alpha:0.7       Epsilon:0.05
Step 228/1000   Reward: 312.04  Alpha:0.7       Epsilon:0.05
Step 229/1000   Reward: 340.84  Alpha:0.7       Epsilon:0.05
Step 230/1000   Reward: 305.2   Alpha:0.7       Epsilon:0.05
Step 231/1000   Reward: 305.82  Alpha:0.7       Epsilon:0.05
Step 232/1000   Reward: 484.96  Alpha:0.7       Epsilon:0.05
Step 233/1000   Reward: 372.88  Alpha:0.7       Epsilon:0.05
Step 234/1000   Reward: 316.9   Alpha:0.7       Epsilon:0.05
Step 235/1000   Reward: 414.58  Alpha:0.7       Epsilon:0.05
Step 236/1000   Reward: 440.14  Alpha:0.7       Epsilon:0.05
Step 237/1000   Reward: 548.62  Alpha:0.7       Epsilon:0.05
Step 238/1000   Reward: 381.16  Alpha:0.7       Epsilon:0.05
Step 239/1000   Reward: 186.2   Alpha:0.7       Epsilon:0.05
Step 240/1000   Reward: 303.72  Alpha:0.7       Epsilon:0.05
Step 241/1000   Reward: 424.42  Alpha:0.7       Epsilon:0.05
Step 242/1000   Reward: 314.72  Alpha:0.7       Epsilon:0.05
Step 243/1000   Reward: 481.06  Alpha:0.7       Epsilon:0.05
Step 244/1000   Reward: 345.02  Alpha:0.7       Epsilon:0.05
Step 245/1000   Reward: 313.06  Alpha:0.7       Epsilon:0.05
Step 246/1000   Reward: 344.54  Alpha:0.7       Epsilon:0.05
Step 247/1000   Reward: 478.7   Alpha:0.7       Epsilon:0.05
Step 248/1000   Reward: 313.58  Alpha:0.7       Epsilon:0.05
Step 249/1000   Reward: 266.7   Alpha:0.7       Epsilon:0.05
Step 250/1000   Reward: 452.54  Alpha:0.7       Epsilon:0.05
Step 251/1000   Reward: 428.92  Alpha:0.7       Epsilon:0.05
Step 252/1000   Reward: 219.44  Alpha:0.7       Epsilon:0.05
Step 253/1000   Reward: 393.5   Alpha:0.7       Epsilon:0.05
Step 254/1000   Reward: 235.48  Alpha:0.7       Epsilon:0.05
Step 255/1000   Reward: 354.56  Alpha:0.7       Epsilon:0.05
Step 256/1000   Reward: 472.28  Alpha:0.7       Epsilon:0.05
Step 257/1000   Reward: 296.4   Alpha:0.7       Epsilon:0.05
Step 258/1000   Reward: 449.74  Alpha:0.7       Epsilon:0.05
Step 259/1000   Reward: 395.44  Alpha:0.7       Epsilon:0.05
Step 260/1000   Reward: 330.24  Alpha:0.7       Epsilon:0.05
Step 261/1000   Reward: 459.74  Alpha:0.7       Epsilon:0.05
Step 262/1000   Reward: 318.5   Alpha:0.7       Epsilon:0.05
Step 263/1000   Reward: 270.54  Alpha:0.7       Epsilon:0.05
Step 264/1000   Reward: 429.24  Alpha:0.7       Epsilon:0.05
Step 265/1000   Reward: 281.56  Alpha:0.7       Epsilon:0.05
Step 266/1000   Reward: 297.8   Alpha:0.7       Epsilon:0.05
Step 267/1000   Reward: 289.6   Alpha:0.7       Epsilon:0.05
Step 268/1000   Reward: 510.62  Alpha:0.7       Epsilon:0.05
Step 269/1000   Reward: 325.22  Alpha:0.7       Epsilon:0.05
Step 270/1000   Reward: 345.56  Alpha:0.7       Epsilon:0.05
Step 271/1000   Reward: 341.28  Alpha:0.7       Epsilon:0.05
Step 272/1000   Reward: 338.08  Alpha:0.7       Epsilon:0.05
```

```
Step 273/1000   Reward: 309.3   Alpha:0.7       Epsilon:0.05
Step 274/1000   Reward: 271.42  Alpha:0.7       Epsilon:0.05
Step 275/1000   Reward: 429.68  Alpha:0.7       Epsilon:0.05
Step 276/1000   Reward: 413.12  Alpha:0.7       Epsilon:0.05
Step 277/1000   Reward: 398.78  Alpha:0.7       Epsilon:0.05
Step 278/1000   Reward: 440.78  Alpha:0.7       Epsilon:0.05
Step 279/1000   Reward: 429.06  Alpha:0.7       Epsilon:0.05
Step 280/1000   Reward: 353.76  Alpha:0.7       Epsilon:0.05
Step 281/1000   Reward: 405.08  Alpha:0.7       Epsilon:0.05
Step 282/1000   Reward: 270.52  Alpha:0.7       Epsilon:0.05
Step 283/1000   Reward: 442.88  Alpha:0.7       Epsilon:0.05
Step 284/1000   Reward: 390.48  Alpha:0.7       Epsilon:0.05
Step 285/1000   Reward: 430.56  Alpha:0.7       Epsilon:0.05
Step 286/1000   Reward: 414.9   Alpha:0.7       Epsilon:0.05
Step 287/1000   Reward: 262.2   Alpha:0.7       Epsilon:0.05
Step 288/1000   Reward: 356.02  Alpha:0.7       Epsilon:0.05
Step 289/1000   Reward: 606.36  Alpha:0.7       Epsilon:0.05
Step 290/1000   Reward: 460.48  Alpha:0.7       Epsilon:0.05
Step 291/1000   Reward: 510.66  Alpha:0.7       Epsilon:0.05
Step 292/1000   Reward: 396.66  Alpha:0.7       Epsilon:0.05
Step 293/1000   Reward: 343.38  Alpha:0.7       Epsilon:0.05
Step 294/1000   Reward: 370.5   Alpha:0.7       Epsilon:0.05
Step 295/1000   Reward: 493.74  Alpha:0.7       Epsilon:0.05
Step 296/1000   Reward: 337.3   Alpha:0.7       Epsilon:0.05
Step 297/1000   Reward: 425.78  Alpha:0.7       Epsilon:0.05
Step 298/1000   Reward: 425.92  Alpha:0.7       Epsilon:0.05
Step 299/1000   Reward: 380.4   Alpha:0.7       Epsilon:0.05
Step 300/1000   Reward: 151.38  Alpha:0.7       Epsilon:0.05
Step 301/1000   Reward: 417.92  Alpha:0.7       Epsilon:0.05
Step 302/1000   Reward: 443.06  Alpha:0.7       Epsilon:0.05
Step 303/1000   Reward: 394.64  Alpha:0.7       Epsilon:0.05
Step 304/1000   Reward: 381.22  Alpha:0.7       Epsilon:0.05
Step 305/1000   Reward: 358.8   Alpha:0.7       Epsilon:0.05
Step 306/1000   Reward: 552.96  Alpha:0.7       Epsilon:0.05
Step 307/1000   Reward: 391.6   Alpha:0.7       Epsilon:0.05
Step 308/1000   Reward: 472.4   Alpha:0.7       Epsilon:0.05
Step 309/1000   Reward: 470.24  Alpha:0.7       Epsilon:0.05
Step 310/1000   Reward: 361.54  Alpha:0.7       Epsilon:0.05
Step 311/1000   Reward: 290.92  Alpha:0.7       Epsilon:0.05
Step 312/1000   Reward: 212.14  Alpha:0.7       Epsilon:0.05
Step 313/1000   Reward: 404.54  Alpha:0.7       Epsilon:0.05
Step 314/1000   Reward: 349.06  Alpha:0.7       Epsilon:0.05
Step 315/1000   Reward: 301.84  Alpha:0.7       Epsilon:0.05
Step 316/1000   Reward: 392.1   Alpha:0.7       Epsilon:0.05
Step 317/1000   Reward: 587.3   Alpha:0.7       Epsilon:0.05
Step 318/1000   Reward: 518.5   Alpha:0.7       Epsilon:0.05
Step 319/1000   Reward: 429.22  Alpha:0.7       Epsilon:0.05
Step 320/1000   Reward: 421.2   Alpha:0.7       Epsilon:0.05
Step 321/1000   Reward: 280.54  Alpha:0.7       Epsilon:0.05
Step 322/1000   Reward: 363.12  Alpha:0.7       Epsilon:0.05
Step 323/1000   Reward: 531.56  Alpha:0.7       Epsilon:0.05
Step 324/1000   Reward: 245.54  Alpha:0.7       Epsilon:0.05
Step 325/1000   Reward: 302.96  Alpha:0.7       Epsilon:0.05
Step 326/1000   Reward: 383.78  Alpha:0.7       Epsilon:0.05
