

MP5

December 1, 2020

1 Deep Q-Learning

Install dependencies for AI gym to run properly (shouldn't take more than a minute). If running on google cloud or running locally, only need to run once. Colab may require installing everytime the vm shuts down.

```
[!]: !pip3 install gym pyvirtualdisplay
      !sudo apt-get install -y xvfb python-opengl ffmpeg

[!]: !pip3 install --upgrade setuptools
      !pip3 install ez_setup
      !pip3 install gym[atari]
```

For this assignment we will implement the Deep Q-Learning algorithm with Experience Replay as described in breakthrough paper **"Playing Atari with Deep Reinforcement Learning"**. We will train an agent to play the famous game of **Breakout**.

```
[!]: %matplotlib inline

import sys
import gym
import torch
import pylab
import random
import numpy as np
from collections import deque
from datetime import datetime
from copy import deepcopy
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
from torch.autograd import Variable
from utils import find_max_lives, check_live, get_frame, get_init_state
from model import DQN
from config import *
import time

import matplotlib.pyplot as plt
%load_ext autoreload
%autoreload 2
```

The autoreload extension is already loaded. To reload it, use:

```
%reload_ext autoreload
```

1.1 Understanding the environment

In the following cell, we initialize our game of **Breakout** and you can see how the environment looks like. For further documentation of the of the environment refer to <https://gym.openai.com/envs>.

In breakout, we will use 3 actions “fire”, “left”, and “right”. “fire” is only used to reset the game when a life is lost, “left” moves the agent left and “right” moves the agent right.

```
[ ]: env = gym.make('BreakoutDeterministic-v4')
      state = env.reset()

[ ]: number_lives = find_max_lives(env)
      state_size = env.observation_space.shape
      action_size = 3 #fire, left, and right
```

1.2 Creating a DQN Agent

Here we create a DQN Agent. This agent is defined in the **agent.py**. The corresponding neural network is defined in the **model.py**. Once you’ve created a working DQN agent, use the code in agent.py to create a double DQN agent in **agent_double.py**. Set the flag “double_dqn” to True to train the double DQN agent.

Evaluation Reward : The average reward received in the past 100 episodes/games.

Frame : Number of frames processed in total.

Memory Size : The current size of the replay memory.

```
[ ]: double_dqn = False # set to True if using double DQN agent

if double_dqn:
    from agent_double import Agent
else:
    from agent import Agent

agent = Agent(action_size)
evaluation_reward = deque(maxlen=evaluation_reward_length)
frame = 0
memory_size = 0
```

1.2.1 Main Training Loop

In this training loop, we do not render the screen because it slows down training significantly. To watch the agent play the game, run the code in next section “Visualize Agent Performance”

```
[ ]: ### single DQN

rewards, episodes = [], []
best_eval_reward = 0
start=time.time()
```

```

for e in range(EPISODES):
    done = False
    score = 0

    history = np.zeros([5, 84, 84], dtype=np.uint8)
    step = 0
    d = False
    state = env.reset()
    next_state = state
    life = number_lives

    get_init_state(history, state)

    while not done:
        step += 1
        frame += 1

        # Perform a fire action if ball is no longer on screen to continue onto
        →next life
        if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
            action = 0
        else:
            action = agent.get_action(np.float32(history[None,:4, :, :]) / 255.)
        state = next_state
        next_state, reward, done, info = env.step(action + 1)

        frame_next_state = get_frame(next_state)
        history[4, :, :] = frame_next_state
        terminal_state = check_live(life, info['ale.lives'])

        life = info['ale.lives']
        r = np.clip(reward, -1, 1)
        r = reward

        # Store the transition in memory
        agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
        # Start training after random sample generation
        if (frame >= train_frame):
            agent.train_policy_net(frame)
            # Update the target network only for Double DQN only
            if double_dqn and (frame % update_target_network_frequency) == 0:
                agent.update_target_net()
            score += reward
            history[:4, :, :] = history[1:, :, :]

        if done:
            evaluation_reward.append(score)

```

```

rewards.append(np.mean(evaluation_reward))
episodes.append(e)
pylab.plot(episodes, rewards, 'b')
pylab.xlabel('Episodes')
pylab.ylabel('Rewards')
pylab.title('Episodes vs Reward')
pylab.savefig("./save_graph/breakout_dqn.png") # save graph for
→training visualization

# every episode, plot the play time
print("episode:", e, " score:", score, " memory length:",
      len(agent.memory), " epsilon:", agent.epsilon, " steps:",
→step,
      " lr:", agent.optimizer.param_groups[0]['lr'], "
→evaluation reward:", np.mean(evaluation_reward))

# if the mean of scores of last 100 episode is bigger than 5 save
→model

### Change this save condition to whatever you prefer ###
if np.mean(evaluation_reward) > 5 and np.mean(evaluation_reward) >
→best_eval_reward:
    torch.save(agent.policy_net, "./save_model/breakout_dqn.pth")
    best_eval_reward = np.mean(evaluation_reward)

```

```

episode: 0  score: 1.0  memory length: 150  epsilon: 1.0  steps: 150  lr:
0.0001  evaluation reward: 1.0
episode: 1  score: 1.0  memory length: 320  epsilon: 1.0  steps: 170  lr:
0.0001  evaluation reward: 1.0
episode: 2  score: 2.0  memory length: 539  epsilon: 1.0  steps: 219  lr:
0.0001  evaluation reward: 1.3333333333333333
episode: 3  score: 3.0  memory length: 783  epsilon: 1.0  steps: 244  lr:
0.0001  evaluation reward: 1.75
episode: 4  score: 2.0  memory length: 965  epsilon: 1.0  steps: 182  lr:
0.0001  evaluation reward: 1.8
episode: 5  score: 1.0  memory length: 1136  epsilon: 1.0  steps: 171
lr: 0.0001  evaluation reward: 1.6666666666666667
episode: 6  score: 0.0  memory length: 1258  epsilon: 1.0  steps: 122
lr: 0.0001  evaluation reward: 1.4285714285714286
episode: 7  score: 1.0  memory length: 1428  epsilon: 1.0  steps: 170
lr: 0.0001  evaluation reward: 1.375
episode: 8  score: 2.0  memory length: 1607  epsilon: 1.0  steps: 179
lr: 0.0001  evaluation reward: 1.4444444444444444
episode: 9  score: 0.0  memory length: 1729  epsilon: 1.0  steps: 122
lr: 0.0001  evaluation reward: 1.3
episode: 10  score: 5.0  memory length: 2067  epsilon: 1.0  steps: 338
lr: 0.0001  evaluation reward: 1.6363636363636365
episode: 11  score: 1.0  memory length: 2236  epsilon: 1.0  steps: 169

```

lr: 0.0001	evaluation reward: 1.5833333333333333		
episode: 12	score: 0.0 memory length: 2359 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.4615384615384615		
episode: 13	score: 1.0 memory length: 2528 epsilon: 1.0	steps: 169	
lr: 0.0001	evaluation reward: 1.4285714285714286		
episode: 14	score: 3.0 memory length: 2759 epsilon: 1.0	steps: 231	
lr: 0.0001	evaluation reward: 1.5333333333333334		
episode: 15	score: 0.0 memory length: 2882 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.4375		
episode: 16	score: 0.0 memory length: 3005 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.3529411764705883		
episode: 17	score: 1.0 memory length: 3156 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.3333333333333333		
episode: 18	score: 2.0 memory length: 3353 epsilon: 1.0	steps: 197	
lr: 0.0001	evaluation reward: 1.368421052631579		
episode: 19	score: 3.0 memory length: 3603 epsilon: 1.0	steps: 250	
lr: 0.0001	evaluation reward: 1.45		
episode: 20	score: 0.0 memory length: 3726 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.380952380952381		
episode: 21	score: 2.0 memory length: 3908 epsilon: 1.0	steps: 182	
lr: 0.0001	evaluation reward: 1.4090909090909092		
episode: 22	score: 2.0 memory length: 4106 epsilon: 1.0	steps: 198	
lr: 0.0001	evaluation reward: 1.434782608695652		
episode: 23	score: 1.0 memory length: 4257 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.4166666666666667		
episode: 24	score: 1.0 memory length: 4408 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.4		
episode: 25	score: 3.0 memory length: 4634 epsilon: 1.0	steps: 226	
lr: 0.0001	evaluation reward: 1.4615384615384615		
episode: 26	score: 0.0 memory length: 4757 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.4074074074074074		
episode: 27	score: 1.0 memory length: 4926 epsilon: 1.0	steps: 169	
lr: 0.0001	evaluation reward: 1.3928571428571428		
episode: 28	score: 0.0 memory length: 5049 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.3448275862068966		
episode: 29	score: 2.0 memory length: 5247 epsilon: 1.0	steps: 198	
lr: 0.0001	evaluation reward: 1.3666666666666667		
episode: 30	score: 2.0 memory length: 5426 epsilon: 1.0	steps: 179	
lr: 0.0001	evaluation reward: 1.3870967741935485		
episode: 31	score: 2.0 memory length: 5645 epsilon: 1.0	steps: 219	
lr: 0.0001	evaluation reward: 1.40625		
episode: 32	score: 2.0 memory length: 5863 epsilon: 1.0	steps: 218	
lr: 0.0001	evaluation reward: 1.4242424242424243		
episode: 33	score: 0.0 memory length: 5986 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.3823529411764706		
episode: 34	score: 1.0 memory length: 6137 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.3714285714285714		
episode: 35	score: 0.0 memory length: 6260 epsilon: 1.0	steps: 123	

lr: 0.0001	evaluation reward: 1.3333333333333333		
episode: 36	score: 1.0 memory length: 6411 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.3243243243243243		
episode: 37	score: 1.0 memory length: 6562 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.3157894736842106		
episode: 38	score: 1.0 memory length: 6713 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.3076923076923077		
episode: 39	score: 2.0 memory length: 6931 epsilon: 1.0	steps: 218	
lr: 0.0001	evaluation reward: 1.325		
episode: 40	score: 0.0 memory length: 7053 epsilon: 1.0	steps: 122	
lr: 0.0001	evaluation reward: 1.2926829268292683		
episode: 41	score: 2.0 memory length: 7251 epsilon: 1.0	steps: 198	
lr: 0.0001	evaluation reward: 1.3095238095238095		
episode: 42	score: 1.0 memory length: 7402 epsilon: 1.0	steps: 151	
lr: 0.0001	evaluation reward: 1.302325581395349		
episode: 43	score: 2.0 memory length: 7622 epsilon: 1.0	steps: 220	
lr: 0.0001	evaluation reward: 1.3181818181818181		
episode: 44	score: 2.0 memory length: 7840 epsilon: 1.0	steps: 218	
lr: 0.0001	evaluation reward: 1.3333333333333333		
episode: 45	score: 3.0 memory length: 8085 epsilon: 1.0	steps: 245	
lr: 0.0001	evaluation reward: 1.3695652173913044		
episode: 46	score: 0.0 memory length: 8207 epsilon: 1.0	steps: 122	
lr: 0.0001	evaluation reward: 1.3404255319148937		
episode: 47	score: 0.0 memory length: 8330 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.3125		
episode: 48	score: 1.0 memory length: 8499 epsilon: 1.0	steps: 169	
lr: 0.0001	evaluation reward: 1.3061224489795917		
episode: 49	score: 1.0 memory length: 8649 epsilon: 1.0	steps: 150	
lr: 0.0001	evaluation reward: 1.3		
episode: 50	score: 1.0 memory length: 8819 epsilon: 1.0	steps: 170	
lr: 0.0001	evaluation reward: 1.2941176470588236		
episode: 51	score: 0.0 memory length: 8941 epsilon: 1.0	steps: 122	
lr: 0.0001	evaluation reward: 1.2692307692307692		
episode: 52	score: 1.0 memory length: 9112 epsilon: 1.0	steps: 171	
lr: 0.0001	evaluation reward: 1.2641509433962264		
episode: 53	score: 2.0 memory length: 9312 epsilon: 1.0	steps: 200	
lr: 0.0001	evaluation reward: 1.2777777777777777		
episode: 54	score: 2.0 memory length: 9510 epsilon: 1.0	steps: 198	
lr: 0.0001	evaluation reward: 1.290909090909091		
episode: 55	score: 2.0 memory length: 9728 epsilon: 1.0	steps: 218	
lr: 0.0001	evaluation reward: 1.3035714285714286		
episode: 56	score: 1.0 memory length: 9878 epsilon: 1.0	steps: 150	
lr: 0.0001	evaluation reward: 1.2982456140350878		
episode: 57	score: 0.0 memory length: 10001 epsilon: 1.0	steps: 123	
lr: 0.0001	evaluation reward: 1.2758620689655173		
episode: 58	score: 2.0 memory length: 10218 epsilon: 1.0	steps: 217	
lr: 0.0001	evaluation reward: 1.2881355932203389		
episode: 59	score: 1.0 memory length: 10369 epsilon: 1.0	steps: 151	

lr: 0.0001	evaluation reward: 1.2833333333333334		
episode: 60	score: 0.0 memory length: 10492	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2622950819672132		
episode: 61	score: 3.0 memory length: 10717	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.2903225806451613		
episode: 62	score: 2.0 memory length: 10918	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.3015873015873016		
episode: 63	score: 2.0 memory length: 11098	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.3125		
episode: 64	score: 2.0 memory length: 11296	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.323076923076923		
episode: 65	score: 3.0 memory length: 11563	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.3484848484848484		
episode: 66	score: 0.0 memory length: 11685	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.328358208955224		
episode: 67	score: 2.0 memory length: 11883	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.338235294117647		
episode: 68	score: 3.0 memory length: 12133	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.3623188405797102		
episode: 69	score: 2.0 memory length: 12351	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.3714285714285714		
episode: 70	score: 3.0 memory length: 12576	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.3943661971830985		
episode: 71	score: 0.0 memory length: 12698	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.375		
episode: 72	score: 3.0 memory length: 12944	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.3972602739726028		
episode: 73	score: 0.0 memory length: 13067	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3783783783783783		
episode: 74	score: 0.0 memory length: 13189	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36		
episode: 75	score: 2.0 memory length: 13405	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.368421052631579		
episode: 76	score: 4.0 memory length: 13663	epsilon: 1.0	steps: 258
lr: 0.0001	evaluation reward: 1.4025974025974026		
episode: 77	score: 0.0 memory length: 13785	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3846153846153846		
episode: 78	score: 3.0 memory length: 14010	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.4050632911392404		
episode: 79	score: 2.0 memory length: 14211	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.4125		
episode: 80	score: 1.0 memory length: 14383	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.4074074074074074		
episode: 81	score: 0.0 memory length: 14506	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3902439024390243		
episode: 82	score: 0.0 memory length: 14629	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3734939759036144		
episode: 83	score: 1.0 memory length: 14779	epsilon: 1.0	steps: 150

lr: 0.0001	evaluation reward: 1.369047619047619		
episode: 84	score: 2.0	memory length: 14977	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.3764705882352941		
episode: 85	score: 1.0	memory length: 15128	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.372093023255814		
episode: 86	score: 2.0	memory length: 15346	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.3793103448275863		
episode: 87	score: 0.0	memory length: 15468	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.3636363636363635		
episode: 88	score: 1.0	memory length: 15636	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.3595505617977528		
episode: 89	score: 0.0	memory length: 15758	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.3444444444444446		
episode: 90	score: 1.0	memory length: 15909	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.3406593406593406		
episode: 91	score: 2.0	memory length: 16107	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.3478260869565217		
episode: 92	score: 3.0	memory length: 16333	epsilon: 1.0 steps: 226
lr: 0.0001	evaluation reward: 1.3655913978494623		
episode: 93	score: 2.0	memory length: 16533	epsilon: 1.0 steps: 200
lr: 0.0001	evaluation reward: 1.372340425531915		
episode: 94	score: 0.0	memory length: 16656	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.3578947368421053		
episode: 95	score: 1.0	memory length: 16825	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.3541666666666667		
episode: 96	score: 0.0	memory length: 16948	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.3402061855670102		
episode: 97	score: 1.0	memory length: 17099	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.336734693877551		
episode: 98	score: 1.0	memory length: 17269	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.3333333333333333		
episode: 99	score: 3.0	memory length: 17520	epsilon: 1.0 steps: 251
lr: 0.0001	evaluation reward: 1.35		
episode: 100	score: 1.0	memory length: 17691	epsilon: 1.0 steps: 171
lr: 0.0001	evaluation reward: 1.35		
episode: 101	score: 0.0	memory length: 17813	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.34		
episode: 102	score: 0.0	memory length: 17936	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 103	score: 1.0	memory length: 18104	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.3		
episode: 104	score: 3.0	memory length: 18331	epsilon: 1.0 steps: 227
lr: 0.0001	evaluation reward: 1.31		
episode: 105	score: 2.0	memory length: 18531	epsilon: 1.0 steps: 200
lr: 0.0001	evaluation reward: 1.32		
episode: 106	score: 3.0	memory length: 18779	epsilon: 1.0 steps: 248
lr: 0.0001	evaluation reward: 1.35		
episode: 107	score: 2.0	memory length: 18998	epsilon: 1.0 steps: 219

lr: 0.0001	evaluation reward: 1.36		
episode: 108	score: 3.0 memory length: 19266	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.37		
episode: 109	score: 3.0 memory length: 19512	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.4		
episode: 110	score: 3.0 memory length: 19740	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.38		
episode: 111	score: 3.0 memory length: 19984	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.4		
episode: 112	score: 0.0 memory length: 20107	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 113	score: 0.0 memory length: 20230	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 114	score: 0.0 memory length: 20353	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36		
episode: 115	score: 4.0 memory length: 20610	epsilon: 1.0	steps: 257
lr: 0.0001	evaluation reward: 1.4		
episode: 116	score: 0.0 memory length: 20732	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4		
episode: 117	score: 1.0 memory length: 20903	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.4		
episode: 118	score: 4.0 memory length: 21198	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.42		
episode: 119	score: 3.0 memory length: 21426	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.42		
episode: 120	score: 1.0 memory length: 21595	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.43		
episode: 121	score: 0.0 memory length: 21718	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 122	score: 2.0 memory length: 21936	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.41		
episode: 123	score: 1.0 memory length: 22105	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.41		
episode: 124	score: 0.0 memory length: 22228	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 125	score: 0.0 memory length: 22350	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37		
episode: 126	score: 3.0 memory length: 22596	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.4		
episode: 127	score: 0.0 memory length: 22719	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 128	score: 2.0 memory length: 22939	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.41		
episode: 129	score: 1.0 memory length: 23089	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4		
episode: 130	score: 2.0 memory length: 23287	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4		
episode: 131	score: 1.0 memory length: 23455	epsilon: 1.0	steps: 168

lr: 0.0001	evaluation reward: 1.39			
episode: 132	score: 0.0	memory length: 23577	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37			
episode: 133	score: 0.0	memory length: 23700	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 134	score: 2.0	memory length: 23919	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.38			
episode: 135	score: 0.0	memory length: 24041	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38			
episode: 136	score: 1.0	memory length: 24191	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.38			
episode: 137	score: 3.0	memory length: 24419	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.4			
episode: 138	score: 0.0	memory length: 24542	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39			
episode: 139	score: 0.0	memory length: 24665	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 140	score: 3.0	memory length: 24912	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.4			
episode: 141	score: 1.0	memory length: 25063	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.39			
episode: 142	score: 2.0	memory length: 25279	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.4			
episode: 143	score: 0.0	memory length: 25402	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38			
episode: 144	score: 3.0	memory length: 25669	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.39			
episode: 145	score: 0.0	memory length: 25791	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36			
episode: 146	score: 0.0	memory length: 25914	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36			
episode: 147	score: 2.0	memory length: 26132	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.38			
episode: 148	score: 1.0	memory length: 26284	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.38			
episode: 149	score: 1.0	memory length: 26454	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.38			
episode: 150	score: 0.0	memory length: 26577	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 151	score: 1.0	memory length: 26748	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.38			
episode: 152	score: 2.0	memory length: 26965	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.39			
episode: 153	score: 2.0	memory length: 27181	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.39			
episode: 154	score: 3.0	memory length: 27446	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.4			
episode: 155	score: 3.0	memory length: 27672	epsilon: 1.0	steps: 226

lr: 0.0001	evaluation reward: 1.41			
episode: 156	score: 1.0	memory length: 27844	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.41			
episode: 157	score: 2.0	memory length: 28061	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.43			
episode: 158	score: 1.0	memory length: 28230	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.42			
episode: 159	score: 2.0	memory length: 28428	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.43			
episode: 160	score: 0.0	memory length: 28551	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43			
episode: 161	score: 0.0	memory length: 28674	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 162	score: 3.0	memory length: 28942	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.41			
episode: 163	score: 1.0	memory length: 29093	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.4			
episode: 164	score: 1.0	memory length: 29262	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39			
episode: 165	score: 5.0	memory length: 29608	epsilon: 1.0	steps: 346
lr: 0.0001	evaluation reward: 1.41			
episode: 166	score: 2.0	memory length: 29807	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.43			
episode: 167	score: 2.0	memory length: 30006	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.43			
episode: 168	score: 0.0	memory length: 30129	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 169	score: 2.0	memory length: 30327	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4			
episode: 170	score: 0.0	memory length: 30450	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 171	score: 0.0	memory length: 30573	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 172	score: 1.0	memory length: 30724	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.35			
episode: 173	score: 2.0	memory length: 30921	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.37			
episode: 174	score: 2.0	memory length: 31139	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.39			
episode: 175	score: 2.0	memory length: 31336	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.39			
episode: 176	score: 2.0	memory length: 31534	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.37			
episode: 177	score: 0.0	memory length: 31656	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37			
episode: 178	score: 0.0	memory length: 31778	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.34			
episode: 179	score: 5.0	memory length: 32081	epsilon: 1.0	steps: 303