Step 327/1000   Reward: 362.5   Alpha:0.7       Epsilon:0.05
Step 328/1000   Reward: 353.62  Alpha:0.7       Epsilon:0.05
Step 329/1000   Reward: 470.46  Alpha:0.7       Epsilon:0.05
Step 330/1000   Reward: 358.84  Alpha:0.7       Epsilon:0.05
Step 331/1000   Reward: 419.88  Alpha:0.7       Epsilon:0.05
```

```
Step 332/1000   Reward: 332.72  Alpha:0.7       Epsilon:0.05
Step 333/1000   Reward: 281.66  Alpha:0.7       Epsilon:0.05
Step 334/1000   Reward: 417.3   Alpha:0.7       Epsilon:0.05
Step 335/1000   Reward: 429.04  Alpha:0.7       Epsilon:0.05
Step 336/1000   Reward: 498.12  Alpha:0.7       Epsilon:0.05
Step 337/1000   Reward: 490.52  Alpha:0.7       Epsilon:0.05
Step 338/1000   Reward: 475.9   Alpha:0.7       Epsilon:0.05
Step 339/1000   Reward: 247.94  Alpha:0.7       Epsilon:0.05
Step 340/1000   Reward: 323.0   Alpha:0.7       Epsilon:0.05
Step 341/1000   Reward: 292.24  Alpha:0.7       Epsilon:0.05
Step 342/1000   Reward: 365.9   Alpha:0.7       Epsilon:0.05
Step 343/1000   Reward: 423.52  Alpha:0.7       Epsilon:0.05
Step 344/1000   Reward: 205.1   Alpha:0.7       Epsilon:0.05
Step 345/1000   Reward: 528.18  Alpha:0.7       Epsilon:0.05
Step 346/1000   Reward: 390.3   Alpha:0.7       Epsilon:0.05
Step 347/1000   Reward: 410.82  Alpha:0.7       Epsilon:0.05
Step 348/1000   Reward: 393.12  Alpha:0.7       Epsilon:0.05
Step 349/1000   Reward: 477.54  Alpha:0.7       Epsilon:0.05
Step 350/1000   Reward: 414.6   Alpha:0.7       Epsilon:0.05
Step 351/1000   Reward: 669.1   Alpha:0.7       Epsilon:0.05
Step 352/1000   Reward: 514.98  Alpha:0.7       Epsilon:0.05
Step 353/1000   Reward: 402.86  Alpha:0.7       Epsilon:0.05
Step 354/1000   Reward: 458.16  Alpha:0.7       Epsilon:0.05
Step 355/1000   Reward: 497.3   Alpha:0.7       Epsilon:0.05
Step 356/1000   Reward: 483.76  Alpha:0.7       Epsilon:0.05
Step 357/1000   Reward: 395.64  Alpha:0.7       Epsilon:0.05
Step 358/1000   Reward: 436.5   Alpha:0.7       Epsilon:0.05
Step 359/1000   Reward: 488.02  Alpha:0.7       Epsilon:0.05
Step 360/1000   Reward: 410.84  Alpha:0.7       Epsilon:0.05
Step 361/1000   Reward: 310.48  Alpha:0.7       Epsilon:0.05
Step 362/1000   Reward: 278.26  Alpha:0.7       Epsilon:0.05
Step 363/1000   Reward: 553.12  Alpha:0.7       Epsilon:0.05
Step 364/1000   Reward: 451.02  Alpha:0.7       Epsilon:0.05
Step 365/1000   Reward: 616.5   Alpha:0.7       Epsilon:0.05
Step 366/1000   Reward: 571.8   Alpha:0.7       Epsilon:0.05
Step 367/1000   Reward: 505.34  Alpha:0.7       Epsilon:0.05
Step 368/1000   Reward: 345.58  Alpha:0.7       Epsilon:0.05
Step 369/1000   Reward: 410.42  Alpha:0.7       Epsilon:0.05
Step 370/1000   Reward: 460.82  Alpha:0.7       Epsilon:0.05
Step 371/1000   Reward: 618.98  Alpha:0.7       Epsilon:0.05
Step 372/1000   Reward: 465.34  Alpha:0.7       Epsilon:0.05
Step 373/1000   Reward: 418.5   Alpha:0.7       Epsilon:0.05
Step 374/1000   Reward: 439.66  Alpha:0.7       Epsilon:0.05
Step 375/1000   Reward: 426.5   Alpha:0.7       Epsilon:0.05
Step 376/1000   Reward: 416.32  Alpha:0.7       Epsilon:0.05
Step 377/1000   Reward: 291.4   Alpha:0.7       Epsilon:0.05
Step 378/1000   Reward: 550.8   Alpha:0.7       Epsilon:0.05
Step 379/1000   Reward: 435.94  Alpha:0.7       Epsilon:0.05
Step 380/1000   Reward: 424.66  Alpha:0.7       Epsilon:0.05
Step 381/1000   Reward: 686.78  Alpha:0.7       Epsilon:0.05
Step 382/1000   Reward: 667.54  Alpha:0.7       Epsilon:0.05
Step 383/1000   Reward: 455.12  Alpha:0.7       Epsilon:0.05
Step 384/1000   Reward: 422.8   Alpha:0.7       Epsilon:0.05
Step 385/1000   Reward: 324.4   Alpha:0.7       Epsilon:0.05
Step 386/1000   Reward: 591.96  Alpha:0.7       Epsilon:0.05
Step 387/1000   Reward: 412.94  Alpha:0.7       Epsilon:0.05
Step 388/1000   Reward: 546.62  Alpha:0.7       Epsilon:0.05
Step 389/1000   Reward: 578.32  Alpha:0.7       Epsilon:0.05
Step 390/1000   Reward: 394.48  Alpha:0.7       Epsilon:0.05
```

```
Step 391/1000   Reward: 578.26  Alpha:0.7       Epsilon:0.05
Step 392/1000   Reward: 516.6   Alpha:0.7       Epsilon:0.05
Step 393/1000   Reward: 394.16  Alpha:0.7       Epsilon:0.05
Step 394/1000   Reward: 454.16  Alpha:0.7       Epsilon:0.05
Step 395/1000   Reward: 518.94  Alpha:0.7       Epsilon:0.05
Step 396/1000   Reward: 558.2   Alpha:0.7       Epsilon:0.05
Step 397/1000   Reward: 669.02  Alpha:0.7       Epsilon:0.05
Step 398/1000   Reward: 393.94  Alpha:0.7       Epsilon:0.05
Step 399/1000   Reward: 354.84  Alpha:0.7       Epsilon:0.05
Step 400/1000   Reward: 480.78  Alpha:0.7       Epsilon:0.05
Step 401/1000   Reward: 436.84  Alpha:0.7       Epsilon:0.05
Step 402/1000   Reward: 427.22  Alpha:0.7       Epsilon:0.05
Step 403/1000   Reward: 329.56  Alpha:0.7       Epsilon:0.05
Step 404/1000   Reward: 439.76  Alpha:0.7       Epsilon:0.05
Step 405/1000   Reward: 503.66  Alpha:0.7       Epsilon:0.05
Step 406/1000   Reward: 372.02  Alpha:0.7       Epsilon:0.05
Step 407/1000   Reward: 340.7   Alpha:0.7       Epsilon:0.05
Step 408/1000   Reward: 426.24  Alpha:0.7       Epsilon:0.05
Step 409/1000   Reward: 521.56  Alpha:0.7       Epsilon:0.05
Step 410/1000   Reward: 594.26  Alpha:0.7       Epsilon:0.05
Step 411/1000   Reward: 485.38  Alpha:0.7       Epsilon:0.05
Step 412/1000   Reward: 483.26  Alpha:0.7       Epsilon:0.05
Step 413/1000   Reward: 511.76  Alpha:0.7       Epsilon:0.05
Step 414/1000   Reward: 360.16  Alpha:0.7       Epsilon:0.05
Step 415/1000   Reward: 636.26  Alpha:0.7       Epsilon:0.05
Step 416/1000   Reward: 545.66  Alpha:0.7       Epsilon:0.05
Step 417/1000   Reward: 482.8   Alpha:0.7       Epsilon:0.05
Step 418/1000   Reward: 429.7   Alpha:0.7       Epsilon:0.05
Step 419/1000   Reward: 584.38  Alpha:0.7       Epsilon:0.05
Step 420/1000   Reward: 625.16  Alpha:0.7       Epsilon:0.05
Step 421/1000   Reward: 418.5   Alpha:0.7       Epsilon:0.05
Step 422/1000   Reward: 473.48  Alpha:0.7       Epsilon:0.05
Step 423/1000   Reward: 461.04  Alpha:0.7       Epsilon:0.05
Step 424/1000   Reward: 518.86  Alpha:0.7       Epsilon:0.05
Step 425/1000   Reward: 676.08  Alpha:0.7       Epsilon:0.05
Step 426/1000   Reward: 587.24  Alpha:0.7       Epsilon:0.05
Step 427/1000   Reward: 490.28  Alpha:0.7       Epsilon:0.05
Step 428/1000   Reward: 636.44  Alpha:0.7       Epsilon:0.05
Step 429/1000   Reward: 507.76  Alpha:0.7       Epsilon:0.05
Step 430/1000   Reward: 532.26  Alpha:0.7       Epsilon:0.05
Step 431/1000   Reward: 254.86  Alpha:0.7       Epsilon:0.05
Step 432/1000   Reward: 352.02  Alpha:0.7       Epsilon:0.05
Step 433/1000   Reward: 433.66  Alpha:0.7       Epsilon:0.05
Step 434/1000   Reward: 437.26  Alpha:0.7       Epsilon:0.05
Step 435/1000   Reward: 558.3   Alpha:0.7       Epsilon:0.05
Step 436/1000   Reward: 501.6   Alpha:0.7       Epsilon:0.05
Step 437/1000   Reward: 669.8   Alpha:0.7       Epsilon:0.05
Step 438/1000   Reward: 667.74  Alpha:0.7       Epsilon:0.05
Step 439/1000   Reward: 448.96  Alpha:0.7       Epsilon:0.05
Step 440/1000   Reward: 511.62  Alpha:0.7       Epsilon:0.05
Step 441/1000   Reward: 676.72  Alpha:0.7       Epsilon:0.05
Step 442/1000   Reward: 523.34  Alpha:0.7       Epsilon:0.05
Step 443/1000   Reward: 637.12  Alpha:0.7       Epsilon:0.05
Step 444/1000   Reward: 374.86  Alpha:0.7       Epsilon:0.05
Step 445/1000   Reward: 379.34  Alpha:0.7       Epsilon:0.05
Step 446/1000   Reward: 607.54  Alpha:0.7       Epsilon:0.05
Step 447/1000   Reward: 527.16  Alpha:0.7       Epsilon:0.05
Step 448/1000   Reward: 506.9   Alpha:0.7       Epsilon:0.05
Step 449/1000   Reward: 559.02  Alpha:0.7       Epsilon:0.05
```

```
Step 450/1000   Reward: 483.1   Alpha:0.7       Epsilon:0.05
Step 451/1000   Reward: 480.2   Alpha:0.7       Epsilon:0.05
Step 452/1000   Reward: 456.06  Alpha:0.7       Epsilon:0.05
Step 453/1000   Reward: 531.32  Alpha:0.7       Epsilon:0.05
Step 454/1000   Reward: 283.8   Alpha:0.7       Epsilon:0.05
Step 455/1000   Reward: 553.66  Alpha:0.7       Epsilon:0.05
Step 456/1000   Reward: 313.86  Alpha:0.7       Epsilon:0.05
Step 457/1000   Reward: 656.14  Alpha:0.7       Epsilon:0.05
Step 458/1000   Reward: 480.02  Alpha:0.7       Epsilon:0.05
Step 459/1000   Reward: 396.28  Alpha:0.7       Epsilon:0.05
Step 460/1000   Reward: 475.18  Alpha:0.7       Epsilon:0.05
Step 461/1000   Reward: 418.38  Alpha:0.7       Epsilon:0.05
Step 462/1000   Reward: 592.32  Alpha:0.7       Epsilon:0.05
Step 463/1000   Reward: 643.44  Alpha:0.7       Epsilon:0.05
Step 464/1000   Reward: 654.12  Alpha:0.7       Epsilon:0.05
Step 465/1000   Reward: 474.04  Alpha:0.7       Epsilon:0.05
Step 466/1000   Reward: 656.88  Alpha:0.7       Epsilon:0.05
Step 467/1000   Reward: 432.7   Alpha:0.7       Epsilon:0.05
Step 468/1000   Reward: 659.94  Alpha:0.7       Epsilon:0.05
Step 469/1000   Reward: 590.9   Alpha:0.7       Epsilon:0.05
Step 470/1000   Reward: 591.42  Alpha:0.7       Epsilon:0.05
Step 471/1000   Reward: 385.56  Alpha:0.7       Epsilon:0.05
Step 472/1000   Reward: 675.18  Alpha:0.7       Epsilon:0.05
Step 473/1000   Reward: 537.88  Alpha:0.7       Epsilon:0.05
Step 474/1000   Reward: 621.32  Alpha:0.7       Epsilon:0.05
Step 475/1000   Reward: 578.6   Alpha:0.7       Epsilon:0.05
Step 476/1000   Reward: 550.4   Alpha:0.7       Epsilon:0.05
Step 477/1000   Reward: 573.86  Alpha:0.7       Epsilon:0.05
Step 478/1000   Reward: 586.16  Alpha:0.7       Epsilon:0.05
Step 479/1000   Reward: 419.98  Alpha:0.7       Epsilon:0.05
Step 480/1000   Reward: 616.04  Alpha:0.7       Epsilon:0.05
Step 481/1000   Reward: 586.04  Alpha:0.7       Epsilon:0.05
Step 482/1000   Reward: 699.54  Alpha:0.7       Epsilon:0.05
Step 483/1000   Reward: 366.02  Alpha:0.7       Epsilon:0.05
Step 484/1000   Reward: 538.52  Alpha:0.7       Epsilon:0.05
Step 485/1000   Reward: 552.08  Alpha:0.7       Epsilon:0.05
Step 486/1000   Reward: 646.64  Alpha:0.7       Epsilon:0.05
Step 487/1000   Reward: 598.38  Alpha:0.7       Epsilon:0.05
Step 488/1000   Reward: 459.66  Alpha:0.7       Epsilon:0.05
Step 489/1000   Reward: 626.08  Alpha:0.7       Epsilon:0.05
Step 490/1000   Reward: 721.22  Alpha:0.7       Epsilon:0.05
Step 491/1000   Reward: 468.78  Alpha:0.7       Epsilon:0.05
Step 492/1000   Reward: 539.96  Alpha:0.7       Epsilon:0.05
Step 493/1000   Reward: 740.92  Alpha:0.7       Epsilon:0.05
Step 494/1000   Reward: 574.26  Alpha:0.7       Epsilon:0.05
Step 495/1000   Reward: 619.26  Alpha:0.7       Epsilon:0.05
Step 496/1000   Reward: 550.04  Alpha:0.7       Epsilon:0.05
Step 497/1000   Reward: 308.36  Alpha:0.7       Epsilon:0.05
Step 498/1000   Reward: 489.58  Alpha:0.7       Epsilon:0.05
Step 499/1000   Reward: 609.0   Alpha:0.7       Epsilon:0.05
Step 500/1000   Reward: 621.56  Alpha:0.7       Epsilon:0.05
Step 501/1000   Reward: 454.78  Alpha:0.7       Epsilon:0.05
Step 502/1000   Reward: 721.82  Alpha:0.7       Epsilon:0.05
Step 503/1000   Reward: 666.44  Alpha:0.7       Epsilon:0.05
Step 504/1000   Reward: 586.54  Alpha:0.7       Epsilon:0.05
Step 505/1000   Reward: 609.0   Alpha:0.7       Epsilon:0.05
Step 506/1000   Reward: 495.08  Alpha:0.7       Epsilon:0.05
Step 507/1000   Reward: 677.28  Alpha:0.7       Epsilon:0.05
Step 508/1000   Reward: 539.74  Alpha:0.7       Epsilon:0.05
```

```
Step 509/1000   Reward: 610.6   Alpha:0.7       Epsilon:0.05
Step 510/1000   Reward: 606.9   Alpha:0.7       Epsilon:0.05
Step 511/1000   Reward: 540.68  Alpha:0.7       Epsilon:0.05
Step 512/1000   Reward: 588.94  Alpha:0.7       Epsilon:0.05
Step 513/1000   Reward: 452.96  Alpha:0.7       Epsilon:0.05
Step 514/1000   Reward: 567.16  Alpha:0.7       Epsilon:0.05
Step 515/1000   Reward: 513.42  Alpha:0.7       Epsilon:0.05
Step 516/1000   Reward: 531.62  Alpha:0.7       Epsilon:0.05
Step 517/1000   Reward: 885.4   Alpha:0.7       Epsilon:0.05
Step 518/1000   Reward: 616.94  Alpha:0.7       Epsilon:0.05
Step 519/1000   Reward: 586.64  Alpha:0.7       Epsilon:0.05
Step 520/1000   Reward: 650.8   Alpha:0.7       Epsilon:0.05
Step 521/1000   Reward: 552.32  Alpha:0.7       Epsilon:0.05
Step 522/1000   Reward: 616.56  Alpha:0.7       Epsilon:0.05
Step 523/1000   Reward: 678.96  Alpha:0.7       Epsilon:0.05
Step 524/1000   Reward: 291.82  Alpha:0.7       Epsilon:0.05