lr: 0.0001	evaluation reward: 1.37		
episode: 180	score: 2.0 memory length: 32279	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.38		
episode: 181	score: 1.0 memory length: 32448	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39		
episode: 182	score: 2.0 memory length: 32646	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41		
episode: 183	score: 2.0 memory length: 32844	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.42		
episode: 184	score: 3.0 memory length: 33107	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.43		
episode: 185	score: 6.0 memory length: 33479	epsilon: 1.0	steps: 372
lr: 0.0001	evaluation reward: 1.48		
episode: 186	score: 0.0 memory length: 33602	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 187	score: 2.0 memory length: 33783	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.48		
episode: 188	score: 4.0 memory length: 34040	epsilon: 1.0	steps: 257
lr: 0.0001	evaluation reward: 1.51		
episode: 189	score: 3.0 memory length: 34286	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.54		
episode: 190	score: 3.0 memory length: 34514	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.56		
episode: 191	score: 1.0 memory length: 34682	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.55		
episode: 192	score: 3.0 memory length: 34948	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.55		
episode: 193	score: 4.0 memory length: 35244	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.57		
episode: 194	score: 4.0 memory length: 35539	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.61		
episode: 195	score: 0.0 memory length: 35661	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.6		
episode: 196	score: 5.0 memory length: 35996	epsilon: 1.0	steps: 335
lr: 0.0001	evaluation reward: 1.65		
episode: 197	score: 1.0 memory length: 36164	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.65		
episode: 198	score: 5.0 memory length: 36471	epsilon: 1.0	steps: 307
lr: 0.0001	evaluation reward: 1.69		
episode: 199	score: 2.0 memory length: 36669	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.68		
episode: 200	score: 0.0 memory length: 36791	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.67		
episode: 201	score: 0.0 memory length: 36914	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67		
episode: 202	score: 0.0 memory length: 37037	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67		
episode: 203	score: 2.0 memory length: 37252	epsilon: 1.0	steps: 215

lr: 0.0001	evaluation reward: 1.68		
episode: 204	score: 0.0 memory length: 37375	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65		
episode: 205	score: 2.0 memory length: 37593	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.65		
episode: 206	score: 1.0 memory length: 37744	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.63		
episode: 207	score: 3.0 memory length: 37991	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.64		
episode: 208	score: 1.0 memory length: 38141	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.62		
episode: 209	score: 1.0 memory length: 38310	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.6		
episode: 210	score: 1.0 memory length: 38479	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.58		
episode: 211	score: 3.0 memory length: 38727	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.58		
episode: 212	score: 0.0 memory length: 38849	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.58		
episode: 213	score: 0.0 memory length: 38971	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.58		
episode: 214	score: 3.0 memory length: 39199	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.61		
episode: 215	score: 1.0 memory length: 39368	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.58		
episode: 216	score: 0.0 memory length: 39491	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58		
episode: 217	score: 2.0 memory length: 39709	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.59		
episode: 218	score: 2.0 memory length: 39907	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.57		
episode: 219	score: 2.0 memory length: 40087	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.56		
episode: 220	score: 2.0 memory length: 40284	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.57		
episode: 221	score: 1.0 memory length: 40455	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.58		
episode: 222	score: 4.0 memory length: 40714	epsilon: 1.0	steps: 259
lr: 0.0001	evaluation reward: 1.6		
episode: 223	score: 5.0 memory length: 41056	epsilon: 1.0	steps: 342
lr: 0.0001	evaluation reward: 1.64		
episode: 224	score: 0.0 memory length: 41179	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64		
episode: 225	score: 0.0 memory length: 41302	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64		
episode: 226	score: 2.0 memory length: 41500	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.63		
episode: 227	score: 1.0 memory length: 41668	epsilon: 1.0	steps: 168

lr: 0.0001	evaluation reward: 1.64		
episode: 228	score: 1.0 memory length: 41837	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.63		
episode: 229	score: 4.0 memory length: 42088	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.66		
episode: 230	score: 2.0 memory length: 42305	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.66		
episode: 231	score: 0.0 memory length: 42428	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65		
episode: 232	score: 0.0 memory length: 42550	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.65		
episode: 233	score: 2.0 memory length: 42748	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.67		
episode: 234	score: 1.0 memory length: 42917	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.66		
episode: 235	score: 1.0 memory length: 43086	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.67		
episode: 236	score: 0.0 memory length: 43209	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66		
episode: 237	score: 3.0 memory length: 43476	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.66		
episode: 238	score: 1.0 memory length: 43645	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.67		
episode: 239	score: 5.0 memory length: 43973	epsilon: 1.0	steps: 328
lr: 0.0001	evaluation reward: 1.72		
episode: 240	score: 2.0 memory length: 44171	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.71		
episode: 241	score: 2.0 memory length: 44368	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.72		
episode: 242	score: 0.0 memory length: 44491	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 243	score: 0.0 memory length: 44614	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 244	score: 2.0 memory length: 44812	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.69		
episode: 245	score: 2.0 memory length: 44992	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.71		
episode: 246	score: 0.0 memory length: 45115	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.71		
episode: 247	score: 2.0 memory length: 45296	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.71		
episode: 248	score: 0.0 memory length: 45419	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 249	score: 0.0 memory length: 45542	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.69		
episode: 250	score: 2.0 memory length: 45759	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.71		
episode: 251	score: 3.0 memory length: 45989	epsilon: 1.0	steps: 230

lr: 0.0001	evaluation reward: 1.73		
episode: 252	score: 1.0 memory length: 46140	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.72		
episode: 253	score: 0.0 memory length: 46263	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 254	score: 4.0 memory length: 46543	epsilon: 1.0	steps: 280
lr: 0.0001	evaluation reward: 1.71		
episode: 255	score: 4.0 memory length: 46817	epsilon: 1.0	steps: 274
lr: 0.0001	evaluation reward: 1.72		
episode: 256	score: 2.0 memory length: 47015	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.73		
episode: 257	score: 2.0 memory length: 47213	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.73		
episode: 258	score: 0.0 memory length: 47336	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.72		
episode: 259	score: 2.0 memory length: 47534	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.72		
episode: 260	score: 2.0 memory length: 47736	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.74		
episode: 261	score: 0.0 memory length: 47858	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.74		
episode: 262	score: 3.0 memory length: 48086	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.74		
episode: 263	score: 0.0 memory length: 48209	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.73		
episode: 264	score: 1.0 memory length: 48360	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.73		
episode: 265	score: 0.0 memory length: 48483	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.68		
episode: 266	score: 3.0 memory length: 48752	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.69		
episode: 267	score: 2.0 memory length: 48951	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.69		
episode: 268	score: 2.0 memory length: 49149	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.71		
episode: 269	score: 1.0 memory length: 49318	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.7		
episode: 270	score: 3.0 memory length: 49565	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.73		
episode: 271	score: 4.0 memory length: 49837	epsilon: 1.0	steps: 272
lr: 0.0001	evaluation reward: 1.77		
episode: 272	score: 2.0 memory length: 50053	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.78		
episode: 273	score: 2.0 memory length: 50271	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.78		
episode: 274	score: 0.0 memory length: 50394	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.76		
episode: 275	score: 2.0 memory length: 50613	epsilon: 1.0	steps: 219

lr: 0.0001	evaluation reward: 1.76		
episode: 276	score: 1.0 memory length: 50781	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.75		
episode: 277	score: 0.0 memory length: 50904	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.75		
episode: 278	score: 1.0 memory length: 51072	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.76		
episode: 279	score: 3.0 memory length: 51297	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.74		
episode: 280	score: 0.0 memory length: 51420	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.72		
episode: 281	score: 4.0 memory length: 51713	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.75		
episode: 282	score: 2.0 memory length: 51911	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.75		
episode: 283	score: 1.0 memory length: 52079	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.74		
episode: 284	score: 0.0 memory length: 52201	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.71		
episode: 285	score: 2.0 memory length: 52401	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.67		
episode: 286	score: 1.0 memory length: 52571	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.68		
episode: 287	score: 1.0 memory length: 52721	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.67		
episode: 288	score: 1.0 memory length: 52892	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.64		
episode: 289	score: 3.0 memory length: 53138	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.64		
episode: 290	score: 0.0 memory length: 53260	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.61		
episode: 291	score: 1.0 memory length: 53431	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.61		
episode: 292	score: 1.0 memory length: 53599	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.59		
episode: 293	score: 1.0 memory length: 53750	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.56		
episode: 294	score: 2.0 memory length: 53947	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.54		
episode: 295	score: 1.0 memory length: 54098	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.55		
episode: 296	score: 1.0 memory length: 54248	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51		
episode: 297	score: 0.0 memory length: 54370	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.5		
episode: 298	score: 2.0 memory length: 54569	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.47		
episode: 299	score: 1.0 memory length: 54739	epsilon: 1.0	steps: 170

lr: 0.0001	evaluation reward: 1.46		
episode: 300	score: 1.0 memory length: 54908	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.47		
episode: 301	score: 5.0 memory length: 55253	epsilon: 1.0	steps: 345
lr: 0.0001	evaluation reward: 1.52		
episode: 302	score: 1.0 memory length: 55422	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.53		
episode: 303	score: 0.0 memory length: 55545	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 304	score: 0.0 memory length: 55668	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 305	score: 1.0 memory length: 55836	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.5		
episode: 306	score: 2.0 memory length: 56056	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.51		
episode: 307	score: 1.0 memory length: 56224	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.49		
episode: 308	score: 0.0 memory length: 56347	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 309	score: 5.0 memory length: 56651	epsilon: 1.0	steps: 304
lr: 0.0001	evaluation reward: 1.52		
episode: 310	score: 0.0 memory length: 56774	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 311	score: 2.0 memory length: 56971	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.5		
episode: 312	score: 1.0 memory length: 57142	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.51		
episode: 313	score: 1.0 memory length: 57310	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.52		
episode: 314	score: 1.0 memory length: 57461	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.5		
episode: 315	score: 0.0 memory length: 57584	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 316	score: 0.0 memory length: 57707	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 317	score: 2.0 memory length: 57925	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.49		
episode: 318	score: 1.0 memory length: 58076	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48		
episode: 319	score: 3.0 memory length: 58301	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.49		
episode: 320	score: 2.0 memory length: 58481	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.49		
episode: 321	score: 2.0 memory length: 58698	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.5		
episode: 322	score: 0.0 memory length: 58820	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.46		
episode: 323	score: 0.0 memory length: 58942	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.41		
episode: 324	score: 1.0 memory length: 59111	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.42		
episode: 325	score: 1.0 memory length: 59262	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.43		
episode: 326	score: 2.0 memory length: 59460	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.43		
episode: 327	score: 0.0 memory length: 59583	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 328	score: 0.0 memory length: 59706	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 329	score: 4.0 memory length: 60002	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.41		
episode: 330	score: 0.0 memory length: 60124	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39		
episode: 331	score: 0.0 memory length: 60246	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39		
episode: 332	score: 2.0 memory length: 60444	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41		
episode: 333	score: 2.0 memory length: 60644	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.41		
episode: 334	score: 2.0 memory length: 60841	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.42		
episode: 335	score: 3.0 memory length: 61054	epsilon: 1.0	steps: 213
lr: 0.0001	evaluation reward: 1.44		
episode: 336	score: 2.0 memory length: 61271	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.46		
episode: 337	score: 4.0 memory length: 61564	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.47		
episode: 338	score: 0.0 memory length: 61687	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 339	score: 0.0 memory length: 61809	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41		
episode: 340	score: 0.0 memory length: 61932	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 341	score: 5.0 memory length: 62294	epsilon: 1.0	steps: 362
lr: 0.0001	evaluation reward: 1.42		
episode: 342	score: 1.0 memory length: 62465	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.43		
episode: 343	score: 2.0 memory length: 62662	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.45		
episode: 344	score: 0.0 memory length: 62784	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43		
episode: 345	score: 3.0 memory length: 63050	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.44		
episode: 346	score: 0.0 memory length: 63173	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 347	score: 2.0 memory length: 63371	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.44		
episode: 348	score: 1.0 memory length: 63540	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.45		
episode: 349	score: 2.0 memory length: 63738	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 350	score: 2.0 memory length: 63938	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.47		
episode: 351	score: 1.0 memory length: 64109	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.45		
episode: 352	score: 1.0 memory length: 64260	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.45		
episode: 353	score: 0.0 memory length: 64382	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45		
episode: 354	score: 0.0 memory length: 64505	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 355	score: 1.0 memory length: 64677	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.38		
episode: 356	score: 1.0 memory length: 64828	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.37		
episode: 357	score: 1.0 memory length: 64978	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.36		
episode: 358	score: 5.0 memory length: 65296	epsilon: 1.0	steps: 318
lr: 0.0001	evaluation reward: 1.41		
episode: 359	score: 2.0 memory length: 65493	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.41		
episode: 360	score: 2.0 memory length: 65691	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41		
episode: 361	score: 2.0 memory length: 65870	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.43		
episode: 362	score: 1.0 memory length: 66020	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.41		
episode: 363	score: 0.0 memory length: 66143	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 364	score: 0.0 memory length: 66265	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4		
episode: 365	score: 1.0 memory length: 66415	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.41		
episode: 366	score: 0.0 memory length: 66538	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38		
episode: 367	score: 0.0 memory length: 66660	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36		
episode: 368	score: 3.0 memory length: 66908	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.37		
episode: 369	score: 0.0 memory length: 67030	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36		
episode: 370	score: 2.0 memory length: 67227	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.35		
episode: 371	score: 0.0 memory length: 67350	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.31		
episode: 372	score: 2.0 memory length: 67548	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.31		
episode: 373	score: 3.0 memory length: 67795	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.32		
episode: 374	score: 0.0 memory length: 67918	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 375	score: 0.0 memory length: 68041	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3		
episode: 376	score: 3.0 memory length: 68290	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.32		
episode: 377	score: 2.0 memory length: 68507	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.34		
episode: 378	score: 2.0 memory length: 68708	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.35		
episode: 379	score: 1.0 memory length: 68878	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.33		
episode: 380	score: 2.0 memory length: 69075	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.35		
episode: 381	score: 3.0 memory length: 69319	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.34		
episode: 382	score: 0.0 memory length: 69442	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 383	score: 0.0 memory length: 69564	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.31		
episode: 384	score: 1.0 memory length: 69734	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.32		
episode: 385	score: 0.0 memory length: 69856	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3		
episode: 386	score: 2.0 memory length: 70053	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.31		
episode: 387	score: 1.0 memory length: 70204	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.31		
episode: 388	score: 3.0 memory length: 70468	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.33		
episode: 389	score: 2.0 memory length: 70649	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.32		
episode: 390	score: 3.0 memory length: 70897	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.35		
episode: 391	score: 0.0 memory length: 71020	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.34		
episode: 392	score: 2.0 memory length: 71218	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.35		
episode: 393	score: 2.0 memory length: 71416	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.36		
episode: 394	score: 1.0 memory length: 71584	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.35		
episode: 395	score: 0.0 memory length: 71707	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.34		
episode: 396	score: 1.0 memory length: 71875	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.34		
episode: 397	score: 1.0 memory length: 72045	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.35		
episode: 398	score: 1.0 memory length: 72195	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.34		
episode: 399	score: 1.0 memory length: 72346	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.34		
episode: 400	score: 4.0 memory length: 72645	epsilon: 1.0	steps: 299
lr: 0.0001	evaluation reward: 1.37		
episode: 401	score: 0.0 memory length: 72768	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 402	score: 2.0 memory length: 72948	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.33		
episode: 403	score: 1.0 memory length: 73119	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.34		
episode: 404	score: 2.0 memory length: 73338	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.36		
episode: 405	score: 0.0 memory length: 73461	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 406	score: 0.0 memory length: 73583	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.33		
episode: 407	score: 0.0 memory length: 73706	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 408	score: 0.0 memory length: 73828	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.32		
episode: 409	score: 0.0 memory length: 73951	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.27		
episode: 410	score: 4.0 memory length: 74225	epsilon: 1.0	steps: 274
lr: 0.0001	evaluation reward: 1.31		
episode: 411	score: 3.0 memory length: 74486	epsilon: 1.0	steps: 261
lr: 0.0001	evaluation reward: 1.32		
episode: 412	score: 1.0 memory length: 74637	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.32		
episode: 413	score: 1.0 memory length: 74806	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.32		
episode: 414	score: 3.0 memory length: 75068	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.34		
episode: 415	score: 2.0 memory length: 75266	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.36		
episode: 416	score: 3.0 memory length: 75536	epsilon: 1.0	steps: 270
lr: 0.0001	evaluation reward: 1.39		
episode: 417	score: 3.0 memory length: 75779	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.4		
episode: 418	score: 0.0 memory length: 75902	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 419	score: 1.0 memory length: 76071	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.37		
episode: 420	score: 0.0 memory length: 76194	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 421	score: 2.0 memory length: 76391	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.35		
episode: 422	score: 0.0 memory length: 76514	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 423	score: 0.0 memory length: 76637	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 424	score: 2.0 memory length: 76834	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.36		
episode: 425	score: 2.0 memory length: 77031	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.37		
episode: 426	score: 0.0 memory length: 77154	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 427	score: 1.0 memory length: 77323	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.36		
episode: 428	score: 0.0 memory length: 77445	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36		
episode: 429	score: 0.0 memory length: 77568	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 430	score: 4.0 memory length: 77843	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.36		
episode: 431	score: 0.0 memory length: 77966	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36		
episode: 432	score: 1.0 memory length: 78135	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.35		
episode: 433	score: 0.0 memory length: 78257	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.33		
episode: 434	score: 1.0 memory length: 78408	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.32		
episode: 435	score: 1.0 memory length: 78559	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3		
episode: 436	score: 0.0 memory length: 78682	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.28		
episode: 437	score: 2.0 memory length: 78880	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.26		
episode: 438	score: 3.0 memory length: 79124	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.29		
episode: 439	score: 3.0 memory length: 79354	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.32		
episode: 440	score: 0.0 memory length: 79476	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.32		
episode: 441	score: 0.0 memory length: 79599	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.27		
episode: 442	score: 0.0 memory length: 79722	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.26		
episode: 443	score: 1.0 memory length: 79873	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.25		
episode: 444	score: 0.0 memory length: 79996	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.25		
episode: 445	score: 0.0 memory length: 80119	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.22		
episode: 446	score: 2.0 memory length: 80316	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.24		
episode: 447	score: 2.0 memory length: 80531	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.24		
episode: 448	score: 6.0 memory length: 80867	epsilon: 1.0	steps: 336
lr: 0.0001	evaluation reward: 1.29		
episode: 449	score: 2.0 memory length: 81064	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.29		
episode: 450	score: 3.0 memory length: 81309	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.3		
episode: 451	score: 1.0 memory length: 81460	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3		
episode: 452	score: 3.0 memory length: 81706	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.32		
episode: 453	score: 0.0 memory length: 81829	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 454	score: 0.0 memory length: 81952	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 455	score: 1.0 memory length: 82122	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.32		
episode: 456	score: 0.0 memory length: 82245	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31		
episode: 457	score: 1.0 memory length: 82414	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.31		
episode: 458	score: 1.0 memory length: 82583	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.27		
episode: 459	score: 1.0 memory length: 82754	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.26		
episode: 460	score: 0.0 memory length: 82877	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.24		
episode: 461	score: 3.0 memory length: 83145	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.25		
episode: 462	score: 0.0 memory length: 83268	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.24		
episode: 463	score: 2.0 memory length: 83448	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.26		
episode: 464	score: 0.0 memory length: 83571	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.26		
episode: 465	score: 0.0 memory length: 83694	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.25		
episode: 466	score: 1.0 memory length: 83845	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.26		
episode: 467	score: 2.0 memory length: 84062	epsilon: 1.0	steps: 217