Step 525/1000   Reward: 710.02  Alpha:0.7       Epsilon:0.05
Step 526/1000   Reward: 510.66  Alpha:0.7       Epsilon:0.05
Step 527/1000   Reward: 442.08  Alpha:0.7       Epsilon:0.05
Step 528/1000   Reward: 542.38  Alpha:0.7       Epsilon:0.05
Step 529/1000   Reward: 481.46  Alpha:0.7       Epsilon:0.05
Step 530/1000   Reward: 398.22  Alpha:0.7       Epsilon:0.05
Step 531/1000   Reward: 674.62  Alpha:0.7       Epsilon:0.05
Step 532/1000   Reward: 602.5   Alpha:0.7       Epsilon:0.05
Step 533/1000   Reward: 479.26  Alpha:0.7       Epsilon:0.05
Step 534/1000   Reward: 422.62  Alpha:0.7       Epsilon:0.05
Step 535/1000   Reward: 345.22  Alpha:0.7       Epsilon:0.05
Step 536/1000   Reward: 564.64  Alpha:0.7       Epsilon:0.05
Step 537/1000   Reward: 637.04  Alpha:0.7       Epsilon:0.05
Step 538/1000   Reward: 631.04  Alpha:0.7       Epsilon:0.05
Step 539/1000   Reward: 598.16  Alpha:0.7       Epsilon:0.05
Step 540/1000   Reward: 620.04  Alpha:0.7       Epsilon:0.05
Step 541/1000   Reward: 416.86  Alpha:0.7       Epsilon:0.05
Step 542/1000   Reward: 664.0   Alpha:0.7       Epsilon:0.05
Step 543/1000   Reward: 366.32  Alpha:0.7       Epsilon:0.05
Step 544/1000   Reward: 508.18  Alpha:0.7       Epsilon:0.05
Step 545/1000   Reward: 682.14  Alpha:0.7       Epsilon:0.05
Step 546/1000   Reward: 623.88  Alpha:0.7       Epsilon:0.05
Step 547/1000   Reward: 663.16  Alpha:0.7       Epsilon:0.05
Step 548/1000   Reward: 690.54  Alpha:0.7       Epsilon:0.05
Step 549/1000   Reward: 705.28  Alpha:0.7       Epsilon:0.05
Step 550/1000   Reward: 506.12  Alpha:0.7       Epsilon:0.05
Step 551/1000   Reward: 503.34  Alpha:0.7       Epsilon:0.05
Step 552/1000   Reward: 608.5   Alpha:0.7       Epsilon:0.05
Step 553/1000   Reward: 578.7   Alpha:0.7       Epsilon:0.05
Step 554/1000   Reward: 555.94  Alpha:0.7       Epsilon:0.05
Step 555/1000   Reward: 582.12  Alpha:0.7       Epsilon:0.05
Step 556/1000   Reward: 600.38  Alpha:0.7       Epsilon:0.05
Step 557/1000   Reward: 666.62  Alpha:0.7       Epsilon:0.05
Step 558/1000   Reward: 601.44  Alpha:0.7       Epsilon:0.05
Step 559/1000   Reward: 464.88  Alpha:0.7       Epsilon:0.05
Step 560/1000   Reward: 627.04  Alpha:0.7       Epsilon:0.05
Step 561/1000   Reward: 447.08  Alpha:0.7       Epsilon:0.05
Step 562/1000   Reward: 818.82  Alpha:0.7       Epsilon:0.05
Step 563/1000   Reward: 569.0   Alpha:0.7       Epsilon:0.05
Step 564/1000   Reward: 674.48  Alpha:0.7       Epsilon:0.05
Step 565/1000   Reward: 789.28  Alpha:0.7       Epsilon:0.05
Step 566/1000   Reward: 715.7   Alpha:0.7       Epsilon:0.05
Step 567/1000   Reward: 617.9   Alpha:0.7       Epsilon:0.05
```

```
Step 568/1000   Reward: 548.8   Alpha:0.7       Epsilon:0.05
Step 569/1000   Reward: 662.0   Alpha:0.7       Epsilon:0.05
Step 570/1000   Reward: 666.22  Alpha:0.7       Epsilon:0.05
Step 571/1000   Reward: 701.0   Alpha:0.7       Epsilon:0.05
Step 572/1000   Reward: 512.56  Alpha:0.7       Epsilon:0.05
Step 573/1000   Reward: 421.3   Alpha:0.7       Epsilon:0.05
Step 574/1000   Reward: 771.72  Alpha:0.7       Epsilon:0.05
Step 575/1000   Reward: 598.34  Alpha:0.7       Epsilon:0.05
Step 576/1000   Reward: 896.6   Alpha:0.7       Epsilon:0.05
Step 577/1000   Reward: 620.96  Alpha:0.7       Epsilon:0.05
Step 578/1000   Reward: 623.3   Alpha:0.7       Epsilon:0.05
Step 579/1000   Reward: 447.18  Alpha:0.7       Epsilon:0.05
Step 580/1000   Reward: 557.5   Alpha:0.7       Epsilon:0.05
Step 581/1000   Reward: 449.9   Alpha:0.7       Epsilon:0.05
Step 582/1000   Reward: 654.1   Alpha:0.7       Epsilon:0.05
Step 583/1000   Reward: 726.54  Alpha:0.7       Epsilon:0.05
Step 584/1000   Reward: 492.22  Alpha:0.7       Epsilon:0.05
Step 585/1000   Reward: 648.04  Alpha:0.7       Epsilon:0.05
Step 586/1000   Reward: 657.58  Alpha:0.7       Epsilon:0.05
Step 587/1000   Reward: 572.56  Alpha:0.7       Epsilon:0.05
Step 588/1000   Reward: 641.22  Alpha:0.7       Epsilon:0.05
Step 589/1000   Reward: 719.86  Alpha:0.7       Epsilon:0.05
Step 590/1000   Reward: 574.86  Alpha:0.7       Epsilon:0.05
Step 591/1000   Reward: 674.46  Alpha:0.7       Epsilon:0.05
Step 592/1000   Reward: 752.26  Alpha:0.7       Epsilon:0.05
Step 593/1000   Reward: 550.06  Alpha:0.7       Epsilon:0.05
Step 594/1000   Reward: 404.66  Alpha:0.7       Epsilon:0.05
Step 595/1000   Reward: 715.68  Alpha:0.7       Epsilon:0.05
Step 596/1000   Reward: 647.56  Alpha:0.7       Epsilon:0.05
Step 597/1000   Reward: 636.32  Alpha:0.7       Epsilon:0.05
Step 598/1000   Reward: 654.56  Alpha:0.7       Epsilon:0.05
Step 599/1000   Reward: 563.66  Alpha:0.7       Epsilon:0.05
Step 600/1000   Reward: 573.44  Alpha:0.7       Epsilon:0.05
Step 601/1000   Reward: 638.36  Alpha:0.7       Epsilon:0.05
Step 602/1000   Reward: 475.94  Alpha:0.7       Epsilon:0.05
Step 603/1000   Reward: 613.7   Alpha:0.7       Epsilon:0.05
Step 604/1000   Reward: 682.42  Alpha:0.7       Epsilon:0.05
Step 605/1000   Reward: 749.66  Alpha:0.7       Epsilon:0.05
Step 606/1000   Reward: 653.52  Alpha:0.7       Epsilon:0.05
Step 607/1000   Reward: 643.26  Alpha:0.7       Epsilon:0.05
Step 608/1000   Reward: 700.8   Alpha:0.7       Epsilon:0.05
Step 609/1000   Reward: 701.44  Alpha:0.7       Epsilon:0.05
Step 610/1000   Reward: 685.72  Alpha:0.7       Epsilon:0.05
Step 611/1000   Reward: 825.1   Alpha:0.7       Epsilon:0.05
Step 612/1000   Reward: 628.6   Alpha:0.7       Epsilon:0.05
Step 613/1000   Reward: 514.9   Alpha:0.7       Epsilon:0.05
Step 614/1000   Reward: 873.32  Alpha:0.7       Epsilon:0.05
Step 615/1000   Reward: 739.08  Alpha:0.7       Epsilon:0.05
Step 616/1000   Reward: 718.74  Alpha:0.7       Epsilon:0.05
Step 617/1000   Reward: 571.08  Alpha:0.7       Epsilon:0.05
Step 618/1000   Reward: 752.38  Alpha:0.7       Epsilon:0.05
Step 619/1000   Reward: 500.54  Alpha:0.7       Epsilon:0.05
Step 620/1000   Reward: 751.3   Alpha:0.7       Epsilon:0.05
Step 621/1000   Reward: 697.02  Alpha:0.7       Epsilon:0.05
Step 622/1000   Reward: 584.76  Alpha:0.7       Epsilon:0.05
Step 623/1000   Reward: 733.9   Alpha:0.7       Epsilon:0.05
Step 624/1000   Reward: 725.4   Alpha:0.7       Epsilon:0.05
Step 625/1000   Reward: 989.3   Alpha:0.7       Epsilon:0.05
Step 626/1000   Reward: 638.34  Alpha:0.7       Epsilon:0.05
```

```
Step 627/1000   Reward: 851.32  Alpha:0.7       Epsilon:0.05
Step 628/1000   Reward: 723.66  Alpha:0.7       Epsilon:0.05
Step 629/1000   Reward: 548.7   Alpha:0.7       Epsilon:0.05
Step 630/1000   Reward: 656.06  Alpha:0.7       Epsilon:0.05
Step 631/1000   Reward: 796.26  Alpha:0.7       Epsilon:0.05
Step 632/1000   Reward: 754.58  Alpha:0.7       Epsilon:0.05
Step 633/1000   Reward: 672.54  Alpha:0.7       Epsilon:0.05
Step 634/1000   Reward: 551.3   Alpha:0.7       Epsilon:0.05
Step 635/1000   Reward: 591.74  Alpha:0.7       Epsilon:0.05
Step 636/1000   Reward: 588.26  Alpha:0.7       Epsilon:0.05
Step 637/1000   Reward: 658.78  Alpha:0.7       Epsilon:0.05
Step 638/1000   Reward: 515.36  Alpha:0.7       Epsilon:0.05
Step 639/1000   Reward: 924.86  Alpha:0.7       Epsilon:0.05
Step 640/1000   Reward: 620.66  Alpha:0.7       Epsilon:0.05
Step 641/1000   Reward: 752.02  Alpha:0.7       Epsilon:0.05
Step 642/1000   Reward: 531.06  Alpha:0.7       Epsilon:0.05
Step 643/1000   Reward: 670.88  Alpha:0.7       Epsilon:0.05
Step 644/1000   Reward: 644.8   Alpha:0.7       Epsilon:0.05
Step 645/1000   Reward: 606.2   Alpha:0.7       Epsilon:0.05
Step 646/1000   Reward: 481.06  Alpha:0.7       Epsilon:0.05
Step 647/1000   Reward: 757.22  Alpha:0.7       Epsilon:0.05
Step 648/1000   Reward: 641.58  Alpha:0.7       Epsilon:0.05
Step 649/1000   Reward: 540.22  Alpha:0.7       Epsilon:0.05
Step 650/1000   Reward: 620.64  Alpha:0.7       Epsilon:0.05
Step 651/1000   Reward: 699.74  Alpha:0.7       Epsilon:0.05
Step 652/1000   Reward: 732.68  Alpha:0.7       Epsilon:0.05
Step 653/1000   Reward: 502.3   Alpha:0.7       Epsilon:0.05
Step 654/1000   Reward: 687.14  Alpha:0.7       Epsilon:0.05
Step 655/1000   Reward: 807.56  Alpha:0.7       Epsilon:0.05
Step 656/1000   Reward: 542.74  Alpha:0.7       Epsilon:0.05
Step 657/1000   Reward: 540.16  Alpha:0.7       Epsilon:0.05
Step 658/1000   Reward: 625.88  Alpha:0.7       Epsilon:0.05
Step 659/1000   Reward: 767.7   Alpha:0.7       Epsilon:0.05
Step 660/1000   Reward: 694.52  Alpha:0.7       Epsilon:0.05
Step 661/1000   Reward: 681.06  Alpha:0.7       Epsilon:0.05
Step 662/1000   Reward: 713.96  Alpha:0.7       Epsilon:0.05
Step 663/1000   Reward: 612.38  Alpha:0.7       Epsilon:0.05
Step 664/1000   Reward: 641.38  Alpha:0.7       Epsilon:0.05
Step 665/1000   Reward: 762.9   Alpha:0.7       Epsilon:0.05
Step 666/1000   Reward: 745.18  Alpha:0.7       Epsilon:0.05
Step 667/1000   Reward: 614.8   Alpha:0.7       Epsilon:0.05
Step 668/1000   Reward: 621.2   Alpha:0.7       Epsilon:0.05
Step 669/1000   Reward: 701.58  Alpha:0.7       Epsilon:0.05
Step 670/1000   Reward: 525.5   Alpha:0.7       Epsilon:0.05
Step 671/1000   Reward: 743.98  Alpha:0.7       Epsilon:0.05
Step 672/1000   Reward: 833.92  Alpha:0.7       Epsilon:0.05
Step 673/1000   Reward: 904.28  Alpha:0.7       Epsilon:0.05
Step 674/1000   Reward: 773.12  Alpha:0.7       Epsilon:0.05
Step 675/1000   Reward: 659.64  Alpha:0.7       Epsilon:0.05
Step 676/1000   Reward: 562.2   Alpha:0.7       Epsilon:0.05
Step 677/1000   Reward: 811.24  Alpha:0.7       Epsilon:0.05
Step 678/1000   Reward: 765.72  Alpha:0.7       Epsilon:0.05
Step 679/1000   Reward: 859.58  Alpha:0.7       Epsilon:0.05
Step 680/1000   Reward: 550.06  Alpha:0.7       Epsilon:0.05
Step 681/1000   Reward: 758.36  Alpha:0.7       Epsilon:0.05
Step 682/1000   Reward: 818.12  Alpha:0.7       Epsilon:0.05
Step 683/1000   Reward: 704.36  Alpha:0.7       Epsilon:0.05
Step 684/1000   Reward: 742.26  Alpha:0.7       Epsilon:0.05
Step 685/1000   Reward: 821.86  Alpha:0.7       Epsilon:0.05
```

```
Step 686/1000   Reward: 818.3   Alpha:0.7       Epsilon:0.05
Step 687/1000   Reward: 1023.56 Alpha:0.7       Epsilon:0.05
Step 688/1000   Reward: 977.28  Alpha:0.7       Epsilon:0.05
Step 689/1000   Reward: 688.1   Alpha:0.7       Epsilon:0.05
Step 690/1000   Reward: 826.64  Alpha:0.7       Epsilon:0.05
Step 691/1000   Reward: 587.3   Alpha:0.7       Epsilon:0.05
Step 692/1000   Reward: 803.6   Alpha:0.7       Epsilon:0.05
Step 693/1000   Reward: 715.78  Alpha:0.7       Epsilon:0.05
Step 694/1000   Reward: 867.12  Alpha:0.7       Epsilon:0.05
Step 695/1000   Reward: 681.56  Alpha:0.7       Epsilon:0.05
Step 696/1000   Reward: 712.62  Alpha:0.7       Epsilon:0.05
Step 697/1000   Reward: 708.78  Alpha:0.7       Epsilon:0.05
Step 698/1000   Reward: 907.9   Alpha:0.7       Epsilon:0.05
Step 699/1000   Reward: 754.56  Alpha:0.7       Epsilon:0.05
Step 700/1000   Reward: 653.48  Alpha:0.7       Epsilon:0.05
Step 701/1000   Reward: 757.02  Alpha:0.7       Epsilon:0.05
Step 702/1000   Reward: 756.72  Alpha:0.7       Epsilon:0.05
Step 703/1000   Reward: 961.26  Alpha:0.7       Epsilon:0.05
Step 704/1000   Reward: 600.58  Alpha:0.7       Epsilon:0.05
Step 705/1000   Reward: 816.58  Alpha:0.7       Epsilon:0.05
Step 706/1000   Reward: 668.54  Alpha:0.7       Epsilon:0.05
Step 707/1000   Reward: 543.5   Alpha:0.7       Epsilon:0.05
Step 708/1000   Reward: 640.34  Alpha:0.7       Epsilon:0.05
Step 709/1000   Reward: 659.16  Alpha:0.7       Epsilon:0.05
Step 710/1000   Reward: 703.58  Alpha:0.7       Epsilon:0.05
Step 711/1000   Reward: 783.64  Alpha:0.7       Epsilon:0.05
Step 712/1000   Reward: 615.34  Alpha:0.7       Epsilon:0.05
Step 713/1000   Reward: 832.64  Alpha:0.7       Epsilon:0.05
Step 714/1000   Reward: 902.02  Alpha:0.7       Epsilon:0.05
Step 715/1000   Reward: 775.54  Alpha:0.7       Epsilon:0.05
Step 716/1000   Reward: 794.5   Alpha:0.7       Epsilon:0.05
Step 717/1000   Reward: 783.08  Alpha:0.7       Epsilon:0.05
Step 718/1000   Reward: 594.82  Alpha:0.7       Epsilon:0.05
Step 719/1000   Reward: 741.2   Alpha:0.7       Epsilon:0.05
Step 720/1000   Reward: 563.6   Alpha:0.7       Epsilon:0.05
Step 721/1000   Reward: 685.88  Alpha:0.7       Epsilon:0.05
Step 722/1000   Reward: 815.9   Alpha:0.7       Epsilon:0.05
Step 723/1000   Reward: 868.56  Alpha:0.7       Epsilon:0.05