lr: 0.0001	evaluation reward: 1.28		
episode: 468	score: 1.0 memory length: 84213	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.26		
episode: 469	score: 1.0 memory length: 84382	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.27		
episode: 470	score: 2.0 memory length: 84564	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.27		
episode: 471	score: 3.0 memory length: 84792	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.3		
episode: 472	score: 1.0 memory length: 84943	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.29		
episode: 473	score: 2.0 memory length: 85141	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.28		
episode: 474	score: 2.0 memory length: 85341	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.3		
episode: 475	score: 0.0 memory length: 85463	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3		
episode: 476	score: 0.0 memory length: 85585	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.27		
episode: 477	score: 1.0 memory length: 85754	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.26		
episode: 478	score: 3.0 memory length: 86002	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.27		
episode: 479	score: 1.0 memory length: 86170	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.27		
episode: 480	score: 0.0 memory length: 86293	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.25		
episode: 481	score: 0.0 memory length: 86416	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.22		
episode: 482	score: 0.0 memory length: 86538	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.22		
episode: 483	score: 4.0 memory length: 86831	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.26		
episode: 484	score: 3.0 memory length: 87078	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.28		
episode: 485	score: 0.0 memory length: 87200	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.28		
episode: 486	score: 0.0 memory length: 87323	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.26		
episode: 487	score: 0.0 memory length: 87445	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.25		
episode: 488	score: 6.0 memory length: 87795	epsilon: 1.0	steps: 350
lr: 0.0001	evaluation reward: 1.28		
episode: 489	score: 2.0 memory length: 87995	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.28		
episode: 490	score: 2.0 memory length: 88192	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.27		
episode: 491	score: 2.0 memory length: 88409	epsilon: 1.0	steps: 217

lr: 0.0001	evaluation reward: 1.29		
episode: 492	score: 1.0 memory length: 88560	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.28		
episode: 493	score: 2.0 memory length: 88757	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.28		
episode: 494	score: 2.0 memory length: 88955	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.29		
episode: 495	score: 1.0 memory length: 89106	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3		
episode: 496	score: 3.0 memory length: 89350	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.32		
episode: 497	score: 3.0 memory length: 89576	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.34		
episode: 498	score: 2.0 memory length: 89774	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.35		
episode: 499	score: 0.0 memory length: 89896	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.34		
episode: 500	score: 2.0 memory length: 90114	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.32		
episode: 501	score: 2.0 memory length: 90312	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.34		
episode: 502	score: 0.0 memory length: 90435	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 503	score: 2.0 memory length: 90632	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.33		
episode: 504	score: 2.0 memory length: 90812	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.33		
episode: 505	score: 2.0 memory length: 91028	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.35		
episode: 506	score: 1.0 memory length: 91197	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.36		
episode: 507	score: 2.0 memory length: 91415	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.38		
episode: 508	score: 1.0 memory length: 91584	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39		
episode: 509	score: 0.0 memory length: 91707	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 510	score: 2.0 memory length: 91905	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.37		
episode: 511	score: 1.0 memory length: 92075	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.35		
episode: 512	score: 0.0 memory length: 92198	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.34		
episode: 513	score: 1.0 memory length: 92368	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.34		
episode: 514	score: 1.0 memory length: 92519	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.32		
episode: 515	score: 0.0 memory length: 92642	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.3		
episode: 516	score: 1.0 memory length: 92812	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.28		
episode: 517	score: 0.0 memory length: 92935	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.25		
episode: 518	score: 3.0 memory length: 93181	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.28		
episode: 519	score: 3.0 memory length: 93407	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.3		
episode: 520	score: 5.0 memory length: 93739	epsilon: 1.0	steps: 332
lr: 0.0001	evaluation reward: 1.35		
episode: 521	score: 1.0 memory length: 93889	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.34		
episode: 522	score: 1.0 memory length: 94059	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.35		
episode: 523	score: 0.0 memory length: 94182	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 524	score: 2.0 memory length: 94382	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.35		
episode: 525	score: 1.0 memory length: 94554	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.34		
episode: 526	score: 1.0 memory length: 94705	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.35		
episode: 527	score: 1.0 memory length: 94874	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.35		
episode: 528	score: 1.0 memory length: 95043	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.36		
episode: 529	score: 1.0 memory length: 95211	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.37		
episode: 530	score: 2.0 memory length: 95409	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.35		
episode: 531	score: 1.0 memory length: 95560	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.36		
episode: 532	score: 0.0 memory length: 95683	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 533	score: 0.0 memory length: 95806	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 534	score: 0.0 memory length: 95929	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.34		
episode: 535	score: 1.0 memory length: 96079	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.34		
episode: 536	score: 4.0 memory length: 96396	epsilon: 1.0	steps: 317
lr: 0.0001	evaluation reward: 1.38		
episode: 537	score: 1.0 memory length: 96565	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.37		
episode: 538	score: 2.0 memory length: 96783	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.36		
episode: 539	score: 2.0 memory length: 96963	epsilon: 1.0	steps: 180

lr: 0.0001	evaluation reward: 1.35		
episode: 540	score: 4.0	memory length: 97259	epsilon: 1.0 steps: 296
lr: 0.0001	evaluation reward: 1.39		
episode: 541	score: 3.0	memory length: 97488	epsilon: 1.0 steps: 229
lr: 0.0001	evaluation reward: 1.42		
episode: 542	score: 4.0	memory length: 97763	epsilon: 1.0 steps: 275
lr: 0.0001	evaluation reward: 1.46		
episode: 543	score: 3.0	memory length: 98028	epsilon: 1.0 steps: 265
lr: 0.0001	evaluation reward: 1.48		
episode: 544	score: 0.0	memory length: 98151	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 545	score: 1.0	memory length: 98301	epsilon: 1.0 steps: 150
lr: 0.0001	evaluation reward: 1.49		
episode: 546	score: 2.0	memory length: 98499	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.49		
episode: 547	score: 3.0	memory length: 98768	epsilon: 1.0 steps: 269
lr: 0.0001	evaluation reward: 1.5		
episode: 548	score: 1.0	memory length: 98919	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.45		
episode: 549	score: 0.0	memory length: 99041	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.43		
episode: 550	score: 1.0	memory length: 99192	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.41		
episode: 551	score: 3.0	memory length: 99458	epsilon: 1.0 steps: 266
lr: 0.0001	evaluation reward: 1.43		
episode: 552	score: 0.0	memory length: 99580	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.4		
episode: 553	score: 1.0	memory length: 99752	epsilon: 1.0 steps: 172
lr: 0.0001	evaluation reward: 1.41		
episode: 554	score: 2.0	memory length: 99971	epsilon: 1.0 steps: 219
lr: 0.0001	evaluation reward: 1.43		
episode: 555	score: 1.0	memory length: 100000	epsilon: 0.9997228000000006
steps: 168	lr: 0.0001	evaluation reward: 1.43	
episode: 556	score: 3.0	memory length: 100000	epsilon: 0.9992377000000165
steps: 245	lr: 0.0001	evaluation reward: 1.46	
episode: 557	score: 2.0	memory length: 100000	epsilon: 0.9988080400000259
steps: 217	lr: 0.0001	evaluation reward: 1.47	
episode: 558	score: 3.0	memory length: 100000	epsilon: 0.9983189800000365
steps: 247	lr: 0.0001	evaluation reward: 1.49	
episode: 559	score: 2.0	memory length: 100000	epsilon: 0.9979229800000451
steps: 200	lr: 0.0001	evaluation reward: 1.5	
episode: 560	score: 3.0	memory length: 100000	epsilon: 0.9973883800000567
steps: 270	lr: 0.0001	evaluation reward: 1.53	
episode: 561	score: 1.0	memory length: 100000	epsilon: 0.9970557400000639
steps: 168	lr: 0.0001	evaluation reward: 1.51	
episode: 562	score: 3.0	memory length: 100000	epsilon: 0.9966003400000738
steps: 230	lr: 0.0001	evaluation reward: 1.54	
episode: 563	score: 4.0	memory length: 100000	epsilon: 0.9960162400000865

steps: 295	lr: 0.0001	evaluation reward: 1.56
episode: 564	score: 0.0	memory length: 100000 epsilon: 0.9957727000000918
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 565	score: 2.0	memory length: 100000 epsilon: 0.9953410600001011
steps: 218	lr: 0.0001	evaluation reward: 1.58
episode: 566	score: 1.0	memory length: 100000 epsilon: 0.9950420800001076
steps: 151	lr: 0.0001	evaluation reward: 1.58
episode: 567	score: 0.0	memory length: 100000 epsilon: 0.9947985400001129
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 568	score: 1.0	memory length: 100000 epsilon: 0.9944995600001194
steps: 151	lr: 0.0001	evaluation reward: 1.56
episode: 569	score: 2.0	memory length: 100000 epsilon: 0.9941095000001279
steps: 197	lr: 0.0001	evaluation reward: 1.57
episode: 570	score: 2.0	memory length: 100000 epsilon: 0.9937095400001366
steps: 202	lr: 0.0001	evaluation reward: 1.57
episode: 571	score: 0.0	memory length: 100000 epsilon: 0.9934660000001418
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 572	score: 0.0	memory length: 100000 epsilon: 0.9932244400001471
steps: 122	lr: 0.0001	evaluation reward: 1.53
episode: 573	score: 2.0	memory length: 100000 epsilon: 0.9928324000001556
steps: 198	lr: 0.0001	evaluation reward: 1.53
episode: 574	score: 3.0	memory length: 100000 epsilon: 0.9923413600001663
steps: 248	lr: 0.0001	evaluation reward: 1.54
episode: 575	score: 2.0	memory length: 100000 epsilon: 0.9919513000001747
steps: 197	lr: 0.0001	evaluation reward: 1.56
episode: 576	score: 0.0	memory length: 100000 epsilon: 0.99170776000018
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 577	score: 1.0	memory length: 100000 epsilon: 0.9913751200001872
steps: 168	lr: 0.0001	evaluation reward: 1.56
episode: 578	score: 1.0	memory length: 100000 epsilon: 0.9910424800001945
steps: 168	lr: 0.0001	evaluation reward: 1.54
episode: 579	score: 3.0	memory length: 100000 epsilon: 0.9905158000002059
steps: 266	lr: 0.0001	evaluation reward: 1.56
episode: 580	score: 2.0	memory length: 100000 epsilon: 0.9901237600002144
steps: 198	lr: 0.0001	evaluation reward: 1.58
episode: 581	score: 0.0	memory length: 100000 epsilon: 0.9898822000002196
steps: 122	lr: 0.0001	evaluation reward: 1.58
episode: 582	score: 1.0	memory length: 100000 epsilon: 0.9895812400002262
steps: 152	lr: 0.0001	evaluation reward: 1.59
episode: 583	score: 2.0	memory length: 100000 epsilon: 0.9891852400002348
steps: 200	lr: 0.0001	evaluation reward: 1.57
episode: 584	score: 4.0	memory length: 100000 epsilon: 0.9886031200002474
steps: 294	lr: 0.0001	evaluation reward: 1.58
episode: 585	score: 1.0	memory length: 100000 epsilon: 0.9883041400002539
steps: 151	lr: 0.0001	evaluation reward: 1.59
episode: 586	score: 2.0	memory length: 100000 epsilon: 0.9879140800002624
steps: 197	lr: 0.0001	evaluation reward: 1.61
episode: 587	score: 2.0	memory length: 100000 epsilon: 0.9874824400002717

steps: 218	lr: 0.0001	evaluation reward: 1.63
episode: 588	score: 2.0	memory length: 100000 epsilon: 0.9870508000002811
steps: 218	lr: 0.0001	evaluation reward: 1.59
episode: 589	score: 1.0	memory length: 100000 epsilon: 0.9867161800002884
steps: 169	lr: 0.0001	evaluation reward: 1.58
episode: 590	score: 4.0	memory length: 100000 epsilon: 0.9862073200002994
steps: 257	lr: 0.0001	evaluation reward: 1.6
episode: 591	score: 3.0	memory length: 100000 epsilon: 0.9857123200003102
steps: 250	lr: 0.0001	evaluation reward: 1.61
episode: 592	score: 1.0	memory length: 100000 epsilon: 0.9853737400003175
steps: 171	lr: 0.0001	evaluation reward: 1.61
episode: 593	score: 0.0	memory length: 100000 epsilon: 0.9851302000003228
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 594	score: 2.0	memory length: 100000 epsilon: 0.9847401400003313
steps: 197	lr: 0.0001	evaluation reward: 1.59
episode: 595	score: 1.0	memory length: 100000 epsilon: 0.9844055200003385
steps: 169	lr: 0.0001	evaluation reward: 1.59
episode: 596	score: 1.0	memory length: 100000 epsilon: 0.9840689200003458
steps: 170	lr: 0.0001	evaluation reward: 1.57
episode: 597	score: 1.0	memory length: 100000 epsilon: 0.9837659800003524
steps: 153	lr: 0.0001	evaluation reward: 1.55
episode: 598	score: 0.0	memory length: 100000 epsilon: 0.9835244200003577
steps: 122	lr: 0.0001	evaluation reward: 1.53
episode: 599	score: 3.0	memory length: 100000 epsilon: 0.9830690200003676
steps: 230	lr: 0.0001	evaluation reward: 1.56
episode: 600	score: 0.0	memory length: 100000 epsilon: 0.9828254800003728
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 601	score: 2.0	memory length: 100000 epsilon: 0.9824354200003813
steps: 197	lr: 0.0001	evaluation reward: 1.54
episode: 602	score: 2.0	memory length: 100000 epsilon: 0.9820453600003898
steps: 197	lr: 0.0001	evaluation reward: 1.56
episode: 603	score: 0.0	memory length: 100000 epsilon: 0.9818018200003951
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 604	score: 0.0	memory length: 100000 epsilon: 0.9815582800004004
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 605	score: 0.0	memory length: 100000 epsilon: 0.9813147400004056
steps: 123	lr: 0.0001	evaluation reward: 1.5
episode: 606	score: 1.0	memory length: 100000 epsilon: 0.9810177400004121
steps: 150	lr: 0.0001	evaluation reward: 1.5
episode: 607	score: 2.0	memory length: 100000 epsilon: 0.9805861000004215
steps: 218	lr: 0.0001	evaluation reward: 1.5
episode: 608	score: 1.0	memory length: 100000 epsilon: 0.9802514800004287
steps: 169	lr: 0.0001	evaluation reward: 1.5
episode: 609	score: 1.0	memory length: 100000 epsilon: 0.979914880000436
steps: 170	lr: 0.0001	evaluation reward: 1.51
episode: 610	score: 3.0	memory length: 100000 epsilon: 0.9794634400004458
steps: 228	lr: 0.0001	evaluation reward: 1.52
episode: 611	score: 2.0	memory length: 100000 epsilon: 0.9790733800004543

steps: 197	lr: 0.0001	evaluation reward: 1.53
episode: 612	score: 0.0	memory length: 100000 epsilon: 0.9788318200004595
steps: 122	lr: 0.0001	evaluation reward: 1.53
episode: 613	score: 3.0	memory length: 100000 epsilon: 0.9783447400004701
steps: 246	lr: 0.0001	evaluation reward: 1.55
episode: 614	score: 0.0	memory length: 100000 epsilon: 0.9781031800004754
steps: 122	lr: 0.0001	evaluation reward: 1.54
episode: 615	score: 0.0	memory length: 100000 epsilon: 0.9778596400004806
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 616	score: 1.0	memory length: 100000 epsilon: 0.9775250200004879
steps: 169	lr: 0.0001	evaluation reward: 1.54
episode: 617	score: 2.0	memory length: 100000 epsilon: 0.9770914000004973
steps: 219	lr: 0.0001	evaluation reward: 1.56
episode: 618	score: 6.0	memory length: 100000 epsilon: 0.9764241400005118
steps: 337	lr: 0.0001	evaluation reward: 1.59
episode: 619	score: 0.0	memory length: 100000 epsilon: 0.9761806000005171
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 620	score: 6.0	memory length: 100000 epsilon: 0.9754856200005322
steps: 351	lr: 0.0001	evaluation reward: 1.57
episode: 621	score: 4.0	memory length: 100000 epsilon: 0.9748678600005456
steps: 312	lr: 0.0001	evaluation reward: 1.6
episode: 622	score: 1.0	memory length: 100000 epsilon: 0.9745688800005521
steps: 151	lr: 0.0001	evaluation reward: 1.6
episode: 623	score: 0.0	memory length: 100000 epsilon: 0.9743273200005573
steps: 122	lr: 0.0001	evaluation reward: 1.6
episode: 624	score: 1.0	memory length: 100000 epsilon: 0.9740283400005638
steps: 151	lr: 0.0001	evaluation reward: 1.59
episode: 625	score: 2.0	memory length: 100000 epsilon: 0.9735947200005732
steps: 219	lr: 0.0001	evaluation reward: 1.6
episode: 626	score: 2.0	memory length: 100000 epsilon: 0.973236340000581
steps: 181	lr: 0.0001	evaluation reward: 1.61
episode: 627	score: 1.0	memory length: 100000 epsilon: 0.9728977600005884
steps: 171	lr: 0.0001	evaluation reward: 1.61
episode: 628	score: 1.0	memory length: 100000 epsilon: 0.9725987800005949
steps: 151	lr: 0.0001	evaluation reward: 1.61
episode: 629	score: 0.0	memory length: 100000 epsilon: 0.9723552400006001
steps: 123	lr: 0.0001	evaluation reward: 1.6
episode: 630	score: 2.0	memory length: 100000 epsilon: 0.9719632000006087
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 631	score: 3.0	memory length: 100000 epsilon: 0.9714800800006191
steps: 244	lr: 0.0001	evaluation reward: 1.62
episode: 632	score: 2.0	memory length: 100000 epsilon: 0.9710880400006277
steps: 198	lr: 0.0001	evaluation reward: 1.64
episode: 633	score: 3.0	memory length: 100000 epsilon: 0.9705712600006389
steps: 261	lr: 0.0001	evaluation reward: 1.67
episode: 634	score: 4.0	memory length: 100000 epsilon: 0.9699871600006515
steps: 295	lr: 0.0001	evaluation reward: 1.71
episode: 635	score: 3.0	memory length: 100000 epsilon: 0.9695000800006621

steps: 246	lr: 0.0001	evaluation reward: 1.73
episode: 636	score: 1.0	memory length: 100000 epsilon: 0.9692011000006686
steps: 151	lr: 0.0001	evaluation reward: 1.7
episode: 637	score: 2.0	memory length: 100000 epsilon: 0.9688090600006771
steps: 198	lr: 0.0001	evaluation reward: 1.71
episode: 638	score: 2.0	memory length: 100000 epsilon: 0.9683714800006866
steps: 221	lr: 0.0001	evaluation reward: 1.71
episode: 639	score: 0.0	memory length: 100000 epsilon: 0.9681279400006919
steps: 123	lr: 0.0001	evaluation reward: 1.69
episode: 640	score: 3.0	memory length: 100000 epsilon: 0.9675953200007035
steps: 269	lr: 0.0001	evaluation reward: 1.68
episode: 641	score: 2.0	memory length: 100000 epsilon: 0.9671676400007128
steps: 216	lr: 0.0001	evaluation reward: 1.67
episode: 642	score: 0.0	memory length: 100000 epsilon: 0.966926080000718
steps: 122	lr: 0.0001	evaluation reward: 1.63
episode: 643	score: 1.0	memory length: 100000 epsilon: 0.9666290800007244
steps: 150	lr: 0.0001	evaluation reward: 1.61
episode: 644	score: 3.0	memory length: 100000 epsilon: 0.966142000000735
steps: 246	lr: 0.0001	evaluation reward: 1.64
episode: 645	score: 1.0	memory length: 100000 epsilon: 0.9658093600007422
steps: 168	lr: 0.0001	evaluation reward: 1.64
episode: 646	score: 1.0	memory length: 100000 epsilon: 0.9654747400007495
steps: 169	lr: 0.0001	evaluation reward: 1.63
episode: 647	score: 1.0	memory length: 100000 epsilon: 0.9651421000007567
steps: 168	lr: 0.0001	evaluation reward: 1.61
episode: 648	score: 2.0	memory length: 100000 epsilon: 0.9647500600007652
steps: 198	lr: 0.0001	evaluation reward: 1.62
episode: 649	score: 1.0	memory length: 100000 epsilon: 0.9644095000007726
steps: 172	lr: 0.0001	evaluation reward: 1.63
episode: 650	score: 0.0	memory length: 100000 epsilon: 0.9641679400007779
steps: 122	lr: 0.0001	evaluation reward: 1.62
episode: 651	score: 3.0	memory length: 100000 epsilon: 0.9636808600007885
steps: 246	lr: 0.0001	evaluation reward: 1.62
episode: 652	score: 1.0	memory length: 100000 epsilon: 0.9633462400007957
steps: 169	lr: 0.0001	evaluation reward: 1.63
episode: 653	score: 1.0	memory length: 100000 epsilon: 0.963011620000803
steps: 169	lr: 0.0001	evaluation reward: 1.63
episode: 654	score: 0.0	memory length: 100000 epsilon: 0.9627700600008082
steps: 122	lr: 0.0001	evaluation reward: 1.61
episode: 655	score: 5.0	memory length: 100000 epsilon: 0.9621245800008222
steps: 326	lr: 0.0001	evaluation reward: 1.65
episode: 656	score: 2.0	memory length: 100000 epsilon: 0.96176620000083
steps: 181	lr: 0.0001	evaluation reward: 1.64
episode: 657	score: 0.0	memory length: 100000 epsilon: 0.9615246400008353
steps: 122	lr: 0.0001	evaluation reward: 1.62
episode: 658	score: 1.0	memory length: 100000 epsilon: 0.9611840800008427
steps: 172	lr: 0.0001	evaluation reward: 1.6
episode: 659	score: 0.0	memory length: 100000 epsilon: 0.9609425200008479