Step 724/1000   Reward: 808.14  Alpha:0.7       Epsilon:0.05
Step 725/1000   Reward: 701.22  Alpha:0.7       Epsilon:0.05
Step 726/1000   Reward: 717.12  Alpha:0.7       Epsilon:0.05
Step 727/1000   Reward: 915.84  Alpha:0.7       Epsilon:0.05
Step 728/1000   Reward: 783.46  Alpha:0.7       Epsilon:0.05
Step 729/1000   Reward: 588.28  Alpha:0.7       Epsilon:0.05
Step 730/1000   Reward: 705.44  Alpha:0.7       Epsilon:0.05
Step 731/1000   Reward: 870.38  Alpha:0.7       Epsilon:0.05
Step 732/1000   Reward: 773.4   Alpha:0.7       Epsilon:0.05
Step 733/1000   Reward: 841.44  Alpha:0.7       Epsilon:0.05
Step 734/1000   Reward: 729.84  Alpha:0.7       Epsilon:0.05
Step 735/1000   Reward: 642.82  Alpha:0.7       Epsilon:0.05
Step 736/1000   Reward: 783.06  Alpha:0.7       Epsilon:0.05
Step 737/1000   Reward: 622.5   Alpha:0.7       Epsilon:0.05
Step 738/1000   Reward: 698.82  Alpha:0.7       Epsilon:0.05
Step 739/1000   Reward: 751.78  Alpha:0.7       Epsilon:0.05
Step 740/1000   Reward: 713.4   Alpha:0.7       Epsilon:0.05
Step 741/1000   Reward: 1139.28 Alpha:0.7       Epsilon:0.05
Step 742/1000   Reward: 729.62  Alpha:0.7       Epsilon:0.05
Step 743/1000   Reward: 735.64  Alpha:0.7       Epsilon:0.05
Step 744/1000   Reward: 973.98  Alpha:0.7       Epsilon:0.05
```

```
Step 745/1000   Reward: 736.04  Alpha:0.7       Epsilon:0.05
Step 746/1000   Reward: 847.08  Alpha:0.7       Epsilon:0.05
Step 747/1000   Reward: 597.24  Alpha:0.7       Epsilon:0.05
Step 748/1000   Reward: 810.4   Alpha:0.7       Epsilon:0.05
Step 749/1000   Reward: 879.0   Alpha:0.7       Epsilon:0.05
Step 750/1000   Reward: 623.18  Alpha:0.7       Epsilon:0.05
Step 751/1000   Reward: 692.38  Alpha:0.7       Epsilon:0.05
Step 752/1000   Reward: 646.42  Alpha:0.7       Epsilon:0.05
Step 753/1000   Reward: 955.4   Alpha:0.7       Epsilon:0.05
Step 754/1000   Reward: 713.58  Alpha:0.7       Epsilon:0.05
Step 755/1000   Reward: 801.42  Alpha:0.7       Epsilon:0.05
Step 756/1000   Reward: 832.54  Alpha:0.7       Epsilon:0.05
Step 757/1000   Reward: 753.2   Alpha:0.7       Epsilon:0.05
Step 758/1000   Reward: 705.9   Alpha:0.7       Epsilon:0.05
Step 759/1000   Reward: 780.04  Alpha:0.7       Epsilon:0.05
Step 760/1000   Reward: 898.66  Alpha:0.7       Epsilon:0.05
Step 761/1000   Reward: 695.26  Alpha:0.7       Epsilon:0.05
Step 762/1000   Reward: 909.2   Alpha:0.7       Epsilon:0.05
Step 763/1000   Reward: 763.76  Alpha:0.7       Epsilon:0.05
Step 764/1000   Reward: 926.98  Alpha:0.7       Epsilon:0.05
Step 765/1000   Reward: 909.76  Alpha:0.7       Epsilon:0.05
Step 766/1000   Reward: 681.3   Alpha:0.7       Epsilon:0.05
Step 767/1000   Reward: 843.46  Alpha:0.7       Epsilon:0.05
Step 768/1000   Reward: 901.76  Alpha:0.7       Epsilon:0.05
Step 769/1000   Reward: 694.94  Alpha:0.7       Epsilon:0.05
Step 770/1000   Reward: 1008.7  Alpha:0.7       Epsilon:0.05
Step 771/1000   Reward: 816.3   Alpha:0.7       Epsilon:0.05
Step 772/1000   Reward: 911.46  Alpha:0.7       Epsilon:0.05
Step 773/1000   Reward: 973.8   Alpha:0.7       Epsilon:0.05
Step 774/1000   Reward: 796.06  Alpha:0.7       Epsilon:0.05
Step 775/1000   Reward: 655.04  Alpha:0.7       Epsilon:0.05
Step 776/1000   Reward: 890.56  Alpha:0.7       Epsilon:0.05
Step 777/1000   Reward: 620.32  Alpha:0.7       Epsilon:0.05
Step 778/1000   Reward: 850.3   Alpha:0.7       Epsilon:0.05
Step 779/1000   Reward: 819.9   Alpha:0.7       Epsilon:0.05
Step 780/1000   Reward: 835.92  Alpha:0.7       Epsilon:0.05
Step 781/1000   Reward: 633.1   Alpha:0.7       Epsilon:0.05
Step 782/1000   Reward: 669.42  Alpha:0.7       Epsilon:0.05
Step 783/1000   Reward: 1040.56 Alpha:0.7       Epsilon:0.05
Step 784/1000   Reward: 962.32  Alpha:0.7       Epsilon:0.05
Step 785/1000   Reward: 713.9   Alpha:0.7       Epsilon:0.05
Step 786/1000   Reward: 745.72  Alpha:0.7       Epsilon:0.05
Step 787/1000   Reward: 646.48  Alpha:0.7       Epsilon:0.05
Step 788/1000   Reward: 480.44  Alpha:0.7       Epsilon:0.05
Step 789/1000   Reward: 607.3   Alpha:0.7       Epsilon:0.05
Step 790/1000   Reward: 670.3   Alpha:0.7       Epsilon:0.05
Step 791/1000   Reward: 1065.84 Alpha:0.7       Epsilon:0.05
Step 792/1000   Reward: 770.14  Alpha:0.7       Epsilon:0.05
Step 793/1000   Reward: 819.84  Alpha:0.7       Epsilon:0.05
Step 794/1000   Reward: 799.06  Alpha:0.7       Epsilon:0.05
Step 795/1000   Reward: 924.0   Alpha:0.7       Epsilon:0.05
Step 796/1000   Reward: 750.16  Alpha:0.7       Epsilon:0.05
Step 797/1000   Reward: 727.74  Alpha:0.7       Epsilon:0.05
Step 798/1000   Reward: 1048.5  Alpha:0.7       Epsilon:0.05
Step 799/1000   Reward: 746.38  Alpha:0.7       Epsilon:0.05
Step 800/1000   Reward: 821.12  Alpha:0.7       Epsilon:0.05
Step 801/1000   Reward: 883.06  Alpha:0.7       Epsilon:0.05
Step 802/1000   Reward: 704.38  Alpha:0.7       Epsilon:0.05
Step 803/1000   Reward: 772.4   Alpha:0.7       Epsilon:0.05
```

```
Step 804/1000   Reward: 797.26  Alpha:0.7       Epsilon:0.05
Step 805/1000   Reward: 1000.46 Alpha:0.7       Epsilon:0.05
Step 806/1000   Reward: 900.14  Alpha:0.7       Epsilon:0.05
Step 807/1000   Reward: 784.84  Alpha:0.7       Epsilon:0.05
Step 808/1000   Reward: 925.06  Alpha:0.7       Epsilon:0.05
Step 809/1000   Reward: 859.22  Alpha:0.7       Epsilon:0.05
Step 810/1000   Reward: 646.32  Alpha:0.7       Epsilon:0.05
Step 811/1000   Reward: 913.2   Alpha:0.7       Epsilon:0.05
Step 812/1000   Reward: 929.52  Alpha:0.7       Epsilon:0.05
Step 813/1000   Reward: 696.76  Alpha:0.7       Epsilon:0.05
Step 814/1000   Reward: 854.32  Alpha:0.7       Epsilon:0.05
Step 815/1000   Reward: 828.24  Alpha:0.7       Epsilon:0.05
Step 816/1000   Reward: 759.18  Alpha:0.7       Epsilon:0.05
Step 817/1000   Reward: 916.58  Alpha:0.7       Epsilon:0.05
Step 818/1000   Reward: 810.6   Alpha:0.7       Epsilon:0.05
Step 819/1000   Reward: 900.02  Alpha:0.7       Epsilon:0.05
Step 820/1000   Reward: 906.32  Alpha:0.7       Epsilon:0.05
Step 821/1000   Reward: 927.98  Alpha:0.7       Epsilon:0.05
Step 822/1000   Reward: 605.5   Alpha:0.7       Epsilon:0.05
Step 823/1000   Reward: 983.08  Alpha:0.7       Epsilon:0.05
Step 824/1000   Reward: 843.18  Alpha:0.7       Epsilon:0.05
Step 825/1000   Reward: 779.26  Alpha:0.7       Epsilon:0.05
Step 826/1000   Reward: 987.1   Alpha:0.7       Epsilon:0.05
Step 827/1000   Reward: 747.5   Alpha:0.7       Epsilon:0.05
Step 828/1000   Reward: 915.44  Alpha:0.7       Epsilon:0.05
Step 829/1000   Reward: 942.24  Alpha:0.7       Epsilon:0.05
Step 830/1000   Reward: 848.74  Alpha:0.7       Epsilon:0.05
Step 831/1000   Reward: 911.68  Alpha:0.7       Epsilon:0.05
Step 832/1000   Reward: 882.24  Alpha:0.7       Epsilon:0.05
Step 833/1000   Reward: 912.18  Alpha:0.7       Epsilon:0.05
Step 834/1000   Reward: 976.88  Alpha:0.7       Epsilon:0.05
Step 835/1000   Reward: 802.8   Alpha:0.7       Epsilon:0.05
Step 836/1000   Reward: 928.0   Alpha:0.7       Epsilon:0.05
Step 837/1000   Reward: 1043.78 Alpha:0.7       Epsilon:0.05
Step 838/1000   Reward: 1122.46 Alpha:0.7       Epsilon:0.05
Step 839/1000   Reward: 743.04  Alpha:0.7       Epsilon:0.05
Step 840/1000   Reward: 755.82  Alpha:0.7       Epsilon:0.05
Step 841/1000   Reward: 834.4   Alpha:0.7       Epsilon:0.05
Step 842/1000   Reward: 922.72  Alpha:0.7       Epsilon:0.05
Step 843/1000   Reward: 844.92  Alpha:0.7       Epsilon:0.05
Step 844/1000   Reward: 888.2   Alpha:0.7       Epsilon:0.05
Step 845/1000   Reward: 670.38  Alpha:0.7       Epsilon:0.05
Step 846/1000   Reward: 806.72  Alpha:0.7       Epsilon:0.05
Step 847/1000   Reward: 868.24  Alpha:0.7       Epsilon:0.05
Step 848/1000   Reward: 934.22  Alpha:0.7       Epsilon:0.05
Step 849/1000   Reward: 890.3   Alpha:0.7       Epsilon:0.05
Step 850/1000   Reward: 924.56  Alpha:0.7       Epsilon:0.05
Step 851/1000   Reward: 758.04  Alpha:0.7       Epsilon:0.05
Step 852/1000   Reward: 771.38  Alpha:0.7       Epsilon:0.05
Step 853/1000   Reward: 1106.88 Alpha:0.7       Epsilon:0.05
Step 854/1000   Reward: 760.24  Alpha:0.7       Epsilon:0.05
Step 855/1000   Reward: 957.14  Alpha:0.7       Epsilon:0.05
Step 856/1000   Reward: 836.46  Alpha:0.7       Epsilon:0.05
Step 857/1000   Reward: 1092.48 Alpha:0.7       Epsilon:0.05
Step 858/1000   Reward: 949.48  Alpha:0.7       Epsilon:0.05
Step 859/1000   Reward: 829.24  Alpha:0.7       Epsilon:0.05
Step 860/1000   Reward: 1005.02 Alpha:0.7       Epsilon:0.05
Step 861/1000   Reward: 1137.06 Alpha:0.7       Epsilon:0.05
Step 862/1000   Reward: 880.28  Alpha:0.7       Epsilon:0.05
```

```
Step 863/1000   Reward: 767.26  Alpha:0.7       Epsilon:0.05
Step 864/1000   Reward: 995.9   Alpha:0.7       Epsilon:0.05
Step 865/1000   Reward: 902.86  Alpha:0.7       Epsilon:0.05
Step 866/1000   Reward: 931.66  Alpha:0.7       Epsilon:0.05
Step 867/1000   Reward: 1029.8  Alpha:0.7       Epsilon:0.05
Step 868/1000   Reward: 788.02  Alpha:0.7       Epsilon:0.05
Step 869/1000   Reward: 1039.22 Alpha:0.7       Epsilon:0.05
Step 870/1000   Reward: 1043.2  Alpha:0.7       Epsilon:0.05
Step 871/1000   Reward: 998.34  Alpha:0.7       Epsilon:0.05
Step 872/1000   Reward: 970.66  Alpha:0.7       Epsilon:0.05
Step 873/1000   Reward: 852.82  Alpha:0.7       Epsilon:0.05
Step 874/1000   Reward: 799.82  Alpha:0.7       Epsilon:0.05
Step 875/1000   Reward: 866.44  Alpha:0.7       Epsilon:0.05
Step 876/1000   Reward: 933.86  Alpha:0.7       Epsilon:0.05
Step 877/1000   Reward: 853.1   Alpha:0.7       Epsilon:0.05
Step 878/1000   Reward: 839.56  Alpha:0.7       Epsilon:0.05
Step 879/1000   Reward: 740.94  Alpha:0.7       Epsilon:0.05
Step 880/1000   Reward: 921.02  Alpha:0.7       Epsilon:0.05
Step 881/1000   Reward: 860.24  Alpha:0.7       Epsilon:0.05
Step 882/1000   Reward: 828.06  Alpha:0.7       Epsilon:0.05
Step 883/1000   Reward: 980.88  Alpha:0.7       Epsilon:0.05
Step 884/1000   Reward: 1001.32 Alpha:0.7       Epsilon:0.05
Step 885/1000   Reward: 698.32  Alpha:0.7       Epsilon:0.05
Step 886/1000   Reward: 765.36  Alpha:0.7       Epsilon:0.05
Step 887/1000   Reward: 837.34  Alpha:0.7       Epsilon:0.05
Step 888/1000   Reward: 994.4   Alpha:0.7       Epsilon:0.05
Step 889/1000   Reward: 899.66  Alpha:0.7       Epsilon:0.05
Step 890/1000   Reward: 743.02  Alpha:0.7       Epsilon:0.05
Step 891/1000   Reward: 951.02  Alpha:0.7       Epsilon:0.05
Step 892/1000   Reward: 759.18  Alpha:0.7       Epsilon:0.05
Step 893/1000   Reward: 980.08  Alpha:0.7       Epsilon:0.05
Step 894/1000   Reward: 1255.08 Alpha:0.7       Epsilon:0.05
Step 895/1000   Reward: 912.02  Alpha:0.7       Epsilon:0.05
Step 896/1000   Reward: 781.32  Alpha:0.7       Epsilon:0.05
Step 897/1000   Reward: 1012.14 Alpha:0.7       Epsilon:0.05
Step 898/1000   Reward: 890.18  Alpha:0.7       Epsilon:0.05
Step 899/1000   Reward: 759.94  Alpha:0.7       Epsilon:0.05
Step 900/1000   Reward: 851.86  Alpha:0.7       Epsilon:0.05
Step 901/1000   Reward: 850.9   Alpha:0.7       Epsilon:0.05
Step 902/1000   Reward: 756.18  Alpha:0.7       Epsilon:0.05
Step 903/1000   Reward: 748.66  Alpha:0.7       Epsilon:0.05
Step 904/1000   Reward: 751.4   Alpha:0.7       Epsilon:0.05
Step 905/1000   Reward: 805.98  Alpha:0.7       Epsilon:0.05
Step 906/1000   Reward: 769.5   Alpha:0.7       Epsilon:0.05
Step 907/1000   Reward: 1058.08 Alpha:0.7       Epsilon:0.05
Step 908/1000   Reward: 1005.02 Alpha:0.7       Epsilon:0.05
Step 909/1000   Reward: 1110.14 Alpha:0.7       Epsilon:0.05
Step 910/1000   Reward: 956.32  Alpha:0.7       Epsilon:0.05
Step 911/1000   Reward: 950.04  Alpha:0.7       Epsilon:0.05
Step 912/1000   Reward: 896.66  Alpha:0.7       Epsilon:0.05
Step 913/1000   Reward: 971.14  Alpha:0.7       Epsilon:0.05
Step 914/1000   Reward: 1118.74 Alpha:0.7       Epsilon:0.05
Step 915/1000   Reward: 1028.04 Alpha:0.7       Epsilon:0.05
Step 916/1000   Reward: 835.14  Alpha:0.7       Epsilon:0.05
Step 917/1000   Reward: 1010.32 Alpha:0.7       Epsilon:0.05
Step 918/1000   Reward: 1206.56 Alpha:0.7       Epsilon:0.05
Step 919/1000   Reward: 940.6   Alpha:0.7       Epsilon:0.05
Step 920/1000   Reward: 648.62  Alpha:0.7       Epsilon:0.05
Step 921/1000   Reward: 1043.4  Alpha:0.7       Epsilon:0.05