steps: 122	lr: 0.0001	evaluation reward: 1.58	
episode: 660	score: 5.0	memory length: 100000	epsilon: 0.9602574400008628
steps: 346	lr: 0.0001	evaluation reward: 1.6	
episode: 661	score: 1.0	memory length: 100000	epsilon: 0.9599168800008702
steps: 172	lr: 0.0001	evaluation reward: 1.6	
episode: 662	score: 0.0	memory length: 100000	epsilon: 0.9596733400008755
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 663	score: 3.0	memory length: 100000	epsilon: 0.9592219000008853
steps: 228	lr: 0.0001	evaluation reward: 1.56	
episode: 664	score: 3.0	memory length: 100000	epsilon: 0.9587328400008959
steps: 247	lr: 0.0001	evaluation reward: 1.59	
episode: 665	score: 0.0	memory length: 100000	epsilon: 0.9584912800009011
steps: 122	lr: 0.0001	evaluation reward: 1.57	
episode: 666	score: 0.0	memory length: 100000	epsilon: 0.9582497200009064
steps: 122	lr: 0.0001	evaluation reward: 1.56	
episode: 667	score: 1.0	memory length: 100000	epsilon: 0.9579151000009136
steps: 169	lr: 0.0001	evaluation reward: 1.57	
episode: 668	score: 0.0	memory length: 100000	epsilon: 0.9576735400009189
steps: 122	lr: 0.0001	evaluation reward: 1.56	
episode: 669	score: 2.0	memory length: 100000	epsilon: 0.9572815000009274
steps: 198	lr: 0.0001	evaluation reward: 1.56	
episode: 670	score: 0.0	memory length: 100000	epsilon: 0.9570379600009327
steps: 123	lr: 0.0001	evaluation reward: 1.54	
episode: 671	score: 0.0	memory length: 100000	epsilon: 0.9567964000009379
steps: 122	lr: 0.0001	evaluation reward: 1.54	
episode: 672	score: 4.0	memory length: 100000	epsilon: 0.9562420000009499
steps: 280	lr: 0.0001	evaluation reward: 1.58	
episode: 673	score: 0.0	memory length: 100000	epsilon: 0.9560004400009552
steps: 122	lr: 0.0001	evaluation reward: 1.56	
episode: 674	score: 2.0	memory length: 100000	epsilon: 0.9556084000009637
steps: 198	lr: 0.0001	evaluation reward: 1.55	
episode: 675	score: 0.0	memory length: 100000	epsilon: 0.955364860000969
steps: 123	lr: 0.0001	evaluation reward: 1.53	
episode: 676	score: 0.0	memory length: 100000	epsilon: 0.9551233000009742
steps: 122	lr: 0.0001	evaluation reward: 1.53	
episode: 677	score: 0.0	memory length: 100000	epsilon: 0.9548797600009795
steps: 123	lr: 0.0001	evaluation reward: 1.52	
episode: 678	score: 4.0	memory length: 100000	epsilon: 0.95439862000099
steps: 243	lr: 0.0001	evaluation reward: 1.55	
episode: 679	score: 2.0	memory length: 100000	epsilon: 0.9540085600009984
steps: 197	lr: 0.0001	evaluation reward: 1.54	
episode: 680	score: 2.0	memory length: 100000	epsilon: 0.9535749400010078
steps: 219	lr: 0.0001	evaluation reward: 1.54	
episode: 681	score: 1.0	memory length: 100000	epsilon: 0.9532759600010143
steps: 151	lr: 0.0001	evaluation reward: 1.55	
episode: 682	score: 0.0	memory length: 100000	epsilon: 0.9530324200010196
steps: 123	lr: 0.0001	evaluation reward: 1.54	
episode: 683	score: 13.0	memory length: 100000	epsilon: 0.9520206400010416

steps: 511	lr: 0.0001	evaluation reward: 1.65
episode: 684	score: 0.0	memory length: 100000 epsilon: 0.9517771000010469
steps: 123	lr: 0.0001	evaluation reward: 1.61
episode: 685	score: 2.0	memory length: 100000 epsilon: 0.9513870400010553
steps: 197	lr: 0.0001	evaluation reward: 1.62
episode: 686	score: 0.0	memory length: 100000 epsilon: 0.9511435000010606
steps: 123	lr: 0.0001	evaluation reward: 1.6
episode: 687	score: 1.0	memory length: 100000 epsilon: 0.9508445200010671
steps: 151	lr: 0.0001	evaluation reward: 1.59
episode: 688	score: 0.0	memory length: 100000 epsilon: 0.9506029600010724
steps: 122	lr: 0.0001	evaluation reward: 1.57
episode: 689	score: 1.0	memory length: 100000 epsilon: 0.9502703200010796
steps: 168	lr: 0.0001	evaluation reward: 1.57
episode: 690	score: 2.0	memory length: 100000 epsilon: 0.9498406600010889
steps: 217	lr: 0.0001	evaluation reward: 1.55
episode: 691	score: 3.0	memory length: 100000 epsilon: 0.9493595200010994
steps: 243	lr: 0.0001	evaluation reward: 1.55
episode: 692	score: 0.0	memory length: 100000 epsilon: 0.9491159800011046
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 693	score: 0.0	memory length: 100000 epsilon: 0.9488724400011099
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 694	score: 0.0	memory length: 100000 epsilon: 0.9486289000011152
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 695	score: 3.0	memory length: 100000 epsilon: 0.9481378600011259
steps: 248	lr: 0.0001	evaluation reward: 1.54
episode: 696	score: 0.0	memory length: 100000 epsilon: 0.9478943200011312
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 697	score: 1.0	memory length: 100000 epsilon: 0.9475973200011376
steps: 150	lr: 0.0001	evaluation reward: 1.53
episode: 698	score: 0.0	memory length: 100000 epsilon: 0.9473537800011429
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 699	score: 0.0	memory length: 100000 epsilon: 0.9471102400011482
steps: 123	lr: 0.0001	evaluation reward: 1.5
episode: 700	score: 5.0	memory length: 100000 epsilon: 0.9465083200011613
steps: 304	lr: 0.0001	evaluation reward: 1.55
episode: 701	score: 1.0	memory length: 100000 epsilon: 0.9462113200011677
steps: 150	lr: 0.0001	evaluation reward: 1.54
episode: 702	score: 1.0	memory length: 100000 epsilon: 0.9459123400011742
steps: 151	lr: 0.0001	evaluation reward: 1.53
episode: 703	score: 0.0	memory length: 100000 epsilon: 0.9456688000011795
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 704	score: 2.0	memory length: 100000 epsilon: 0.945276760001188
steps: 198	lr: 0.0001	evaluation reward: 1.55
episode: 705	score: 2.0	memory length: 100000 epsilon: 0.9448867000011965
steps: 197	lr: 0.0001	evaluation reward: 1.57
episode: 706	score: 2.0	memory length: 100000 epsilon: 0.9444530800012059
steps: 219	lr: 0.0001	evaluation reward: 1.58
episode: 707	score: 0.0	memory length: 100000 epsilon: 0.9442095400012112

steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 708	score: 2.0	memory length: 100000 epsilon: 0.9438115600012198
steps: 201	lr: 0.0001	evaluation reward: 1.57
episode: 709	score: 0.0	memory length: 100000 epsilon: 0.9435680200012251
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 710	score: 0.0	memory length: 100000 epsilon: 0.9433244800012304
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 711	score: 0.0	memory length: 100000 epsilon: 0.9430829200012356
steps: 122	lr: 0.0001	evaluation reward: 1.51
episode: 712	score: 3.0	memory length: 100000 epsilon: 0.9425918800012463
steps: 248	lr: 0.0001	evaluation reward: 1.54
episode: 713	score: 1.0	memory length: 100000 epsilon: 0.9422572600012535
steps: 169	lr: 0.0001	evaluation reward: 1.52
episode: 714	score: 2.0	memory length: 100000 epsilon: 0.9418592800012622
steps: 201	lr: 0.0001	evaluation reward: 1.54
episode: 715	score: 1.0	memory length: 100000 epsilon: 0.9415207000012695
steps: 171	lr: 0.0001	evaluation reward: 1.55
episode: 716	score: 2.0	memory length: 100000 epsilon: 0.941128660001278
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 717	score: 2.0	memory length: 100000 epsilon: 0.9407366200012865
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 718	score: 1.0	memory length: 100000 epsilon: 0.9404000200012939
steps: 170	lr: 0.0001	evaluation reward: 1.51
episode: 719	score: 1.0	memory length: 100000 epsilon: 0.9401030200013003
steps: 150	lr: 0.0001	evaluation reward: 1.52
episode: 720	score: 1.0	memory length: 100000 epsilon: 0.9397703800013075
steps: 168	lr: 0.0001	evaluation reward: 1.47
episode: 721	score: 2.0	memory length: 100000 epsilon: 0.939378340001316
steps: 198	lr: 0.0001	evaluation reward: 1.45
episode: 722	score: 2.0	memory length: 100000 epsilon: 0.9389863000013245
steps: 198	lr: 0.0001	evaluation reward: 1.46
episode: 723	score: 4.0	memory length: 100000 epsilon: 0.9384695200013358
steps: 261	lr: 0.0001	evaluation reward: 1.5
episode: 724	score: 5.0	memory length: 100000 epsilon: 0.9377884000013506
steps: 344	lr: 0.0001	evaluation reward: 1.54
episode: 725	score: 2.0	memory length: 100000 epsilon: 0.9373963600013591
steps: 198	lr: 0.0001	evaluation reward: 1.54
episode: 726	score: 9.0	memory length: 100000 epsilon: 0.9366835600013745
steps: 360	lr: 0.0001	evaluation reward: 1.61
episode: 727	score: 1.0	memory length: 100000 epsilon: 0.936386560001381
steps: 150	lr: 0.0001	evaluation reward: 1.61
episode: 728	score: 0.0	memory length: 100000 epsilon: 0.9361450000013862
steps: 122	lr: 0.0001	evaluation reward: 1.6
episode: 729	score: 1.0	memory length: 100000 epsilon: 0.9358044400013936
steps: 172	lr: 0.0001	evaluation reward: 1.61
episode: 730	score: 5.0	memory length: 100000 epsilon: 0.9351352000014082
steps: 338	lr: 0.0001	evaluation reward: 1.64
episode: 731	score: 3.0	memory length: 100000 epsilon: 0.9346065400014196

steps: 267	lr: 0.0001	evaluation reward: 1.64
episode: 732	score: 2.0	memory length: 100000 epsilon: 0.9342125200014282
steps: 199	lr: 0.0001	evaluation reward: 1.64
episode: 733	score: 1.0	memory length: 100000 epsilon: 0.9339095800014348
steps: 153	lr: 0.0001	evaluation reward: 1.62
episode: 734	score: 2.0	memory length: 100000 epsilon: 0.9335175400014433
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 735	score: 2.0	memory length: 100000 epsilon: 0.9330859000014526
steps: 218	lr: 0.0001	evaluation reward: 1.59
episode: 736	score: 0.0	memory length: 100000 epsilon: 0.9328423600014579
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 737	score: 2.0	memory length: 100000 epsilon: 0.9324127000014673
steps: 217	lr: 0.0001	evaluation reward: 1.58
episode: 738	score: 2.0	memory length: 100000 epsilon: 0.9319830400014766
steps: 217	lr: 0.0001	evaluation reward: 1.58
episode: 739	score: 0.0	memory length: 100000 epsilon: 0.9317395000014819
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 740	score: 0.0	memory length: 100000 epsilon: 0.9314959600014872
steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 741	score: 0.0	memory length: 100000 epsilon: 0.9312524200014924
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 742	score: 2.0	memory length: 100000 epsilon: 0.9308544400015011
steps: 201	lr: 0.0001	evaluation reward: 1.55
episode: 743	score: 1.0	memory length: 100000 epsilon: 0.9305554600015076
steps: 151	lr: 0.0001	evaluation reward: 1.55
episode: 744	score: 0.0	memory length: 100000 epsilon: 0.9303139000015128
steps: 122	lr: 0.0001	evaluation reward: 1.52
episode: 745	score: 2.0	memory length: 100000 epsilon: 0.9299218600015213
steps: 198	lr: 0.0001	evaluation reward: 1.53
episode: 746	score: 2.0	memory length: 100000 epsilon: 0.9294902200015307
steps: 218	lr: 0.0001	evaluation reward: 1.54
episode: 747	score: 1.0	memory length: 100000 epsilon: 0.9291912400015372
steps: 151	lr: 0.0001	evaluation reward: 1.54
episode: 748	score: 0.0	memory length: 100000 epsilon: 0.9289496800015424
steps: 122	lr: 0.0001	evaluation reward: 1.52
episode: 749	score: 2.0	memory length: 100000 epsilon: 0.928553680001551
steps: 200	lr: 0.0001	evaluation reward: 1.53
episode: 750	score: 1.0	memory length: 100000 epsilon: 0.9282190600015583
steps: 169	lr: 0.0001	evaluation reward: 1.54
episode: 751	score: 1.0	memory length: 100000 epsilon: 0.9278844400015656
steps: 169	lr: 0.0001	evaluation reward: 1.52
episode: 752	score: 4.0	memory length: 100000 epsilon: 0.9273379600015774
steps: 276	lr: 0.0001	evaluation reward: 1.55
episode: 753	score: 4.0	memory length: 100000 epsilon: 0.9267360400015905
steps: 304	lr: 0.0001	evaluation reward: 1.58
episode: 754	score: 0.0	memory length: 100000 epsilon: 0.9264925000015958
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 755	score: 2.0	memory length: 100000 epsilon: 0.9261004600016043

steps: 198	lr: 0.0001	evaluation reward: 1.55
episode: 756	score: 1.0	memory length: 100000 epsilon: 0.9257638600016116
steps: 170	lr: 0.0001	evaluation reward: 1.54
episode: 757	score: 0.0	memory length: 100000 epsilon: 0.9255203200016169
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 758	score: 1.0	memory length: 100000 epsilon: 0.9251817400016242
steps: 171	lr: 0.0001	evaluation reward: 1.54
episode: 759	score: 0.0	memory length: 100000 epsilon: 0.9249382000016295
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 760	score: 2.0	memory length: 100000 epsilon: 0.924546160001638
steps: 198	lr: 0.0001	evaluation reward: 1.51
episode: 761	score: 2.0	memory length: 100000 epsilon: 0.9241541200016465
steps: 198	lr: 0.0001	evaluation reward: 1.52
episode: 762	score: 1.0	memory length: 100000 epsilon: 0.9238135600016539
steps: 172	lr: 0.0001	evaluation reward: 1.53
episode: 763	score: 0.0	memory length: 100000 epsilon: 0.9235720000016592
steps: 122	lr: 0.0001	evaluation reward: 1.5
episode: 764	score: 1.0	memory length: 100000 epsilon: 0.9232354000016665
steps: 170	lr: 0.0001	evaluation reward: 1.48
episode: 765	score: 5.0	memory length: 100000 epsilon: 0.9225602200016811
steps: 341	lr: 0.0001	evaluation reward: 1.53
episode: 766	score: 0.0	memory length: 100000 epsilon: 0.9223166800016864
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 767	score: 2.0	memory length: 100000 epsilon: 0.9218850400016958
steps: 218	lr: 0.0001	evaluation reward: 1.54
episode: 768	score: 0.0	memory length: 100000 epsilon: 0.9216415000017011
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 769	score: 0.0	memory length: 100000 epsilon: 0.9213979600017064
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 770	score: 1.0	memory length: 100000 epsilon: 0.9210633400017136
steps: 169	lr: 0.0001	evaluation reward: 1.53
episode: 771	score: 2.0	memory length: 100000 epsilon: 0.9206673400017222
steps: 200	lr: 0.0001	evaluation reward: 1.55
episode: 772	score: 0.0	memory length: 100000 epsilon: 0.9204238000017275
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 773	score: 1.0	memory length: 100000 epsilon: 0.920124820001734
steps: 151	lr: 0.0001	evaluation reward: 1.52
episode: 774	score: 1.0	memory length: 100000 epsilon: 0.9198258400017405
steps: 151	lr: 0.0001	evaluation reward: 1.51
episode: 775	score: 1.0	memory length: 100000 epsilon: 0.9194892400017478
steps: 170	lr: 0.0001	evaluation reward: 1.52
episode: 776	score: 0.0	memory length: 100000 epsilon: 0.9192457000017531
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 777	score: 0.0	memory length: 100000 epsilon: 0.9190021600017584
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 778	score: 1.0	memory length: 100000 epsilon: 0.9186695200017656
steps: 168	lr: 0.0001	evaluation reward: 1.49
episode: 779	score: 2.0	memory length: 100000 epsilon: 0.9182695600017743

steps: 202	lr: 0.0001	evaluation reward: 1.49
episode: 780	score: 2.0	memory length: 100000 epsilon: 0.9178775200017828
steps: 198	lr: 0.0001	evaluation reward: 1.49
episode: 781	score: 0.0	memory length: 100000 epsilon: 0.917635960001788
steps: 122	lr: 0.0001	evaluation reward: 1.48
episode: 782	score: 1.0	memory length: 100000 epsilon: 0.9172973800017954
steps: 171	lr: 0.0001	evaluation reward: 1.49
episode: 783	score: 1.0	memory length: 100000 epsilon: 0.9169627600018027
steps: 169	lr: 0.0001	evaluation reward: 1.37
episode: 784	score: 0.0	memory length: 100000 epsilon: 0.9167192200018079
steps: 123	lr: 0.0001	evaluation reward: 1.37
episode: 785	score: 2.0	memory length: 100000 epsilon: 0.9162915400018172
steps: 216	lr: 0.0001	evaluation reward: 1.37
episode: 786	score: 0.0	memory length: 100000 epsilon: 0.9160499800018225
steps: 122	lr: 0.0001	evaluation reward: 1.37
episode: 787	score: 2.0	memory length: 100000 epsilon: 0.915657940001831
steps: 198	lr: 0.0001	evaluation reward: 1.38
episode: 788	score: 3.0	memory length: 100000 epsilon: 0.9151312600018424
steps: 266	lr: 0.0001	evaluation reward: 1.41
episode: 789	score: 2.0	memory length: 100000 epsilon: 0.9147313000018511
steps: 202	lr: 0.0001	evaluation reward: 1.42
episode: 790	score: 1.0	memory length: 100000 epsilon: 0.9143966800018584
steps: 169	lr: 0.0001	evaluation reward: 1.41
episode: 791	score: 0.0	memory length: 100000 epsilon: 0.9141531400018637
steps: 123	lr: 0.0001	evaluation reward: 1.38
episode: 792	score: 3.0	memory length: 100000 epsilon: 0.9137036800018734
steps: 227	lr: 0.0001	evaluation reward: 1.41
episode: 793	score: 1.0	memory length: 100000 epsilon: 0.9133670800018807
steps: 170	lr: 0.0001	evaluation reward: 1.42
episode: 794	score: 3.0	memory length: 100000 epsilon: 0.9129176200018905
steps: 227	lr: 0.0001	evaluation reward: 1.45
episode: 795	score: 2.0	memory length: 100000 epsilon: 0.9125275600018989
steps: 197	lr: 0.0001	evaluation reward: 1.44
episode: 796	score: 4.0	memory length: 100000 epsilon: 0.9119335600019118
steps: 300	lr: 0.0001	evaluation reward: 1.48
episode: 797	score: 0.0	memory length: 100000 epsilon: 0.9116900200019171
steps: 123	lr: 0.0001	evaluation reward: 1.47
episode: 798	score: 2.0	memory length: 100000 epsilon: 0.9112979800019256
steps: 198	lr: 0.0001	evaluation reward: 1.49
episode: 799	score: 0.0	memory length: 100000 epsilon: 0.9110544400019309
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 800	score: 2.0	memory length: 100000 epsilon: 0.9106624000019394
steps: 198	lr: 0.0001	evaluation reward: 1.46
episode: 801	score: 0.0	memory length: 100000 epsilon: 0.9104208400019447
steps: 122	lr: 0.0001	evaluation reward: 1.45
episode: 802	score: 0.0	memory length: 100000 epsilon: 0.9101792800019499
steps: 122	lr: 0.0001	evaluation reward: 1.44
episode: 803	score: 1.0	memory length: 100000 epsilon: 0.9098822800019564

steps: 150	lr: 0.0001	evaluation reward: 1.45
episode: 804	score: 2.0	memory length: 100000 epsilon: 0.9094902400019649
steps: 198	lr: 0.0001	evaluation reward: 1.45
episode: 805	score: 0.0	memory length: 100000 epsilon: 0.9092486800019701
steps: 122	lr: 0.0001	evaluation reward: 1.43
episode: 806	score: 2.0	memory length: 100000 epsilon: 0.9088210000019794
steps: 216	lr: 0.0001	evaluation reward: 1.43
episode: 807	score: 3.0	memory length: 100000 epsilon: 0.9082903600019909
steps: 268	lr: 0.0001	evaluation reward: 1.46
episode: 808	score: 1.0	memory length: 100000 epsilon: 0.9079557400019982
steps: 169	lr: 0.0001	evaluation reward: 1.45
episode: 809	score: 2.0	memory length: 100000 epsilon: 0.9075597400020068
steps: 200	lr: 0.0001	evaluation reward: 1.47
episode: 810	score: 0.0	memory length: 100000 epsilon: 0.9073162000020121
steps: 123	lr: 0.0001	evaluation reward: 1.47
episode: 811	score: 0.0	memory length: 100000 epsilon: 0.9070726600020174
steps: 123	lr: 0.0001	evaluation reward: 1.47
episode: 812	score: 1.0	memory length: 100000 epsilon: 0.9067756600020238
steps: 150	lr: 0.0001	evaluation reward: 1.45
episode: 813	score: 3.0	memory length: 100000 epsilon: 0.9062489800020352
steps: 266	lr: 0.0001	evaluation reward: 1.47
episode: 814	score: 2.0	memory length: 100000 epsilon: 0.9058589200020437
steps: 197	lr: 0.0001	evaluation reward: 1.47
episode: 815	score: 2.0	memory length: 100000 epsilon: 0.905429260002053
steps: 217	lr: 0.0001	evaluation reward: 1.48
episode: 816	score: 2.0	memory length: 100000 epsilon: 0.9050372200020615
steps: 198	lr: 0.0001	evaluation reward: 1.48
episode: 817	score: 2.0	memory length: 100000 epsilon: 0.9046451800020701
steps: 198	lr: 0.0001	evaluation reward: 1.48
episode: 818	score: 1.0	memory length: 100000 epsilon: 0.9043125400020773
steps: 168	lr: 0.0001	evaluation reward: 1.48
episode: 819	score: 2.0	memory length: 100000 epsilon: 0.9039205000020858
steps: 198	lr: 0.0001	evaluation reward: 1.49
episode: 820	score: 3.0	memory length: 100000 epsilon: 0.9034334200020964
steps: 246	lr: 0.0001	evaluation reward: 1.51
episode: 821	score: 2.0	memory length: 100000 epsilon: 0.9030413800021049
steps: 198	lr: 0.0001	evaluation reward: 1.51
episode: 822	score: 3.0	memory length: 100000 epsilon: 0.9025939000021146
steps: 226	lr: 0.0001	evaluation reward: 1.52
episode: 823	score: 0.0	memory length: 100000 epsilon: 0.9023503600021199
steps: 123	lr: 0.0001	evaluation reward: 1.48
episode: 824	score: 0.0	memory length: 100000 epsilon: 0.9021068200021252
steps: 123	lr: 0.0001	evaluation reward: 1.43
episode: 825	score: 0.0	memory length: 100000 epsilon: 0.9018652600021304
steps: 122	lr: 0.0001	evaluation reward: 1.41
episode: 826	score: 3.0	memory length: 100000 epsilon: 0.901376200002141
steps: 247	lr: 0.0001	evaluation reward: 1.35
episode: 827	score: 4.0	memory length: 100000 epsilon: 0.9008277400021529

steps: 277	lr: 0.0001	evaluation reward: 1.38
episode: 828	score: 3.0	memory length: 100000 epsilon: 0.9003802600021626
steps: 226	lr: 0.0001	evaluation reward: 1.41
episode: 829	score: 5.0	memory length: 100000 epsilon: 0.8996991400021774
steps: 344	lr: 0.0001	evaluation reward: 1.45
episode: 830	score: 2.0	memory length: 100000 epsilon: 0.8993090800021859
steps: 197	lr: 0.0001	evaluation reward: 1.42
episode: 831	score: 1.0	memory length: 100000 epsilon: 0.8989744600021932
steps: 169	lr: 0.0001	evaluation reward: 1.4
episode: 832	score: 2.0	memory length: 100000 epsilon: 0.8986061800022012
steps: 186	lr: 0.0001	evaluation reward: 1.4
episode: 833	score: 1.0	memory length: 100000 epsilon: 0.8982715600022084
steps: 169	lr: 0.0001	evaluation reward: 1.4
episode: 834	score: 2.0	memory length: 100000 epsilon: 0.8979131800022162
steps: 181	lr: 0.0001	evaluation reward: 1.4
episode: 835	score: 0.0	memory length: 100000 epsilon: 0.8976696400022215
steps: 123	lr: 0.0001	evaluation reward: 1.38
episode: 836	score: 0.0	memory length: 100000 epsilon: 0.8974280800022267
steps: 122	lr: 0.0001	evaluation reward: 1.38
episode: 837	score: 1.0	memory length: 100000 epsilon: 0.8970895000022341
steps: 171	lr: 0.0001	evaluation reward: 1.37
episode: 838	score: 0.0	memory length: 100000 epsilon: 0.8968479400022393
steps: 122	lr: 0.0001	evaluation reward: 1.35
episode: 839	score: 1.0	memory length: 100000 epsilon: 0.8965153000022466
steps: 168	lr: 0.0001	evaluation reward: 1.36
episode: 840	score: 4.0	memory length: 100000 epsilon: 0.8958916000022601
steps: 315	lr: 0.0001	evaluation reward: 1.4
episode: 841	score: 0.0	memory length: 100000 epsilon: 0.8956480600022654
steps: 123	lr: 0.0001	evaluation reward: 1.4
episode: 842	score: 2.0	memory length: 100000 epsilon: 0.8952580000022738
steps: 197	lr: 0.0001	evaluation reward: 1.4
episode: 843	score: 0.0	memory length: 100000 epsilon: 0.8950144600022791
steps: 123	lr: 0.0001	evaluation reward: 1.39
episode: 844	score: 1.0	memory length: 100000 epsilon: 0.8946739000022865
steps: 172	lr: 0.0001	evaluation reward: 1.4
episode: 845	score: 0.0	memory length: 100000 epsilon: 0.8944323400022918
steps: 122	lr: 0.0001	evaluation reward: 1.38
episode: 846	score: 2.0	memory length: 100000 epsilon: 0.8940007000023011
steps: 218	lr: 0.0001	evaluation reward: 1.38
episode: 847	score: 4.0	memory length: 100000 epsilon: 0.8934502600023131
steps: 278	lr: 0.0001	evaluation reward: 1.41
episode: 848	score: 0.0	memory length: 100000 epsilon: 0.8932067200023184
steps: 123	lr: 0.0001	evaluation reward: 1.41
episode: 849	score: 3.0	memory length: 100000 epsilon: 0.8927592400023281
steps: 226	lr: 0.0001	evaluation reward: 1.42
episode: 850	score: 0.0	memory length: 100000 epsilon: 0.8925157000023334
steps: 123	lr: 0.0001	evaluation reward: 1.41
episode: 851	score: 1.0	memory length: 100000 epsilon: 0.8922167200023399

steps: 151	lr: 0.0001	evaluation reward: 1.41
episode: 852	score: 2.0	memory length: 100000 epsilon: 0.8918266600023483
steps: 197	lr: 0.0001	evaluation reward: 1.39
episode: 853	score: 3.0	memory length: 100000 epsilon: 0.891381160002358
steps: 225	lr: 0.0001	evaluation reward: 1.38
episode: 854	score: 2.0	memory length: 100000 epsilon: 0.8909475400023674
steps: 219	lr: 0.0001	evaluation reward: 1.4
episode: 855	score: 3.0	memory length: 100000 epsilon: 0.890460460002378
steps: 246	lr: 0.0001	evaluation reward: 1.41
episode: 856	score: 3.0	memory length: 100000 epsilon: 0.8900110000023878
steps: 227	lr: 0.0001	evaluation reward: 1.43
episode: 857	score: 0.0	memory length: 100000 epsilon: 0.889769440002393
steps: 122	lr: 0.0001	evaluation reward: 1.43
episode: 858	score: 1.0	memory length: 100000 epsilon: 0.8894348200024003
steps: 169	lr: 0.0001	evaluation reward: 1.43
episode: 859	score: 3.0	memory length: 100000 epsilon: 0.8889101200024117
steps: 265	lr: 0.0001	evaluation reward: 1.46
episode: 860	score: 2.0	memory length: 100000 epsilon: 0.8885497600024195
steps: 182	lr: 0.0001	evaluation reward: 1.46
episode: 861	score: 2.0	memory length: 100000 epsilon: 0.888157720002428
steps: 198	lr: 0.0001	evaluation reward: 1.46
episode: 862	score: 2.0	memory length: 100000 epsilon: 0.8877241000024374
steps: 219	lr: 0.0001	evaluation reward: 1.47
episode: 863	score: 1.0	memory length: 100000 epsilon: 0.8873875000024447
steps: 170	lr: 0.0001	evaluation reward: 1.48
episode: 864	score: 2.0	memory length: 100000 epsilon: 0.8869915000024533
steps: 200	lr: 0.0001	evaluation reward: 1.49
episode: 865	score: 3.0	memory length: 100000 epsilon: 0.8864668000024647
steps: 265	lr: 0.0001	evaluation reward: 1.47
episode: 866	score: 1.0	memory length: 100000 epsilon: 0.8861698000024711
steps: 150	lr: 0.0001	evaluation reward: 1.48
episode: 867	score: 2.0	memory length: 100000 epsilon: 0.8857322200024806
steps: 221	lr: 0.0001	evaluation reward: 1.48
episode: 868	score: 1.0	memory length: 100000 epsilon: 0.8853976000024879
steps: 169	lr: 0.0001	evaluation reward: 1.49
episode: 869	score: 0.0	memory length: 100000 epsilon: 0.8851560400024931
steps: 122	lr: 0.0001	evaluation reward: 1.49
episode: 870	score: 0.0	memory length: 100000 epsilon: 0.8849125000024984
steps: 123	lr: 0.0001	evaluation reward: 1.48
episode: 871	score: 2.0	memory length: 100000 epsilon: 0.8844828400025078
steps: 217	lr: 0.0001	evaluation reward: 1.48
episode: 872	score: 1.0	memory length: 100000 epsilon: 0.884148220002515
steps: 169	lr: 0.0001	evaluation reward: 1.49
episode: 873	score: 1.0	memory length: 100000 epsilon: 0.8838492400025215
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 874	score: 0.0	memory length: 100000 epsilon: 0.8836076800025268
steps: 122	lr: 0.0001	evaluation reward: 1.48
episode: 875	score: 2.0	memory length: 100000 epsilon: 0.8831740600025362

steps: 219	lr: 0.0001	evaluation reward: 1.49
episode: 876	score: 1.0	memory length: 100000 epsilon: 0.8828374600025435
steps: 170	lr: 0.0001	evaluation reward: 1.5
episode: 877	score: 4.0	memory length: 100000 epsilon: 0.8822533600025562
steps: 295	lr: 0.0001	evaluation reward: 1.54
episode: 878	score: 3.0	memory length: 100000 epsilon: 0.8817662800025667
steps: 246	lr: 0.0001	evaluation reward: 1.56
episode: 879	score: 3.0	memory length: 100000 epsilon: 0.8812415800025781
steps: 265	lr: 0.0001	evaluation reward: 1.57
episode: 880	score: 1.0	memory length: 100000 epsilon: 0.8809069600025854
steps: 169	lr: 0.0001	evaluation reward: 1.56
episode: 881	score: 4.0	memory length: 100000 epsilon: 0.8803327600025979
steps: 290	lr: 0.0001	evaluation reward: 1.6
episode: 882	score: 5.0	memory length: 100000 epsilon: 0.8796476800026127
steps: 346	lr: 0.0001	evaluation reward: 1.64
episode: 883	score: 2.0	memory length: 100000 epsilon: 0.8792180200026221
steps: 217	lr: 0.0001	evaluation reward: 1.65
episode: 884	score: 2.0	memory length: 100000 epsilon: 0.8787844000026315
steps: 219	lr: 0.0001	evaluation reward: 1.67
episode: 885	score: 0.0	memory length: 100000 epsilon: 0.8785408600026368
steps: 123	lr: 0.0001	evaluation reward: 1.65
episode: 886	score: 1.0	memory length: 100000 epsilon: 0.878208220002644
steps: 168	lr: 0.0001	evaluation reward: 1.66
episode: 887	score: 3.0	memory length: 100000 epsilon: 0.8777211400026546
steps: 246	lr: 0.0001	evaluation reward: 1.67
episode: 888	score: 1.0	memory length: 100000 epsilon: 0.877422160002661
steps: 151	lr: 0.0001	evaluation reward: 1.65
episode: 889	score: 1.0	memory length: 100000 epsilon: 0.8771231800026675
steps: 151	lr: 0.0001	evaluation reward: 1.64
episode: 890	score: 0.0	memory length: 100000 epsilon: 0.8768816200026728
steps: 122	lr: 0.0001	evaluation reward: 1.63
episode: 891	score: 3.0	memory length: 100000 epsilon: 0.8764242400026827
steps: 231	lr: 0.0001	evaluation reward: 1.66
episode: 892	score: 1.0	memory length: 100000 epsilon: 0.87608962000269
steps: 169	lr: 0.0001	evaluation reward: 1.64
episode: 893	score: 5.0	memory length: 100000 epsilon: 0.875489680002703
steps: 303	lr: 0.0001	evaluation reward: 1.68
episode: 894	score: 2.0	memory length: 100000 epsilon: 0.8750600200027123
steps: 217	lr: 0.0001	evaluation reward: 1.67
episode: 895	score: 1.0	memory length: 100000 epsilon: 0.8747234200027196
steps: 170	lr: 0.0001	evaluation reward: 1.66
episode: 896	score: 3.0	memory length: 100000 epsilon: 0.8742779200027293
steps: 225	lr: 0.0001	evaluation reward: 1.65
episode: 897	score: 1.0	memory length: 100000 epsilon: 0.8739789400027358
steps: 151	lr: 0.0001	evaluation reward: 1.66
episode: 898	score: 0.0	memory length: 100000 epsilon: 0.8737354000027411
steps: 123	lr: 0.0001	evaluation reward: 1.64
episode: 899	score: 1.0	memory length: 100000 epsilon: 0.8734384000027475

steps: 150	lr: 0.0001	evaluation reward: 1.65
episode: 900	score: 1.0	memory length: 100000 epsilon: 0.8731018000027548
steps: 170	lr: 0.0001	evaluation reward: 1.64
episode: 901	score: 2.0	memory length: 100000 epsilon: 0.8726701600027642
steps: 218	lr: 0.0001	evaluation reward: 1.66
episode: 902	score: 3.0	memory length: 100000 epsilon: 0.8721830800027748
steps: 246	lr: 0.0001	evaluation reward: 1.69
episode: 903	score: 1.0	memory length: 100000 epsilon: 0.8718464800027821
steps: 170	lr: 0.0001	evaluation reward: 1.69
episode: 904	score: 0.0	memory length: 100000 epsilon: 0.8716049200027873
steps: 122	lr: 0.0001	evaluation reward: 1.67
episode: 905	score: 0.0	memory length: 100000 epsilon: 0.8713613800027926
steps: 123	lr: 0.0001	evaluation reward: 1.67
episode: 906	score: 1.0	memory length: 100000 epsilon: 0.8710643800027991
steps: 150	lr: 0.0001	evaluation reward: 1.66
episode: 907	score: 0.0	memory length: 100000 epsilon: 0.8708208400028044
steps: 123	lr: 0.0001	evaluation reward: 1.63
episode: 908	score: 5.0	memory length: 100000 epsilon: 0.8701535800028188
steps: 337	lr: 0.0001	evaluation reward: 1.67
episode: 909	score: 2.0	memory length: 100000 epsilon: 0.8697615400028273
steps: 198	lr: 0.0001	evaluation reward: 1.67
episode: 910	score: 0.0	memory length: 100000 epsilon: 0.8695199800028326
steps: 122	lr: 0.0001	evaluation reward: 1.67
episode: 911	score: 3.0	memory length: 100000 epsilon: 0.8690309200028432
steps: 247	lr: 0.0001	evaluation reward: 1.7
episode: 912	score: 0.0	memory length: 100000 epsilon: 0.8687893600028485
steps: 122	lr: 0.0001	evaluation reward: 1.69
episode: 913	score: 0.0	memory length: 100000 epsilon: 0.8685458200028537
steps: 123	lr: 0.0001	evaluation reward: 1.66
episode: 914	score: 2.0	memory length: 100000 epsilon: 0.8681102200028632
steps: 220	lr: 0.0001	evaluation reward: 1.66
episode: 915	score: 0.0	memory length: 100000 epsilon: 0.8678666800028685
steps: 123	lr: 0.0001	evaluation reward: 1.64
episode: 916	score: 3.0	memory length: 100000 epsilon: 0.8673776200028791
steps: 247	lr: 0.0001	evaluation reward: 1.65
episode: 917	score: 2.0	memory length: 100000 epsilon: 0.8669855800028876
steps: 198	lr: 0.0001	evaluation reward: 1.65
episode: 918	score: 3.0	memory length: 100000 epsilon: 0.8665400800028973
steps: 225	lr: 0.0001	evaluation reward: 1.67
episode: 919	score: 0.0	memory length: 100000 epsilon: 0.8662985200029025
steps: 122	lr: 0.0001	evaluation reward: 1.65
episode: 920	score: 2.0	memory length: 100000 epsilon: 0.865908460002911
steps: 197	lr: 0.0001	evaluation reward: 1.64
episode: 921	score: 2.0	memory length: 100000 epsilon: 0.8654807800029203
steps: 216	lr: 0.0001	evaluation reward: 1.64
episode: 922	score: 4.0	memory length: 100000 epsilon: 0.864851140002934
steps: 318	lr: 0.0001	evaluation reward: 1.65
episode: 923	score: 2.0	memory length: 100000 epsilon: 0.8644175200029434

steps: 219	lr: 0.0001	evaluation reward: 1.67
episode: 924	score: 0.0	memory length: 100000 epsilon: 0.8641759600029486
steps: 122	lr: 0.0001	evaluation reward: 1.67
episode: 925	score: 2.0	memory length: 100000 epsilon: 0.8637859000029571
steps: 197	lr: 0.0001	evaluation reward: 1.69
episode: 926	score: 1.0	memory length: 100000 epsilon: 0.8634512800029643
steps: 169	lr: 0.0001	evaluation reward: 1.67
episode: 927	score: 3.0	memory length: 100000 epsilon: 0.862958260002975
steps: 249	lr: 0.0001	evaluation reward: 1.66
episode: 928	score: 0.0	memory length: 100000 epsilon: 0.8627147200029803
steps: 123	lr: 0.0001	evaluation reward: 1.63
episode: 929	score: 0.0	memory length: 100000 epsilon: 0.8624711800029856
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 930	score: 1.0	memory length: 100000 epsilon: 0.8621385400029928
steps: 168	lr: 0.0001	evaluation reward: 1.57
episode: 931	score: 1.0	memory length: 100000 epsilon: 0.8617979800030002
steps: 172	lr: 0.0001	evaluation reward: 1.57
episode: 932	score: 3.0	memory length: 100000 epsilon: 0.8612713000030117
steps: 266	lr: 0.0001	evaluation reward: 1.58
episode: 933	score: 3.0	memory length: 100000 epsilon: 0.8608258000030213
steps: 225	lr: 0.0001	evaluation reward: 1.6
episode: 934	score: 3.0	memory length: 100000 epsilon: 0.860378320003031
steps: 226	lr: 0.0001	evaluation reward: 1.61
episode: 935	score: 4.0	memory length: 100000 epsilon: 0.8597566000030445
steps: 314	lr: 0.0001	evaluation reward: 1.65
episode: 936	score: 1.0	memory length: 100000 epsilon: 0.8594200000030519
steps: 170	lr: 0.0001	evaluation reward: 1.66
episode: 937	score: 1.0	memory length: 100000 epsilon: 0.8590873600030591
steps: 168	lr: 0.0001	evaluation reward: 1.66
episode: 938	score: 0.0	memory length: 100000 epsilon: 0.8588438200030644
steps: 123	lr: 0.0001	evaluation reward: 1.66
episode: 939	score: 4.0	memory length: 100000 epsilon: 0.8582933800030763
steps: 278	lr: 0.0001	evaluation reward: 1.69
episode: 940	score: 1.0	memory length: 100000 epsilon: 0.8579963800030828
steps: 150	lr: 0.0001	evaluation reward: 1.66
episode: 941	score: 1.0	memory length: 100000 epsilon: 0.85766176000309
steps: 169	lr: 0.0001	evaluation reward: 1.67
episode: 942	score: 3.0	memory length: 100000 epsilon: 0.8571350800031015
steps: 266	lr: 0.0001	evaluation reward: 1.68
episode: 943	score: 0.0	memory length: 100000 epsilon: 0.8568915400031067
steps: 123	lr: 0.0001	evaluation reward: 1.68
episode: 944	score: 3.0	memory length: 100000 epsilon: 0.8564064400031173
steps: 245	lr: 0.0001	evaluation reward: 1.7
episode: 945	score: 2.0	memory length: 100000 epsilon: 0.8560163800031257
steps: 197	lr: 0.0001	evaluation reward: 1.72
episode: 946	score: 1.0	memory length: 100000 epsilon: 0.855679780003133
steps: 170	lr: 0.0001	evaluation reward: 1.71
episode: 947	score: 0.0	memory length: 100000 epsilon: 0.8554362400031383

steps: 123	lr: 0.0001	evaluation reward: 1.67
episode: 948	score: 4.0	memory length: 100000 epsilon: 0.854808580003152
steps: 317	lr: 0.0001	evaluation reward: 1.71
episode: 949	score: 2.0	memory length: 100000 epsilon: 0.8544165400031605
steps: 198	lr: 0.0001	evaluation reward: 1.7
episode: 950	score: 2.0	memory length: 100000 epsilon: 0.8539849000031698
steps: 218	lr: 0.0001	evaluation reward: 1.72
episode: 951	score: 1.0	memory length: 100000 epsilon: 0.8536859200031763
steps: 151	lr: 0.0001	evaluation reward: 1.72
episode: 952	score: 3.0	memory length: 100000 epsilon: 0.853238440003186
steps: 226	lr: 0.0001	evaluation reward: 1.73
episode: 953	score: 2.0	memory length: 100000 epsilon: 0.8528424400031946
steps: 200	lr: 0.0001	evaluation reward: 1.72
episode: 954	score: 2.0	memory length: 100000 epsilon: 0.8524504000032032
steps: 198	lr: 0.0001	evaluation reward: 1.72
episode: 955	score: 1.0	memory length: 100000 epsilon: 0.8521118200032105
steps: 171	lr: 0.0001	evaluation reward: 1.7
episode: 956	score: 0.0	memory length: 100000 epsilon: 0.8518682800032158
steps: 123	lr: 0.0001	evaluation reward: 1.67
episode: 957	score: 1.0	memory length: 100000 epsilon: 0.851535640003223
steps: 168	lr: 0.0001	evaluation reward: 1.68
episode: 958	score: 1.0	memory length: 100000 epsilon: 0.8511970600032304
steps: 171	lr: 0.0001	evaluation reward: 1.68
episode: 959	score: 2.0	memory length: 100000 epsilon: 0.8507654200032397
steps: 218	lr: 0.0001	evaluation reward: 1.67
episode: 960	score: 1.0	memory length: 100000 epsilon: 0.8504664400032462
steps: 151	lr: 0.0001	evaluation reward: 1.66
episode: 961	score: 2.0	memory length: 100000 epsilon: 0.8500387600032555
steps: 216	lr: 0.0001	evaluation reward: 1.66
episode: 962	score: 0.0	memory length: 100000 epsilon: 0.8497972000032608
steps: 122	lr: 0.0001	evaluation reward: 1.64
episode: 963	score: 2.0	memory length: 100000 epsilon: 0.8494031800032693
steps: 199	lr: 0.0001	evaluation reward: 1.65
episode: 964	score: 0.0	memory length: 100000 epsilon: 0.8491596400032746
steps: 123	lr: 0.0001	evaluation reward: 1.63
episode: 965	score: 4.0	memory length: 100000 epsilon: 0.8485814800032871
steps: 292	lr: 0.0001	evaluation reward: 1.64
episode: 966	score: 0.0	memory length: 100000 epsilon: 0.8483399200032924
steps: 122	lr: 0.0001	evaluation reward: 1.63
episode: 967	score: 2.0	memory length: 100000 epsilon: 0.8479498600033009
steps: 197	lr: 0.0001	evaluation reward: 1.63
episode: 968	score: 0.0	memory length: 100000 epsilon: 0.8477063200033061
steps: 123	lr: 0.0001	evaluation reward: 1.62
episode: 969	score: 3.0	memory length: 100000 epsilon: 0.8472192400033167
steps: 246	lr: 0.0001	evaluation reward: 1.65
episode: 970	score: 3.0	memory length: 100000 epsilon: 0.8466945400033281
steps: 265	lr: 0.0001	evaluation reward: 1.68
episode: 971	score: 2.0	memory length: 100000 epsilon: 0.8462629000033375

steps: 218	lr: 0.0001	evaluation reward: 1.68
episode: 972	score: 0.0	memory length: 100000 epsilon: 0.8460193600033428
steps: 123	lr: 0.0001	evaluation reward: 1.67
episode: 973	score: 0.0	memory length: 100000 epsilon: 0.845777800003348
steps: 122	lr: 0.0001	evaluation reward: 1.66
episode: 974	score: 2.0	memory length: 100000 epsilon: 0.8453857600033565
steps: 198	lr: 0.0001	evaluation reward: 1.68
episode: 975	score: 1.0	memory length: 100000 epsilon: 0.845086780003363
steps: 151	lr: 0.0001	evaluation reward: 1.67
episode: 976	score: 3.0	memory length: 100000 epsilon: 0.8445977200033736
steps: 247	lr: 0.0001	evaluation reward: 1.69
episode: 977	score: 2.0	memory length: 100000 epsilon: 0.844168060003383
steps: 217	lr: 0.0001	evaluation reward: 1.67
episode: 978	score: 2.0	memory length: 100000 epsilon: 0.8437760200033915
steps: 198	lr: 0.0001	evaluation reward: 1.66
episode: 979	score: 2.0	memory length: 100000 epsilon: 0.8433839800034
steps: 198	lr: 0.0001	evaluation reward: 1.65
episode: 980	score: 3.0	memory length: 100000 epsilon: 0.8428929400034106
steps: 248	lr: 0.0001	evaluation reward: 1.67
episode: 981	score: 2.0	memory length: 100000 epsilon: 0.8425028800034191
steps: 197	lr: 0.0001	evaluation reward: 1.65
episode: 982	score: 0.0	memory length: 100000 epsilon: 0.8422593400034244
steps: 123	lr: 0.0001	evaluation reward: 1.6
episode: 983	score: 1.0	memory length: 100000 epsilon: 0.8419247200034317
steps: 169	lr: 0.0001	evaluation reward: 1.59
episode: 984	score: 6.0	memory length: 100000 epsilon: 0.8412376600034466
steps: 347	lr: 0.0001	evaluation reward: 1.63
episode: 985	score: 5.0	memory length: 100000 epsilon: 0.840667420003459
steps: 288	lr: 0.0001	evaluation reward: 1.68
episode: 986	score: 0.0	memory length: 100000 epsilon: 0.8404258600034642
steps: 122	lr: 0.0001	evaluation reward: 1.67
episode: 987	score: 3.0	memory length: 100000 epsilon: 0.8399011600034756
steps: 265	lr: 0.0001	evaluation reward: 1.67
episode: 988	score: 0.0	memory length: 100000 epsilon: 0.8396576200034809
steps: 123	lr: 0.0001	evaluation reward: 1.66
episode: 989	score: 3.0	memory length: 100000 epsilon: 0.8391705400034914
steps: 246	lr: 0.0001	evaluation reward: 1.68
episode: 990	score: 1.0	memory length: 100000 epsilon: 0.838867600003498
steps: 153	lr: 0.0001	evaluation reward: 1.69
episode: 991	score: 4.0	memory length: 100000 epsilon: 0.8382399400035117
steps: 317	lr: 0.0001	evaluation reward: 1.7
episode: 992	score: 0.0	memory length: 100000 epsilon: 0.8379964000035169
steps: 123	lr: 0.0001	evaluation reward: 1.69
episode: 993	score: 0.0	memory length: 100000 epsilon: 0.8377528600035222
steps: 123	lr: 0.0001	evaluation reward: 1.64
episode: 994	score: 0.0	memory length: 100000 epsilon: 0.8375093200035275
steps: 123	lr: 0.0001	evaluation reward: 1.62
episode: 995	score: 3.0	memory length: 100000 epsilon: 0.8370222400035381

steps: 246	lr: 0.0001	evaluation reward: 1.64	
episode: 996	score: 1.0	memory length: 100000	epsilon: 0.8366836600035454
steps: 171	lr: 0.0001	evaluation reward: 1.62	
episode: 997	score: 2.0	memory length: 100000	epsilon: 0.8362520200035548
steps: 218	lr: 0.0001	evaluation reward: 1.63	
episode: 998	score: 1.0	memory length: 100000	epsilon: 0.8359530400035613
steps: 151	lr: 0.0001	evaluation reward: 1.64	
episode: 999	score: 2.0	memory length: 100000	epsilon: 0.8355194200035707
steps: 219	lr: 0.0001	evaluation reward: 1.65	
episode: 1000	score: 1.0	memory length: 100000	epsilon: 0.835184800003578
steps: 169	lr: 0.0001	evaluation reward: 1.65	
episode: 1001	score: 3.0	memory length: 100000	epsilon: 0.8346581200035894
steps: 266	lr: 0.0001	evaluation reward: 1.66	
episode: 1002	score: 1.0	memory length: 100000	epsilon: 0.8343215200035967
steps: 170	lr: 0.0001	evaluation reward: 1.64	
episode: 1003	score: 0.0	memory length: 100000	epsilon: 0.834077980003602
steps: 123	lr: 0.0001	evaluation reward: 1.63	
episode: 1004	score: 0.0	memory length: 100000	epsilon: 0.8338344400036073
steps: 123	lr: 0.0001	evaluation reward: 1.63	
episode: 1005	score: 1.0	memory length: 100000	epsilon: 0.8334998200036146
steps: 169	lr: 0.0001	evaluation reward: 1.64	
episode: 1006	score: 2.0	memory length: 100000	epsilon: 0.8331454000036222
steps: 179	lr: 0.0001	evaluation reward: 1.65	
episode: 1007	score: 2.0	memory length: 100000	epsilon: 0.8327533600036308
steps: 198	lr: 0.0001	evaluation reward: 1.67	
episode: 1008	score: 1.0	memory length: 100000	epsilon: 0.8324147800036381
steps: 171	lr: 0.0001	evaluation reward: 1.63	
episode: 1009	score: 0.0	memory length: 100000	epsilon: 0.8321712400036434
steps: 123	lr: 0.0001	evaluation reward: 1.61	
episode: 1010	score: 0.0	memory length: 100000	epsilon: 0.8319277000036487
steps: 123	lr: 0.0001	evaluation reward: 1.61	
episode: 1011	score: 2.0	memory length: 100000	epsilon: 0.8314940800036581
steps: 219	lr: 0.0001	evaluation reward: 1.6	
episode: 1012	score: 1.0	memory length: 100000	epsilon: 0.8311614400036653
steps: 168	lr: 0.0001	evaluation reward: 1.61	
episode: 1013	score: 2.0	memory length: 100000	epsilon: 0.8307317800036746
steps: 217	lr: 0.0001	evaluation reward: 1.63	
episode: 1014	score: 2.0	memory length: 100000	epsilon: 0.8303377600036832
steps: 199	lr: 0.0001	evaluation reward: 1.63	
episode: 1015	score: 1.0	memory length: 100000	epsilon: 0.8300407600036896
steps: 150	lr: 0.0001	evaluation reward: 1.64	
episode: 1016	score: 1.0	memory length: 100000	epsilon: 0.8297061400036969
steps: 169	lr: 0.0001	evaluation reward: 1.62	
episode: 1017	score: 0.0	memory length: 100000	epsilon: 0.8294645800037022
steps: 122	lr: 0.0001	evaluation reward: 1.6	
episode: 1018	score: 2.0	memory length: 100000	epsilon: 0.8290289800037116
steps: 220	lr: 0.0001	evaluation reward: 1.59	
episode: 1019	score: 2.0	memory length: 100000	epsilon: 0.8286369400037201

steps: 198	lr: 0.0001	evaluation reward: 1.61	
episode: 1020	score: 0.0	memory length: 100000	epsilon: 0.8283953800037254
steps: 122	lr: 0.0001	evaluation reward: 1.59	
episode: 1021	score: 1.0	memory length: 100000	epsilon: 0.8280627400037326
steps: 168	lr: 0.0001	evaluation reward: 1.58	
episode: 1022	score: 0.0	memory length: 100000	epsilon: 0.8278192000037379
steps: 123	lr: 0.0001	evaluation reward: 1.54	
episode: 1023	score: 1.0	memory length: 100000	epsilon: 0.8274845800037451
steps: 169	lr: 0.0001	evaluation reward: 1.53	
episode: 1024	score: 1.0	memory length: 100000	epsilon: 0.8271875800037516
steps: 150	lr: 0.0001	evaluation reward: 1.54	
episode: 1025	score: 1.0	memory length: 100000	epsilon: 0.826847020003759
steps: 172	lr: 0.0001	evaluation reward: 1.53	
episode: 1026	score: 5.0	memory length: 100000	epsilon: 0.8262886600037711
steps: 282	lr: 0.0001	evaluation reward: 1.57	
episode: 1027	score: 2.0	memory length: 100000	epsilon: 0.8258629600037803
steps: 215	lr: 0.0001	evaluation reward: 1.56	
episode: 1028	score: 5.0	memory length: 100000	epsilon: 0.825235300003794
steps: 317	lr: 0.0001	evaluation reward: 1.61	
episode: 1029	score: 2.0	memory length: 100000	epsilon: 0.8248432600038025
steps: 198	lr: 0.0001	evaluation reward: 1.63	
episode: 1030	score: 1.0	memory length: 100000	epsilon: 0.824544280003809
steps: 151	lr: 0.0001	evaluation reward: 1.63	
episode: 1031	score: 1.0	memory length: 100000	epsilon: 0.8242076800038163
steps: 170	lr: 0.0001	evaluation reward: 1.63	
episode: 1032	score: 3.0	memory length: 100000	epsilon: 0.8236790200038278
steps: 267	lr: 0.0001	evaluation reward: 1.63	
episode: 1033	score: 1.0	memory length: 100000	epsilon: 0.8233384600038351
steps: 172	lr: 0.0001	evaluation reward: 1.61	
episode: 1034	score: 0.0	memory length: 100000	epsilon: 0.8230949200038404
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 1035	score: 1.0	memory length: 100000	epsilon: 0.8227959400038469
steps: 151	lr: 0.0001	evaluation reward: 1.55	
episode: 1036	score: 3.0	memory length: 100000	epsilon: 0.8223425200038568
steps: 229	lr: 0.0001	evaluation reward: 1.57	
episode: 1037	score: 1.0	memory length: 100000	epsilon: 0.8220435400038633
steps: 151	lr: 0.0001	evaluation reward: 1.57	
episode: 1038	score: 4.0	memory length: 100000	epsilon: 0.8214594400038759
steps: 295	lr: 0.0001	evaluation reward: 1.61	
episode: 1039	score: 2.0	memory length: 100000	epsilon: 0.8211050200038836
steps: 179	lr: 0.0001	evaluation reward: 1.59	
episode: 1040	score: 0.0	memory length: 100000	epsilon: 0.8208634600038889
steps: 122	lr: 0.0001	evaluation reward: 1.58	
episode: 1041	score: 4.0	memory length: 100000	epsilon: 0.8202773800039016
steps: 296	lr: 0.0001	evaluation reward: 1.61	
episode: 1042	score: 2.0	memory length: 100000	epsilon: 0.8198853400039101
steps: 198	lr: 0.0001	evaluation reward: 1.6	
episode: 1043	score: 0.0	memory length: 100000	epsilon: 0.8196418000039154

steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 1044	score: 3.0	memory length: 100000	epsilon: 0.8191190800039267
steps: 264	lr: 0.0001	evaluation reward: 1.6	
episode: 1045	score: 0.0	memory length: 100000	epsilon: 0.818877520003932
steps: 122	lr: 0.0001	evaluation reward: 1.58	
episode: 1046	score: 1.0	memory length: 100000	epsilon: 0.8185785400039385
steps: 151	lr: 0.0001	evaluation reward: 1.58	
episode: 1047	score: 0.0	memory length: 100000	epsilon: 0.8183350000039438
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 1048	score: 2.0	memory length: 100000	epsilon: 0.8179390000039524
steps: 200	lr: 0.0001	evaluation reward: 1.56	
episode: 1049	score: 5.0	memory length: 100000	epsilon: 0.8172578800039672
steps: 344	lr: 0.0001	evaluation reward: 1.59	
episode: 1050	score: 0.0	memory length: 100000	epsilon: 0.8170143400039724
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 1051	score: 0.0	memory length: 100000	epsilon: 0.8167708000039777
steps: 123	lr: 0.0001	evaluation reward: 1.56	
episode: 1052	score: 1.0	memory length: 100000	epsilon: 0.8164718200039842
steps: 151	lr: 0.0001	evaluation reward: 1.54	
episode: 1053	score: 0.0	memory length: 100000	epsilon: 0.8162282800039895
steps: 123	lr: 0.0001	evaluation reward: 1.52	
episode: 1054	score: 0.0	memory length: 100000	epsilon: 0.8159867200039947
steps: 122	lr: 0.0001	evaluation reward: 1.5	
episode: 1055	score: 3.0	memory length: 100000	epsilon: 0.8155392400040045
steps: 226	lr: 0.0001	evaluation reward: 1.52	
episode: 1056	score: 4.0	memory length: 100000	epsilon: 0.8149591000040171
steps: 293	lr: 0.0001	evaluation reward: 1.56	
episode: 1057	score: 2.0	memory length: 100000	epsilon: 0.8145235000040265
steps: 220	lr: 0.0001	evaluation reward: 1.57	
episode: 1058	score: 1.0	memory length: 100000	epsilon: 0.8141908600040337
steps: 168	lr: 0.0001	evaluation reward: 1.57	
episode: 1059	score: 1.0	memory length: 100000	epsilon: 0.813856240004041
steps: 169	lr: 0.0001	evaluation reward: 1.56	
episode: 1060	score: 2.0	memory length: 100000	epsilon: 0.8134642000040495
steps: 198	lr: 0.0001	evaluation reward: 1.57	
episode: 1061	score: 8.0	memory length: 100000	epsilon: 0.8125553800040692
steps: 459	lr: 0.0001	evaluation reward: 1.63	
episode: 1062	score: 4.0	memory length: 100000	epsilon: 0.8119811800040817
steps: 290	lr: 0.0001	evaluation reward: 1.67	
episode: 1063	score: 2.0	memory length: 100000	epsilon: 0.8115911200040902
steps: 197	lr: 0.0001	evaluation reward: 1.67	
episode: 1064	score: 2.0	memory length: 100000	epsilon: 0.8111594800040995
steps: 218	lr: 0.0001	evaluation reward: 1.69	
episode: 1065	score: 3.0	memory length: 100000	epsilon: 0.8106743800041101
steps: 245	lr: 0.0001	evaluation reward: 1.68	
episode: 1066	score: 1.0	memory length: 100000	epsilon: 0.8103734200041166
steps: 152	lr: 0.0001	evaluation reward: 1.69	
episode: 1067	score: 2.0	memory length: 100000	epsilon: 0.8100190000041243

steps: 179	lr: 0.0001	evaluation reward: 1.69	
episode: 1068	score: 2.0	memory length: 100000	epsilon: 0.8096289400041328
steps: 197	lr: 0.0001	evaluation reward: 1.71	
episode: 1069	score: 2.0	memory length: 100000	epsilon: 0.809201260004142
steps: 216	lr: 0.0001	evaluation reward: 1.7	
episode: 1070	score: 2.0	memory length: 100000	epsilon: 0.8088409000041499
steps: 182	lr: 0.0001	evaluation reward: 1.69	
episode: 1071	score: 2.0	memory length: 100000	epsilon: 0.8084508400041583
steps: 197	lr: 0.0001	evaluation reward: 1.69	
episode: 1072	score: 0.0	memory length: 100000	epsilon: 0.8082073000041636
steps: 123	lr: 0.0001	evaluation reward: 1.69	
episode: 1073	score: 3.0	memory length: 100000	epsilon: 0.8076727000041752
steps: 270	lr: 0.0001	evaluation reward: 1.72	
episode: 1074	score: 1.0	memory length: 100000	epsilon: 0.8073400600041825
steps: 168	lr: 0.0001	evaluation reward: 1.71	
episode: 1075	score: 2.0	memory length: 100000	epsilon: 0.8069500000041909
steps: 197	lr: 0.0001	evaluation reward: 1.72	
episode: 1076	score: 1.0	memory length: 100000	epsilon: 0.8066153800041982
steps: 169	lr: 0.0001	evaluation reward: 1.7	
episode: 1077	score: 0.0	memory length: 100000	epsilon: 0.8063718400042035
steps: 123	lr: 0.0001	evaluation reward: 1.68	
episode: 1078	score: 0.0	memory length: 100000	epsilon: 0.8061283000042088
steps: 123	lr: 0.0001	evaluation reward: 1.66	
episode: 1079	score: 2.0	memory length: 100000	epsilon: 0.8056927000042182
steps: 220	lr: 0.0001	evaluation reward: 1.66	
episode: 1080	score: 3.0	memory length: 100000	epsilon: 0.8052769000042272
steps: 210	lr: 0.0001	evaluation reward: 1.66	
episode: 1081	score: 1.0	memory length: 100000	epsilon: 0.8049403000042346
steps: 170	lr: 0.0001	evaluation reward: 1.65	
episode: 1082	score: 0.0	memory length: 100000	epsilon: 0.8046967600042398
steps: 123	lr: 0.0001	evaluation reward: 1.65	
episode: 1083	score: 1.0	memory length: 100000	epsilon: 0.8043601600042471
steps: 170	lr: 0.0001	evaluation reward: 1.65	
episode: 1084	score: 3.0	memory length: 100000	epsilon: 0.8038730800042577
steps: 246	lr: 0.0001	evaluation reward: 1.62	
episode: 1085	score: 3.0	memory length: 100000	epsilon: 0.8033384800042693
steps: 270	lr: 0.0001	evaluation reward: 1.6	
episode: 1086	score: 1.0	memory length: 100000	epsilon: 0.8029979200042767
steps: 172	lr: 0.0001	evaluation reward: 1.61	
episode: 1087	score: 0.0	memory length: 100000	epsilon: 0.802754380004282
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 1088	score: 2.0	memory length: 100000	epsilon: 0.8023643200042905
steps: 197	lr: 0.0001	evaluation reward: 1.6	
episode: 1089	score: 3.0	memory length: 100000	epsilon: 0.8019128800043003
steps: 228	lr: 4e-05	evaluation reward: 1.6	
episode: 1090	score: 4.0	memory length: 100000	epsilon: 0.8013664000043121
steps: 276	lr: 4e-05	evaluation reward: 1.63	
episode: 1091	score: 1.0	memory length: 100000	epsilon: 0.8010337600043194

steps: 168	lr: 4e-05	evaluation reward: 1.6	
episode: 1092	score: 0.0	memory length: 100000	epsilon: 0.8007902200043246
steps: 123	lr: 4e-05	evaluation reward: 1.6	
episode: 1093	score: 6.0	memory length: 100000	epsilon: 0.7999645600043426
steps: 417	lr: 4e-05	evaluation reward: 1.66	
episode: 1094	score: 0.0	memory length: 100000	epsilon: 0.7997210200043479
steps: 123	lr: 4e-05	evaluation reward: 1.66	
episode: 1095	score: 2.0	memory length: 100000	epsilon: 0.7993309600043563
steps: 197	lr: 4e-05	evaluation reward: 1.65	
episode: 1096	score: 0.0	memory length: 100000	epsilon: 0.7990874200043616
steps: 123	lr: 4e-05	evaluation reward: 1.64	
episode: 1097	score: 1.0	memory length: 100000	epsilon: 0.7987528000043689
steps: 169	lr: 4e-05	evaluation reward: 1.63	
episode: 1098	score: 0.0	memory length: 100000	epsilon: 0.7985092600043742
steps: 123	lr: 4e-05	evaluation reward: 1.62	
episode: 1099	score: 1.0	memory length: 100000	epsilon: 0.7982122600043806
steps: 150	lr: 4e-05	evaluation reward: 1.61	
episode: 1100	score: 3.0	memory length: 100000	epsilon: 0.7977608200043904
steps: 228	lr: 4e-05	evaluation reward: 1.63	
episode: 1101	score: 1.0	memory length: 100000	epsilon: 0.7974262000043977
steps: 169	lr: 4e-05	evaluation reward: 1.61	
episode: 1102	score: 0.0	memory length: 100000	epsilon: 0.7971846400044029
steps: 122	lr: 4e-05	evaluation reward: 1.6	
episode: 1103	score: 1.0	memory length: 100000	epsilon: 0.7968856600044094
steps: 151	lr: 4e-05	evaluation reward: 1.61	
episode: 1104	score: 0.0	memory length: 100000	epsilon: 0.7966421200044147
steps: 123	lr: 4e-05	evaluation reward: 1.61	
episode: 1105	score: 2.0	memory length: 100000	epsilon: 0.7962500800044232
steps: 198	lr: 4e-05	evaluation reward: 1.62	
episode: 1106	score: 1.0	memory length: 100000	epsilon: 0.7959530800044297
steps: 150	lr: 4e-05	evaluation reward: 1.61	
episode: 1107	score: 2.0	memory length: 100000	epsilon: 0.795523420004439
steps: 217	lr: 4e-05	evaluation reward: 1.61	
episode: 1108	score: 2.0	memory length: 100000	epsilon: 0.7950917800044484
steps: 218	lr: 4e-05	evaluation reward: 1.62	
episode: 1109	score: 2.0	memory length: 100000	epsilon: 0.7946601400044577
steps: 218	lr: 4e-05	evaluation reward: 1.64	
episode: 1110	score: 2.0	memory length: 100000	epsilon: 0.7942225600044672
steps: 221	lr: 4e-05	evaluation reward: 1.66	
episode: 1111	score: 1.0	memory length: 100000	epsilon: 0.7938859600044745
steps: 170	lr: 4e-05	evaluation reward: 1.65	
episode: 1112	score: 3.0	memory length: 100000	epsilon: 0.7934404600044842
steps: 225	lr: 4e-05	evaluation reward: 1.67	
episode: 1113	score: 1.0	memory length: 100000	epsilon: 0.7931058400044915
steps: 169	lr: 4e-05	evaluation reward: 1.66	
episode: 1114	score: 3.0	memory length: 100000	epsilon: 0.7926128200045022
steps: 249	lr: 4e-05	evaluation reward: 1.67	
episode: 1115	score: 3.0	memory length: 100000	epsilon: 0.7921257400045127

steps: 246	lr: 4e-05	evaluation reward: 1.69	
episode: 1116	score: 2.0	memory length: 100000	epsilon: 0.7917317200045213
steps: 199	lr: 4e-05	evaluation reward: 1.7	
episode: 1117	score: 1.0	memory length: 100000	epsilon: 0.7914327400045278
steps: 151	lr: 4e-05	evaluation reward: 1.71	
episode: 1118	score: 0.0	memory length: 100000	epsilon: 0.7911892000045331
steps: 123	lr: 4e-05	evaluation reward: 1.69	
episode: 1119	score: 2.0	memory length: 100000	epsilon: 0.7907575600045424
steps: 218	lr: 4e-05	evaluation reward: 1.69	
episode: 1120	score: 2.0	memory length: 100000	epsilon: 0.7903675000045509
steps: 197	lr: 4e-05	evaluation reward: 1.71	
episode: 1121	score: 2.0	memory length: 100000	epsilon: 0.7899754600045594
steps: 198	lr: 4e-05	evaluation reward: 1.72	
episode: 1122	score: 0.0	memory length: 100000	epsilon: 0.7897319200045647
steps: 123	lr: 4e-05	evaluation reward: 1.72	
episode: 1123	score: 3.0	memory length: 100000	epsilon: 0.7892428600045753
steps: 247	lr: 4e-05	evaluation reward: 1.74	
episode: 1124	score: 0.0	memory length: 100000	epsilon: 0.7889993200045806
steps: 123	lr: 4e-05	evaluation reward: 1.73	
episode: 1125	score: 3.0	memory length: 100000	epsilon: 0.7885518400045903
steps: 226	lr: 4e-05	evaluation reward: 1.75	
episode: 1126	score: 3.0	memory length: 100000	epsilon: 0.7880687200046008
steps: 244	lr: 4e-05	evaluation reward: 1.73	
episode: 1127	score: 2.0	memory length: 100000	epsilon: 0.7876786600046093
steps: 197	lr: 4e-05	evaluation reward: 1.73	
episode: 1128	score: 1.0	memory length: 100000	epsilon: 0.7873440400046166
steps: 169	lr: 4e-05	evaluation reward: 1.69	
episode: 1129	score: 0.0	memory length: 100000	epsilon: 0.7871005000046218
steps: 123	lr: 4e-05	evaluation reward: 1.67	
episode: 1130	score: 2.0	memory length: 100000	epsilon: 0.7867025200046305
steps: 201	lr: 4e-05	evaluation reward: 1.68	
episode: 1131	score: 0.0	memory length: 100000	epsilon: 0.7864609600046357
steps: 122	lr: 4e-05	evaluation reward: 1.67	
episode: 1132	score: 5.0	memory length: 100000	epsilon: 0.7858649800046487
steps: 301	lr: 4e-05	evaluation reward: 1.69	
episode: 1133	score: 0.0	memory length: 100000	epsilon: 0.7856234200046539
steps: 122	lr: 4e-05	evaluation reward: 1.68	
episode: 1134	score: 0.0	memory length: 100000	epsilon: 0.7853818600046591
steps: 122	lr: 4e-05	evaluation reward: 1.68	
episode: 1135	score: 0.0	memory length: 100000	epsilon: 0.7851383200046644
steps: 123	lr: 4e-05	evaluation reward: 1.67	
episode: 1136	score: 3.0	memory length: 100000	epsilon: 0.784649260004675
steps: 247	lr: 4e-05	evaluation reward: 1.67	
episode: 1137	score: 2.0	memory length: 100000	epsilon: 0.7842592000046835
steps: 197	lr: 4e-05	evaluation reward: 1.68	
episode: 1138	score: 0.0	memory length: 100000	epsilon: 0.7840156600046888
steps: 123	lr: 4e-05	evaluation reward: 1.64	
episode: 1139	score: 0.0	memory length: 100000	epsilon: 0.783774100004694

steps: 122	lr: 4e-05	evaluation reward: 1.62	
episode: 1140	score: 3.0	memory length: 100000	epsilon: 0.7833286000047037
steps: 225	lr: 4e-05	evaluation reward: 1.65	
episode: 1141	score: 2.0	memory length: 100000	epsilon: 0.7829385400047122
steps: 197	lr: 4e-05	evaluation reward: 1.63	
episode: 1142	score: 1.0	memory length: 100000	epsilon: 0.7826415400047186
steps: 150	lr: 4e-05	evaluation reward: 1.62	
episode: 1143	score: 2.0	memory length: 100000	epsilon: 0.7822831600047264
steps: 181	lr: 4e-05	evaluation reward: 1.64	
episode: 1144	score: 1.0	memory length: 100000	epsilon: 0.7819465600047337
steps: 170	lr: 4e-05	evaluation reward: 1.62	
episode: 1145	score: 3.0	memory length: 100000	epsilon: 0.7814990800047434
steps: 226	lr: 4e-05	evaluation reward: 1.65	
episode: 1146	score: 0.0	memory length: 100000	epsilon: 0.7812555400047487
steps: 123	lr: 4e-05	evaluation reward: 1.64	
episode: 1147	score: 2.0	memory length: 100000	epsilon: 0.7808654800047572
steps: 197	lr: 4e-05	evaluation reward: 1.66	
episode: 1148	score: 3.0	memory length: 100000	epsilon: 0.780412060004767
steps: 229	lr: 4e-05	evaluation reward: 1.67	
episode: 1149	score: 0.0	memory length: 100000	epsilon: 0.7801685200047723
steps: 123	lr: 4e-05	evaluation reward: 1.62	
episode: 1150	score: 0.0	memory length: 100000	epsilon: 0.7799269600047776
steps: 122	lr: 4e-05	evaluation reward: 1.62	
episode: 1151	score: 0.0	memory length: 100000	epsilon: 0.7796834200047829
steps: 123	lr: 4e-05	evaluation reward: 1.62	
episode: 1152	score: 3.0	memory length: 100000	epsilon: 0.7792339600047926
steps: 227	lr: 4e-05	evaluation reward: 1.64	
episode: 1153	score: 3.0	memory length: 100000	epsilon: 0.778707280004804
steps: 266	lr: 4e-05	evaluation reward: 1.67	
episode: 1154	score: 0.0	memory length: 100000	epsilon: 0.7784637400048093
steps: 123	lr: 4e-05	evaluation reward: 1.67	
episode: 1155	score: 0.0	memory length: 100000	epsilon: 0.7782221800048146
steps: 122	lr: 4e-05	evaluation reward: 1.64	
episode: 1156	score: 3.0	memory length: 100000	epsilon: 0.7777370800048251
steps: 245	lr: 4e-05	evaluation reward: 1.63	
episode: 1157	score: 1.0	memory length: 100000	epsilon: 0.7774381000048316
steps: 151	lr: 4e-05	evaluation reward: 1.62	
episode: 1158	score: 0.0	memory length: 100000	epsilon: 0.7771945600048369
steps: 123	lr: 4e-05	evaluation reward: 1.61	
episode: 1159	score: 2.0	memory length: 100000	epsilon: 0.7767629200048463
steps: 218	lr: 4e-05	evaluation reward: 1.62	
episode: 1160	score: 0.0	memory length: 100000	epsilon: 0.7765193800048515
steps: 123	lr: 4e-05	evaluation reward: 1.6	
episode: 1161	score: 1.0	memory length: 100000	epsilon: 0.776222380004858
steps: 150	lr: 4e-05	evaluation reward: 1.53	
episode: 1162	score: 2.0	memory length: 100000	epsilon: 0.7758303400048665
steps: 198	lr: 4e-05	evaluation reward: 1.51	
episode: 1163	score: 2.0	memory length: 100000	epsilon: 0.775440280004875

steps: 197	lr: 4e-05	evaluation reward: 1.51	
episode: 1164	score: 5.0	memory length: 100000	epsilon: 0.7747967800048889
steps: 325	lr: 4e-05	evaluation reward: 1.54	
episode: 1165	score: 1.0	memory length: 100000	epsilon: 0.7744978000048954
steps: 151	lr: 4e-05	evaluation reward: 1.52	
episode: 1166	score: 3.0	memory length: 100000	epsilon: 0.7740146800049059
steps: 244	lr: 4e-05	evaluation reward: 1.54	
episode: 1167	score: 0.0	memory length: 100000	epsilon: 0.7737731200049112
steps: 122	lr: 4e-05	evaluation reward: 1.52	
episode: 1168	score: 3.0	memory length: 100000	epsilon: 0.7732860400049217
steps: 246	lr: 4e-05	evaluation reward: 1.53	
episode: 1169	score: 3.0	memory length: 100000	epsilon: 0.7727989600049323
steps: 246	lr: 4e-05	evaluation reward: 1.54	
episode: 1170	score: 4.0	memory length: 100000	epsilon: 0.772216840004945
steps: 294	lr: 4e-05	evaluation reward: 1.56	
episode: 1171	score: 5.0	memory length: 100000	epsilon: 0.7715357200049597
steps: 344	lr: 4e-05	evaluation reward: 1.59	
episode: 1172	score: 3.0	memory length: 100000	epsilon: 0.7710486400049703
steps: 246	lr: 4e-05	evaluation reward: 1.62	
episode: 1173	score: 1.0	memory length: 100000	epsilon: 0.7707120400049776
steps: 170	lr: 4e-05	evaluation reward: 1.6	
episode: 1174	score: 4.0	memory length: 100000	epsilon: 0.7702031800049887
steps: 257	lr: 4e-05	evaluation reward: 1.63	
episode: 1175	score: 2.0	memory length: 100000	epsilon: 0.7697695600049981
steps: 219	lr: 4e-05	evaluation reward: 1.63	
episode: 1176	score: 0.0	memory length: 100000	epsilon: 0.7695260200050034
steps: 123	lr: 4e-05	evaluation reward: 1.62	
episode: 1177	score: 1.0	memory length: 100000	epsilon: 0.7691874400050107
steps: 171	lr: 4e-05	evaluation reward: 1.63	
episode: 1178	score: 1.0	memory length: 100000	epsilon: 0.768852820005018
steps: 169	lr: 4e-05	evaluation reward: 1.64	
episode: 1179	score: 1.0	memory length: 100000	epsilon: 0.7685162200050253
steps: 170	lr: 4e-05	evaluation reward: 1.63	
episode: 1180	score: 0.0	memory length: 100000	epsilon: 0.7682726800050306
steps: 123	lr: 4e-05	evaluation reward: 1.6	
episode: 1181	score: 1.0	memory length: 100000	epsilon: 0.7679400400050378
steps: 168	lr: 4e-05	evaluation reward: 1.6	
episode: 1182	score: 3.0	memory length: 100000	epsilon: 0.7674094000050493
steps: 268	lr: 4e-05	evaluation reward: 1.63	
episode: 1183	score: 1.0	memory length: 100000	epsilon: 0.7670747800050566
steps: 169	lr: 4e-05	evaluation reward: 1.63	
episode: 1184	score: 2.0	memory length: 100000	epsilon: 0.7666827400050651
steps: 198	lr: 4e-05	evaluation reward: 1.62	
episode: 1185	score: 1.0	memory length: 100000	epsilon: 0.7663481200050724
steps: 169	lr: 4e-05	evaluation reward: 1.6	
episode: 1186	score: 2.0	memory length: 100000	epsilon: 0.7659560800050809
steps: 198	lr: 4e-05	evaluation reward: 1.61	
episode: 1187	score: 0.0	memory length: 100000	epsilon: 0.7657125400050862

steps: 123	lr: 4e-05	evaluation reward: 1.61	
episode: 1188	score: 0.0	memory length: 100000	epsilon: 0.7654690000050914
steps: 123	lr: 4e-05	evaluation reward: 1.59	
episode: 1189	score: 1.0	memory length: 100000	epsilon: 0.7651700200050979
steps: 151	lr: 4e-05	evaluation reward: 1.57	
episode: 1190	score: 4.0	memory length: 100000	epsilon: 0.7645799800051107
steps: 298	lr: 4e-05	evaluation reward: 1.57	
episode: 1191	score: 0.0	memory length: 100000	epsilon: 0.764336440005116
steps: 123	lr: 4e-05	evaluation reward: 1.56	
episode: 1192	score: 3.0	memory length: 100000	epsilon: 0.7638157000051273
steps: 263	lr: 4e-05	evaluation reward: 1.59	
episode: 1193	score: 0.0	memory length: 100000	epsilon: 0.7635721600051326
steps: 123	lr: 4e-05	evaluation reward: 1.53	
episode: 1194	score: 4.0	memory length: 100000	epsilon: 0.7629940000051452
steps: 292	lr: 4e-05	evaluation reward: 1.57	
episode: 1195	score: 5.0	memory length: 100000	epsilon: 0.7624099000051578
steps: 295	lr: 4e-05	evaluation reward: 1.6	
episode: 1196	score: 1.0	memory length: 100000	epsilon: 0.7621129000051643
steps: 150	lr: 4e-05	evaluation reward: 1.61	
episode: 1197	score: 2.0	memory length: 100000	epsilon: 0.7616812600051737
steps: 218	lr: 4e-05	evaluation reward: 1.62	
episode: 1198	score: 4.0	memory length: 100000	epsilon: 0.7610951800051864
steps: 296	lr: 4e-05	evaluation reward: 1.66	
episode: 1199	score: 4.0	memory length: 100000	epsilon: 0.7605764200051977
steps: 262	lr: 4e-05	evaluation reward: 1.69	
episode: 1200	score: 0.0	memory length: 100000	epsilon: 0.7603328800052029
steps: 123	lr: 4e-05	evaluation reward: 1.66	
episode: 1201	score: 4.0	memory length: 100000	epsilon: 0.7598497600052134
steps: 244	lr: 4e-05	evaluation reward: 1.69	
episode: 1202	score: 1.0	memory length: 100000	epsilon: 0.7595507800052199
steps: 151	lr: 4e-05	evaluation reward: 1.7	
episode: 1203	score: 1.0	memory length: 100000	epsilon: 0.7592122000052273
steps: 171	lr: 4e-05	evaluation reward: 1.7	
episode: 1204	score: 2.0	memory length: 100000	epsilon: 0.7588221400052357
steps: 197	lr: 4e-05	evaluation reward: 1.72	
episode: 1205	score: 1.0	memory length: 100000	epsilon: 0.758487520005243
steps: 169	lr: 4e-05	evaluation reward: 1.71	
episode: 1206	score: 2.0	memory length: 100000	epsilon: 0.7580954800052515
steps: 198	lr: 4e-05	evaluation reward: 1.72	
episode: 1207	score: 4.0	memory length: 100000	epsilon: 0.7575450400052635
steps: 278	lr: 4e-05	evaluation reward: 1.74	
episode: 1208	score: 2.0	memory length: 100000	epsilon: 0.7571153800052728
steps: 217	lr: 4e-05	evaluation reward: 1.74	
episode: 1209	score: 2.0	memory length: 100000	epsilon: 0.7567233400052813
steps: 198	lr: 4e-05	evaluation reward: 1.74	
episode: 1210	score: 1.0	memory length: 100000	epsilon: 0.7564243600052878
steps: 151	lr: 4e-05	evaluation reward: 1.73	
episode: 1211	score: 2.0	memory length: 100000	epsilon: 0.7560343000052963

steps: 197	lr: 4e-05	evaluation reward: 1.74	
episode: 1212	score: 3.0	memory length: 100000	epsilon: 0.7555076200053077
steps: 266	lr: 4e-05	evaluation reward: 1.74	
episode: 1213	score: 6.0	memory length: 100000	epsilon: 0.7547631400053239
steps: 376	lr: 4e-05	evaluation reward: 1.79	
episode: 1214	score: 1.0	memory length: 100000	epsilon: 0.7544641600053303
steps: 151	lr: 4e-05	evaluation reward: 1.77	
episode: 1215	score: 1.0	memory length: 100000	epsilon: 0.7541275600053376
steps: 170	lr: 4e-05	evaluation reward: 1.75	
episode: 1216	score: 0.0	memory length: 100000	epsilon: 0.7538840200053429
steps: 123	lr: 4e-05	evaluation reward: 1.73	
episode: 1217	score: 1.0	memory length: 100000	epsilon: 0.7535494000053502
steps: 169	lr: 4e-05	evaluation reward: 1.73	
episode: 1218	score: 0.0	memory length: 100000	epsilon: 0.7533058600053555
steps: 123	lr: 4e-05	evaluation reward: 1.73	
episode: 1219	score: 3.0	memory length: 100000	epsilon: 0.752822740005366
steps: 244	lr: 4e-05	evaluation reward: 1.74	
episode: 1220	score: 1.0	memory length: 100000	epsilon: 0.7525237600053725
steps: 151	lr: 4e-05	evaluation reward: 1.73	
episode: 1221	score: 2.0	memory length: 100000	epsilon: 0.7521653800053802
steps: 181	lr: 4e-05	evaluation reward: 1.73	
episode: 1222	score: 0.0	memory length: 100000	epsilon: 0.7519238200053855
steps: 122	lr: 4e-05	evaluation reward: 1.73	
episode: 1223	score: 0.0	memory length: 100000	epsilon: 0.7516802800053908
steps: 123	lr: 4e-05	evaluation reward: 1.7	
episode: 1224	score: 4.0	memory length: 100000	epsilon: 0.7511753800054017
steps: 255	lr: 4e-05	evaluation reward: 1.74	
episode: 1225	score: 3.0	memory length: 100000	epsilon: 0.7507279000054115
steps: 226	lr: 4e-05	evaluation reward: 1.74	
episode: 1226	score: 2.0	memory length: 100000	epsilon: 0.7503002200054207
steps: 216	lr: 4e-05	evaluation reward: 1.73	
episode: 1227	score: 0.0	memory length: 100000	epsilon: 0.750056680005426
steps: 123	lr: 4e-05	evaluation reward: 1.71	
episode: 1228	score: 0.0	memory length: 100000	epsilon: 0.7498151200054313
steps: 122	lr: 4e-05	evaluation reward: 1.7	
episode: 1229	score: 5.0	memory length: 100000	epsilon: 0.7491023200054467
steps: 360	lr: 4e-05	evaluation reward: 1.75	
episode: 1230	score: 2.0	memory length: 100000	epsilon: 0.748674640005456
steps: 216	lr: 4e-05	evaluation reward: 1.75	
episode: 1231	score: 2.0	memory length: 100000	epsilon: 0.7482845800054645
steps: 197	lr: 4e-05	evaluation reward: 1.77	
episode: 1232	score: 1.0	memory length: 100000	epsilon: 0.7479499600054718
steps: 169	lr: 4e-05	evaluation reward: 1.73	
episode: 1233	score: 2.0	memory length: 100000	epsilon: 0.7475579200054803
steps: 198	lr: 4e-05	evaluation reward: 1.75	
episode: 1234	score: 2.0	memory length: 100000	epsilon: 0.7471658800054888
steps: 198	lr: 4e-05	evaluation reward: 1.77	
episode: 1235	score: 3.0	memory length: 100000	epsilon: 0.7466768200054994

steps: 247	lr: 4e-05	evaluation reward: 1.8	
episode: 1236	score: 1.0	memory length: 100000	epsilon: 0.7463422000055067
steps: 169	lr: 4e-05	evaluation reward: 1.78	
episode: 1237	score: 3.0	memory length: 100000	epsilon: 0.7458947200055164
steps: 226	lr: 4e-05	evaluation reward: 1.79	
episode: 1238	score: 1.0	memory length: 100000	epsilon: 0.7455581200055237
steps: 170	lr: 4e-05	evaluation reward: 1.8	
episode: 1239	score: 1.0	memory length: 100000	epsilon: 0.745221520005531
steps: 170	lr: 4e-05	evaluation reward: 1.81	
episode: 1240	score: 0.0	memory length: 100000	epsilon: 0.7449799600055362
steps: 122	lr: 4e-05	evaluation reward: 1.78	
episode: 1241	score: 1.0	memory length: 100000	epsilon: 0.7446809800055427
steps: 151	lr: 4e-05	evaluation reward: 1.77	
episode: 1242	score: 2.0	memory length: 100000	epsilon: 0.7443226000055505
steps: 181	lr: 4e-05	evaluation reward: 1.78	
episode: 1243	score: 2.0	memory length: 100000	epsilon: 0.7438909600055599
steps: 218	lr: 4e-05	evaluation reward: 1.78	
episode: 1244	score: 2.0	memory length: 100000	epsilon: 0.7434533800055694
steps: 221	lr: 4e-05	evaluation reward: 1.79	
episode: 1245	score: 1.0	memory length: 100000	epsilon: 0.7431187600055766
steps: 169	lr: 4e-05	evaluation reward: 1.77	
episode: 1246	score: 2.0	memory length: 100000	epsilon: 0.7426930600055859
steps: 215	lr: 4e-05	evaluation reward: 1.79	
episode: 1247	score: 4.0	memory length: 100000	epsilon: 0.7420673800055995
steps: 316	lr: 4e-05	evaluation reward: 1.81	
episode: 1248	score: 0.0	memory length: 100000	epsilon: 0.7418238400056048
steps: 123	lr: 4e-05	evaluation reward: 1.78	
episode: 1249	score: 2.0	memory length: 100000	epsilon: 0.7413922000056141
steps: 218	lr: 4e-05	evaluation reward: 1.8	
episode: 1250	score: 0.0	memory length: 100000	epsilon: 0.7411486600056194
steps: 123	lr: 4e-05	evaluation reward: 1.8	
episode: 1251	score: 3.0	memory length: 100000	epsilon: 0.74065960000563
steps: 247	lr: 4e-05	evaluation reward: 1.83	
episode: 1252	score: 2.0	memory length: 100000	epsilon: 0.7402636000056386
steps: 200	lr: 4e-05	evaluation reward: 1.82	
episode: 1253	score: 2.0	memory length: 100000	epsilon: 0.7398715600056471
steps: 198	lr: 4e-05	evaluation reward: 1.81	
episode: 1254	score: 2.0	memory length: 100000	epsilon: 0.7394795200056556
steps: 198	lr: 4e-05	evaluation reward: 1.83	
episode: 1255	score: 3.0	memory length: 100000	epsilon: 0.7390320400056654
steps: 226	lr: 4e-05	evaluation reward: 1.86	
episode: 1256	score: 2.0	memory length: 100000	epsilon: 0.7386004000056747
steps: 218	lr: 4e-05	evaluation reward: 1.85	
episode: 1257	score: 3.0	memory length: 100000	epsilon: 0.7381529200056844
steps: 226	lr: 4e-05	evaluation reward: 1.87	
episode: 1258	score: 2.0	memory length: 100000	epsilon: 0.7377193000056939
steps: 219	lr: 4e-05	evaluation reward: 1.89	
episode: 1259	score: 1.0	memory length: 100000	epsilon: 0.7374203200057003

steps: 151	lr: 4e-05	evaluation reward: 1.88	
episode: 1260	score: 0.0	memory length: 100000	epsilon: 0.7371787600057056
steps: 122	lr: 4e-05	evaluation reward: 1.88	
episode: 1261	score: 3.0	memory length: 100000	epsilon: 0.7367312800057153
steps: 226	lr: 4e-05	evaluation reward: 1.9	
episode: 1262	score: 2.0	memory length: 100000	epsilon: 0.7362937000057248
steps: 221	lr: 4e-05	evaluation reward: 1.9	
episode: 1263	score: 2.0	memory length: 100000	epsilon: 0.7359016600057333
steps: 198	lr: 4e-05	evaluation reward: 1.9	
episode: 1264	score: 1.0	memory length: 100000	epsilon: 0.7355650600057406
steps: 170	lr: 4e-05	evaluation reward: 1.86	
episode: 1265	score: 1.0	memory length: 100000	epsilon: 0.7352680600057471
steps: 150	lr: 4e-05	evaluation reward: 1.86	
episode: 1266	score: 1.0	memory length: 100000	epsilon: 0.7349314600057544
steps: 170	lr: 4e-05	evaluation reward: 1.84	
episode: 1267	score: 2.0	memory length: 100000	epsilon: 0.7345057600057636
steps: 215	lr: 4e-05	evaluation reward: 1.86	
episode: 1268	score: 0.0	memory length: 100000	epsilon: 0.7342622200057689
steps: 123	lr: 4e-05	evaluation reward: 1.83	
episode: 1269	score: 2.0	memory length: 100000	epsilon: 0.7339078000057766
steps: 179	lr: 4e-05	evaluation reward: 1.82	
episode: 1270	score: 4.0	memory length: 100000	epsilon: 0.7333256800057892
steps: 294	lr: 4e-05	evaluation reward: 1.82	
episode: 1271	score: 0.0	memory length: 100000	epsilon: 0.7330821400057945
steps: 123	lr: 4e-05	evaluation reward: 1.77	
episode: 1272	score: 1.0	memory length: 100000	epsilon: 0.7327415800058019
steps: 172	lr: 4e-05	evaluation reward: 1.75	
episode: 1273	score: 2.0	memory length: 100000	epsilon: 0.7323495400058104
steps: 198	lr: 4e-05	evaluation reward: 1.76	
episode: 1274	score: 3.0	memory length: 100000	epsilon: 0.7318981000058202
steps: 228	lr: 4e-05	evaluation reward: 1.75	
episode: 1275	score: 0.0	memory length: 100000	epsilon: 0.7316565400058255
steps: 122	lr: 4e-05	evaluation reward: 1.73	
episode: 1276	score: 4.0	memory length: 100000	epsilon: 0.7310744200058381
steps: 294	lr: 4e-05	evaluation reward: 1.77	
episode: 1277	score: 1.0	memory length: 100000	epsilon: 0.7307774200058446
steps: 150	lr: 4e-05	evaluation reward: 1.77	
episode: 1278	score: 1.0	memory length: 100000	epsilon: 0.7304447800058518
steps: 168	lr: 4e-05	evaluation reward: 1.77	
episode: 1279	score: 0.0	memory length: 100000	epsilon: 0.730203220005857
steps: 122	lr: 4e-05	evaluation reward: 1.76	
episode: 1280	score: 4.0	memory length: 100000	epsilon: 0.7296191200058697
steps: 295	lr: 4e-05	evaluation reward: 1.8	
episode: 1281	score: 0.0	memory length: 100000	epsilon: 0.729377560005875
steps: 122	lr: 4e-05	evaluation reward: 1.79	
episode: 1282	score: 2.0	memory length: 100000	epsilon: 0.7289835400058835
steps: 199	lr: 4e-05	evaluation reward: 1.78	
episode: 1283	score: 2.0	memory length: 100000	epsilon: 0.7285499200058929

steps: 219	lr: 4e-05	evaluation reward: 1.79	
episode: 1284	score: 1.0	memory length: 100000	epsilon: 0.7282113400059003
steps: 171	lr: 4e-05	evaluation reward: 1.78	
episode: 1285	score: 1.0	memory length: 100000	epsilon: 0.7279123600059068
steps: 151	lr: 4e-05	evaluation reward: 1.78	
episode: 1286	score: 2.0	memory length: 100000	epsilon: 0.7275203200059153
steps: 198	lr: 4e-05	evaluation reward: 1.78	
episode: 1287	score: 1.0	memory length: 100000	epsilon: 0.7272213400059218
steps: 151	lr: 4e-05	evaluation reward: 1.79	
episode: 1288	score: 2.0	memory length: 100000	epsilon: 0.7268293000059303
steps: 198	lr: 4e-05	evaluation reward: 1.81	
episode: 1289	score: 3.0	memory length: 100000	epsilon: 0.7263382600059409
steps: 248	lr: 4e-05	evaluation reward: 1.83	
episode: 1290	score: 1.0	memory length: 100000	epsilon: 0.7260412600059474
steps: 150	lr: 4e-05	evaluation reward: 1.8	
episode: 1291	score: 1.0	memory length: 100000	epsilon: 0.7257046600059547
steps: 170	lr: 4e-05	evaluation reward: 1.81	
episode: 1292	score: 1.0	memory length: 100000	epsilon: 0.7254056800059612
steps: 151	lr: 4e-05	evaluation reward: 1.79	
episode: 1293	score: 3.0	memory length: 100000	epsilon: 0.724954240005971
steps: 228	lr: 4e-05	evaluation reward: 1.82	
episode: 1294	score: 1.0	memory length: 100000	epsilon: 0.7246196200059782
steps: 169	lr: 4e-05	evaluation reward: 1.79	
episode: 1295	score: 4.0	memory length: 100000	epsilon: 0.72407710000599
steps: 274	lr: 4e-05	evaluation reward: 1.78	
episode: 1296	score: 2.0	memory length: 100000	epsilon: 0.7236850600059985
steps: 198	lr: 4e-05	evaluation reward: 1.79	
episode: 1297	score: 5.0	memory length: 100000	epsilon: 0.7229999800060134
steps: 346	lr: 4e-05	evaluation reward: 1.82	
episode: 1298	score: 1.0	memory length: 100000	epsilon: 0.7226653600060207
steps: 169	lr: 4e-05	evaluation reward: 1.79	
episode: 1299	score: 2.0	memory length: 100000	epsilon: 0.72223372000603
steps: 218	lr: 4e-05	evaluation reward: 1.77	
episode: 1300	score: 0.0	memory length: 100000	epsilon: 0.7219901800060353
steps: 123	lr: 4e-05	evaluation reward: 1.77	
episode: 1301	score: 2.0	memory length: 100000	epsilon: 0.7215941800060439
steps: 200	lr: 4e-05	evaluation reward: 1.75	
episode: 1302	score: 2.0	memory length: 100000	epsilon: 0.7211645200060532
steps: 217	lr: 4e-05	evaluation reward: 1.76	
episode: 1303	score: 3.0	memory length: 100000	epsilon: 0.7206774400060638
steps: 246	lr: 4e-05	evaluation reward: 1.78	
episode: 1304	score: 1.0	memory length: 100000	epsilon: 0.720344800006071
steps: 168	lr: 4e-05	evaluation reward: 1.77	
episode: 1305	score: 1.0	memory length: 100000	epsilon: 0.7200458200060775
steps: 151	lr: 4e-05	evaluation reward: 1.77	
episode: 1306	score: 1.0	memory length: 100000	epsilon: 0.7197092200060848
steps: 170	lr: 4e-05	evaluation reward: 1.76	
episode: 1307	score: 2.0	memory length: 100000	epsilon: 0.7193488600060927

steps: 182	lr: 4e-05	evaluation reward: 1.74	
episode: 1308	score: 0.0	memory length: 100000	epsilon: 0.7191073000060979
steps: 122	lr: 4e-05	evaluation reward: 1.72	
episode: 1309	score: 2.0	memory length: 100000	epsilon: 0.7186756600061073
steps: 218	lr: 4e-05	evaluation reward: 1.72	
episode: 1310	score: 4.0	memory length: 100000	epsilon: 0.7180539400061208
steps: 314	lr: 4e-05	evaluation reward: 1.75	
episode: 1311	score: 2.0	memory length: 100000	epsilon: 0.7176619000061293
steps: 198	lr: 4e-05	evaluation reward: 1.75	
episode: 1312	score: 2.0	memory length: 100000	epsilon: 0.7172678800061378
steps: 199	lr: 4e-05	evaluation reward: 1.74	
episode: 1313	score: 1.0	memory length: 100000	epsilon: 0.7169332600061451
steps: 169	lr: 4e-05	evaluation reward: 1.69	
episode: 1314	score: 2.0	memory length: 100000	epsilon: 0.7164976600061546
steps: 220	lr: 4e-05	evaluation reward: 1.7	
episode: 1315	score: 3.0	memory length: 100000	epsilon: 0.715972960006166
steps: 265	lr: 4e-05	evaluation reward: 1.72	
episode: 1316	score: 5.0	memory length: 100000	epsilon: 0.7153631200061792
steps: 308	lr: 4e-05	evaluation reward: 1.77	
episode: 1317	score: 3.0	memory length: 100000	epsilon: 0.7148443600061904
steps: 262	lr: 4e-05	evaluation reward: 1.79	
episode: 1318	score: 2.0	memory length: 100000	epsilon: 0.7144879600061982
steps: 180	lr: 4e-05	evaluation reward: 1.81	
episode: 1319	score: 2.0	memory length: 100000	epsilon: 0.7140979000062067
steps: 197	lr: 4e-05	evaluation reward: 1.8	
episode: 1320	score: 4.0	memory length: 100000	epsilon: 0.7134702400062203
steps: 317	lr: 4e-05	evaluation reward: 1.83	
episode: 1321	score: 3.0	memory length: 100000	epsilon: 0.71302276000623
steps: 226	lr: 4e-05	evaluation reward: 1.84	
episode: 1322	score: 3.0	memory length: 100000	epsilon: 0.7125337000062406
steps: 247	lr: 4e-05	evaluation reward: 1.87	
episode: 1323	score: 3.0	memory length: 100000	epsilon: 0.7120882000062503
steps: 225	lr: 4e-05	evaluation reward: 1.9	
episode: 1324	score: 1.0	memory length: 100000	epsilon: 0.7117555600062575
steps: 168	lr: 4e-05	evaluation reward: 1.87	
episode: 1325	score: 4.0	memory length: 100000	epsilon: 0.7112110600062693
steps: 275	lr: 4e-05	evaluation reward: 1.88	
episode: 1326	score: 1.0	memory length: 100000	epsilon: 0.7108764400062766
steps: 169	lr: 4e-05	evaluation reward: 1.87	
episode: 1327	score: 1.0	memory length: 100000	epsilon: 0.7105735000062832
steps: 153	lr: 4e-05	evaluation reward: 1.88	
episode: 1328	score: 1.0	memory length: 100000	epsilon: 0.7102725400062897
steps: 152	lr: 4e-05	evaluation reward: 1.89	
episode: 1329	score: 1.0	memory length: 100000	epsilon: 0.709937920006297
steps: 169	lr: 4e-05	evaluation reward: 1.85	
episode: 1330	score: 1.0	memory length: 100000	epsilon: 0.7096033000063042
steps: 169	lr: 4e-05	evaluation reward: 1.84	
episode: 1331	score: 4.0	memory length: 100000	epsilon: 0.7090885000063154

steps: 260	lr: 4e-05	evaluation reward: 1.86	
episode: 1332	score: 3.0	memory length: 100000	epsilon: 0.7086370600063252
steps: 228	lr: 4e-05	evaluation reward: 1.88	
episode: 1333	score: 2.0	memory length: 100000	epsilon: 0.7082410600063338
steps: 200	lr: 4e-05	evaluation reward: 1.88	
episode: 1334	score: 2.0	memory length: 100000	epsilon: 0.7078490200063423
steps: 198	lr: 4e-05	evaluation reward: 1.88	
episode: 1335	score: 3.0	memory length: 100000	epsilon: 0.7073599600063529
steps: 247	lr: 4e-05	evaluation reward: 1.88	
episode: 1336	score: 0.0	memory length: 100000	epsilon: 0.7071184000063582
steps: 122	lr: 4e-05	evaluation reward: 1.87	
episode: 1337	score: 3.0	memory length: 100000	epsilon: 0.7065897400063696
steps: 267	lr: 4e-05	evaluation reward: 1.87	
episode: 1338	score: 3.0	memory length: 100000	epsilon: 0.706067020006381
steps: 264	lr: 4e-05	evaluation reward: 1.89	
episode: 1339	score: 0.0	memory length: 100000	epsilon: 0.7058254600063862
steps: 122	lr: 4e-05	evaluation reward: 1.88	
episode: 1340	score: 2.0	memory length: 100000	epsilon: 0.7054274800063949
steps: 201	lr: 4e-05	evaluation reward: 1.9	
episode: 1341	score: 1.0	memory length: 100000	epsilon: 0.7050928600064021
steps: 169	lr: 4e-05	evaluation reward: 1.9	
episode: 1342	score: 3.0	memory length: 100000	epsilon: 0.7046018200064128
steps: 248	lr: 4e-05	evaluation reward: 1.91	
episode: 1343	score: 8.0	memory length: 100000	epsilon: 0.7039088200064278
steps: 350	lr: 4e-05	evaluation reward: 1.97	
episode: 1344	score: 3.0	memory length: 100000	epsilon: 0.7034197600064385
steps: 247	lr: 4e-05	evaluation reward: 1.98	
episode: 1345	score: 0.0	memory length: 100000	epsilon: 0.7031782000064437
steps: 122	lr: 4e-05	evaluation reward: 1.97	
episode: 1346	score: 3.0	memory length: 100000	epsilon: 0.7027307200064534
steps: 226	lr: 4e-05	evaluation reward: 1.98	
episode: 1347	score: 3.0	memory length: 100000	epsilon: 0.702245620006464
steps: 245	lr: 4e-05	evaluation reward: 1.97	
episode: 1348	score: 0.0	memory length: 100000	epsilon: 0.7020040600064692
steps: 122	lr: 4e-05	evaluation reward: 1.97	
episode: 1349	score: 3.0	memory length: 100000	epsilon: 0.7014773800064806
steps: 266	lr: 4e-05	evaluation reward: 1.98	
episode: 1350	score: 2.0	memory length: 100000	epsilon: 0.7011150400064885
steps: 183	lr: 4e-05	evaluation reward: 2.0	
episode: 1351	score: 3.0	memory length: 100000	epsilon: 0.7006675600064982
steps: 226	lr: 4e-05	evaluation reward: 2.0	
episode: 1352	score: 0.0	memory length: 100000	epsilon: 0.7004240200065035
steps: 123	lr: 4e-05	evaluation reward: 1.98	
episode: 1353	score: 4.0	memory length: 100000	epsilon: 0.6998359600065163
steps: 297	lr: 4e-05	evaluation reward: 2.0	
episode: 1354	