```

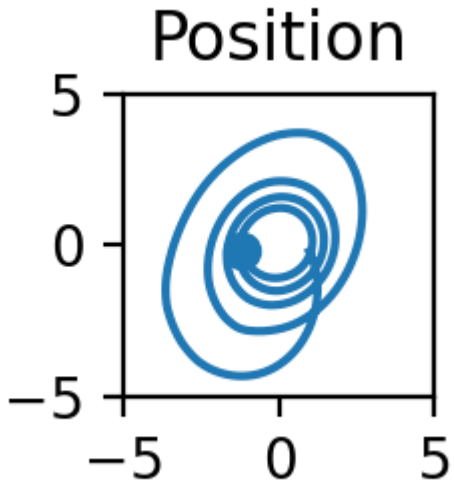
```
Step 922/1000   Reward: 938.66  Alpha:0.7       Epsilon:0.05
Step 923/1000   Reward: 815.32  Alpha:0.7       Epsilon:0.05
Step 924/1000   Reward: 761.7   Alpha:0.7       Epsilon:0.05
Step 925/1000   Reward: 1141.08 Alpha:0.7       Epsilon:0.05
Step 926/1000   Reward: 1043.8  Alpha:0.7       Epsilon:0.05
Step 927/1000   Reward: 866.76  Alpha:0.7       Epsilon:0.05
Step 928/1000   Reward: 762.86  Alpha:0.7       Epsilon:0.05
Step 929/1000   Reward: 718.12  Alpha:0.7       Epsilon:0.05
Step 930/1000   Reward: 1056.42 Alpha:0.7       Epsilon:0.05
Step 931/1000   Reward: 1065.3  Alpha:0.7       Epsilon:0.05
Step 932/1000   Reward: 1104.02 Alpha:0.7       Epsilon:0.05
Step 933/1000   Reward: 987.38  Alpha:0.7       Epsilon:0.05
Step 934/1000   Reward: 994.26  Alpha:0.7       Epsilon:0.05
Step 935/1000   Reward: 913.56  Alpha:0.7       Epsilon:0.05
Step 936/1000   Reward: 909.14  Alpha:0.7       Epsilon:0.05
Step 937/1000   Reward: 857.14  Alpha:0.7       Epsilon:0.05
Step 938/1000   Reward: 785.24  Alpha:0.7       Epsilon:0.05
Step 939/1000   Reward: 984.76  Alpha:0.7       Epsilon:0.05
Step 940/1000   Reward: 943.92  Alpha:0.7       Epsilon:0.05
Step 941/1000   Reward: 684.18  Alpha:0.7       Epsilon:0.05
Step 942/1000   Reward: 1125.94 Alpha:0.7       Epsilon:0.05
Step 943/1000   Reward: 915.28  Alpha:0.7       Epsilon:0.05
Step 944/1000   Reward: 826.06  Alpha:0.7       Epsilon:0.05
Step 945/1000   Reward: 831.46  Alpha:0.7       Epsilon:0.05
Step 946/1000   Reward: 1026.12 Alpha:0.7       Epsilon:0.05
Step 947/1000   Reward: 795.54  Alpha:0.7       Epsilon:0.05
Step 948/1000   Reward: 759.4   Alpha:0.7       Epsilon:0.05
Step 949/1000   Reward: 792.74  Alpha:0.7       Epsilon:0.05
Step 950/1000   Reward: 1092.66 Alpha:0.7       Epsilon:0.05
Step 951/1000   Reward: 750.68  Alpha:0.7       Epsilon:0.05
Step 952/1000   Reward: 1012.1  Alpha:0.7       Epsilon:0.05
Step 953/1000   Reward: 946.0   Alpha:0.7       Epsilon:0.05
Step 954/1000   Reward: 846.92  Alpha:0.7       Epsilon:0.05
Step 955/1000   Reward: 750.98  Alpha:0.7       Epsilon:0.05
Step 956/1000   Reward: 748.22  Alpha:0.7       Epsilon:0.05
Step 957/1000   Reward: 889.92  Alpha:0.7       Epsilon:0.05
Step 958/1000   Reward: 857.3   Alpha:0.7       Epsilon:0.05
Step 959/1000   Reward: 940.38  Alpha:0.7       Epsilon:0.05
Step 960/1000   Reward: 938.7   Alpha:0.7       Epsilon:0.05
Step 961/1000   Reward: 897.14  Alpha:0.7       Epsilon:0.05
Step 962/1000   Reward: 901.3   Alpha:0.7       Epsilon:0.05
Step 963/1000   Reward: 934.04  Alpha:0.7       Epsilon:0.05
Step 964/1000   Reward: 917.92  Alpha:0.7       Epsilon:0.05
Step 965/1000   Reward: 1028.22 Alpha:0.7       Epsilon:0.05
Step 966/1000   Reward: 796.66  Alpha:0.7       Epsilon:0.05
Step 967/1000   Reward: 923.78  Alpha:0.7       Epsilon:0.05
Step 968/1000   Reward: 744.02  Alpha:0.7       Epsilon:0.05
Step 969/1000   Reward: 1147.52 Alpha:0.7       Epsilon:0.05
Step 970/1000   Reward: 1340.16 Alpha:0.7       Epsilon:0.05
Step 971/1000   Reward: 951.22  Alpha:0.7       Epsilon:0.05
Step 972/1000   Reward: 1005.02 Alpha:0.7       Epsilon:0.05
Step 973/1000   Reward: 1030.88 Alpha:0.7       Epsilon:0.05
Step 974/1000   Reward: 1120.4  Alpha:0.7       Epsilon:0.05
Step 975/1000   Reward: 1078.32 Alpha:0.7       Epsilon:0.05
Step 976/1000   Reward: 893.5   Alpha:0.7       Epsilon:0.05
Step 977/1000   Reward: 883.24  Alpha:0.7       Epsilon:0.05
Step 978/1000   Reward: 918.92  Alpha:0.7       Epsilon:0.05
Step 979/1000   Reward: 1012.5  Alpha:0.7       Epsilon:0.05
Step 980/1000   Reward: 984.86  Alpha:0.7       Epsilon:0.05
```

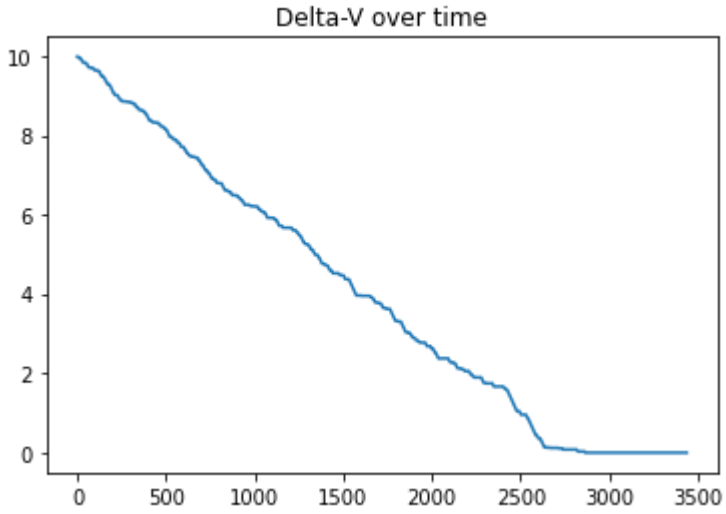
```
Step 981/1000   Reward: 1020.74 Alpha:0.7       Epsilon:0.05
Step 982/1000   Reward: 1062.12 Alpha:0.7       Epsilon:0.05
Step 983/1000   Reward: 1016.82 Alpha:0.7       Epsilon:0.05
Step 984/1000   Reward: 1088.04 Alpha:0.7       Epsilon:0.05
Step 985/1000   Reward: 1269.62 Alpha:0.7       Epsilon:0.05
Step 986/1000   Reward: 822.86  Alpha:0.7       Epsilon:0.05
Step 987/1000   Reward: 808.5   Alpha:0.7       Epsilon:0.05
Step 988/1000   Reward: 997.2   Alpha:0.7       Epsilon:0.05
Step 989/1000   Reward: 640.8   Alpha:0.7       Epsilon:0.05
Step 990/1000   Reward: 973.2   Alpha:0.7       Epsilon:0.05
Step 991/1000   Reward: 907.4   Alpha:0.7       Epsilon:0.05
Step 992/1000   Reward: 867.68  Alpha:0.7       Epsilon:0.05
Step 993/1000   Reward: 1010.88 Alpha:0.7       Epsilon:0.05
Step 994/1000   Reward: 993.24  Alpha:0.7       Epsilon:0.05
Step 995/1000   Reward: 892.86  Alpha:0.7       Epsilon:0.05
Step 996/1000   Reward: 915.38  Alpha:0.7       Epsilon:0.05
Step 997/1000   Reward: 1016.1  Alpha:0.7       Epsilon:0.05
Step 998/1000   Reward: 1124.86 Alpha:0.7       Epsilon:0.05
```

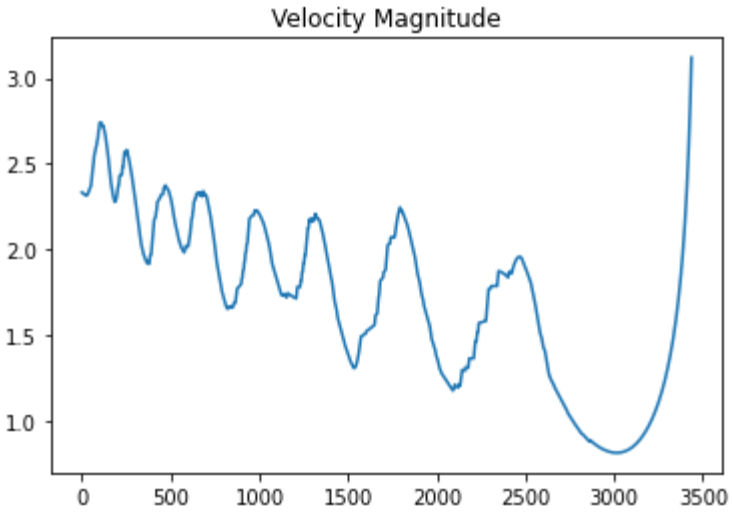
In [652…

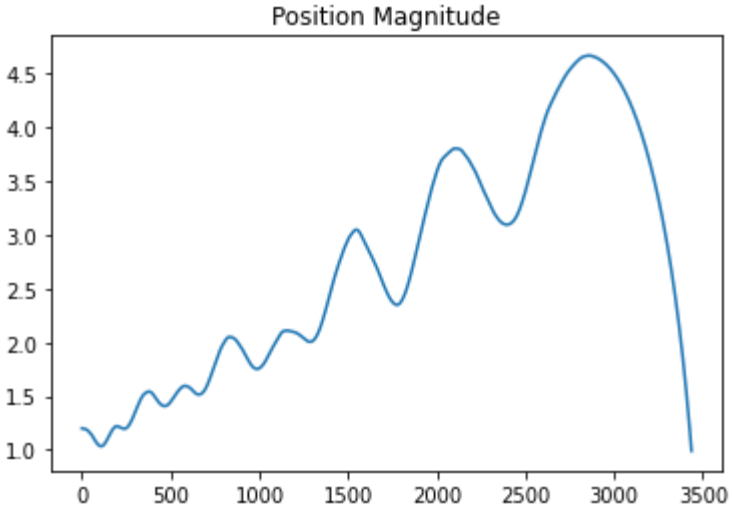
```
plot_policy(ql_Q_98_50000)
```

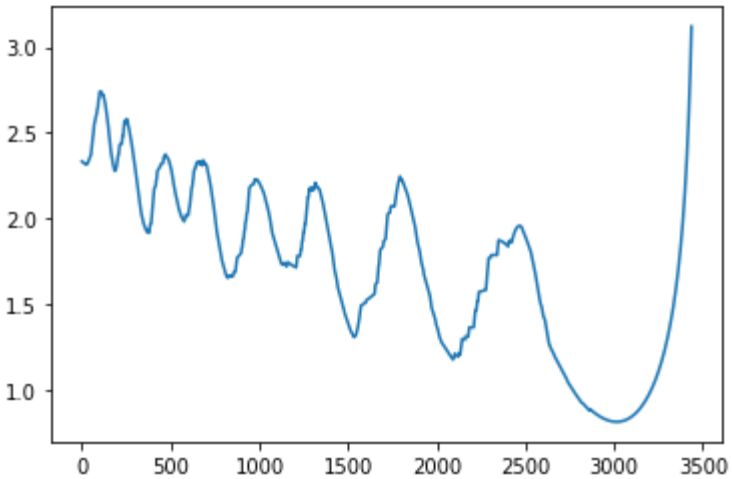
```
2780.0
1.1997688888303473
```





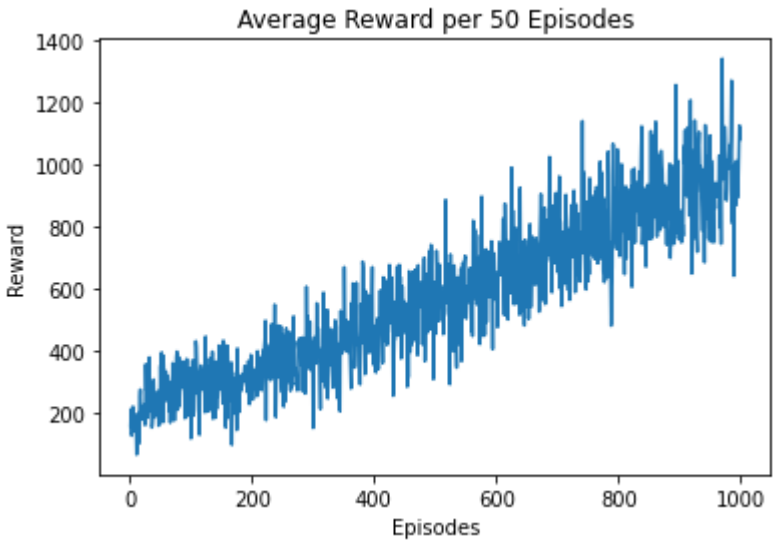

Velocity Magnitude
3.0
2.5
2.0
1.5
1.0
0
500
1000
1500
2000
2500
3000
3500


Position Magnitude
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0
500
1000
1500
2000
2500
3000
3500


3.0
2.5
2.0
1.5
1.0
0
500
1000
1500
2000
2500
3000
3500

In [670…

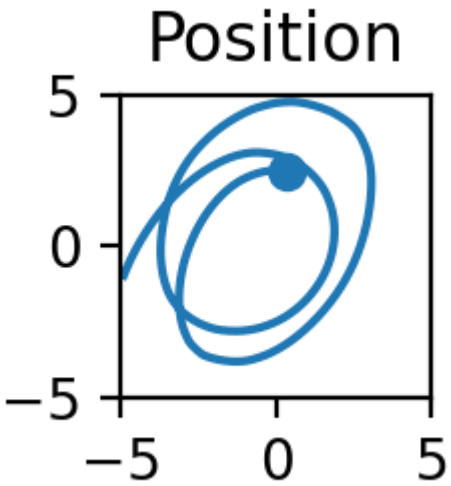
```
def plot_result():
    plt.plot(np.arange(1, len(qlearning_result)), qlearning_result[1:])
    plt.title("Average Reward per 50 Episodes")
    plt.ylabel("Reward")
    plt.xlabel("Episodes")
    plt.show()

plot_result()
```

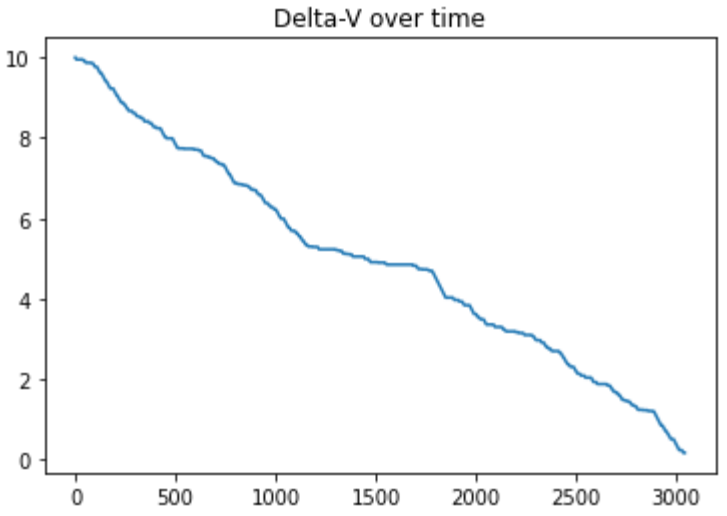


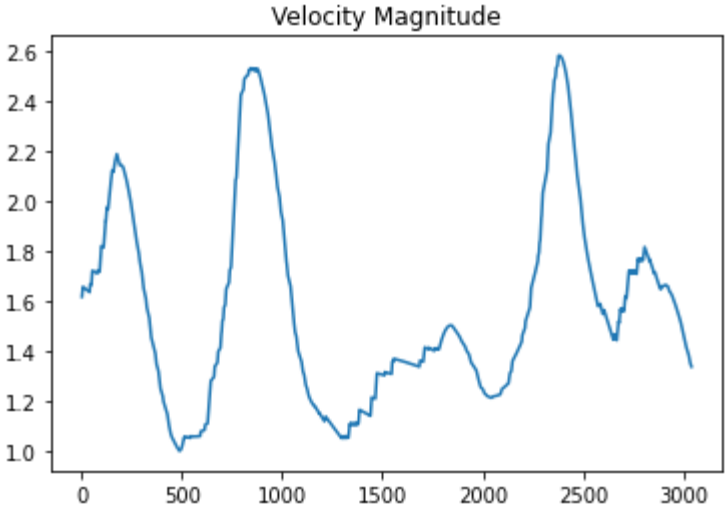
In [668…

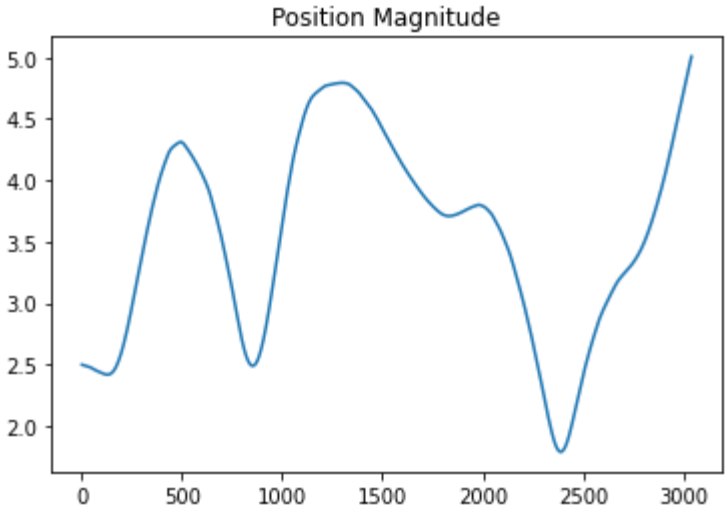
```
plot_policy(ql_Q_98_50000)
```

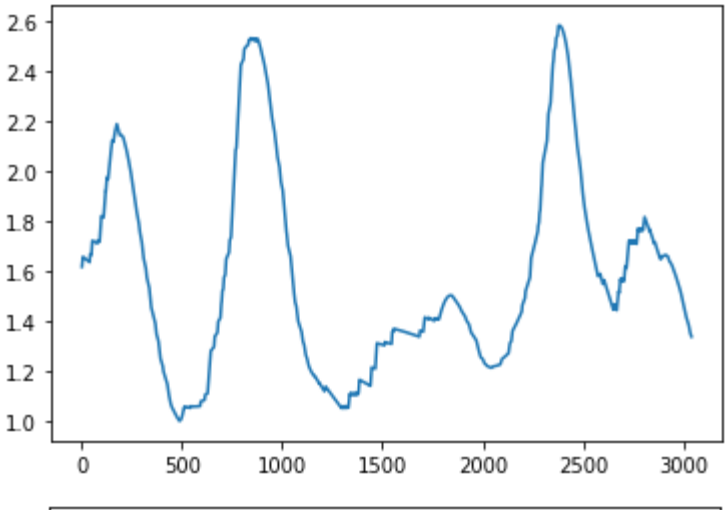
```
2385.0
2.4999475816055776
```




Delta-V over time
10
8
6
4
2
0
0
500
1000
1500
2000
2500
3000


Velocity Magnitude
2.6
2.4
2.2
2.0
1.8
1.6
1.4
1.2
1.0
0
500
1000
1500
2000
2500
3000


Position Magnitude
5.0
4.5
4.0
3.5
3.0
2.5
2.0
0
500
1000
1500
2000
2500
3000



In [ ]: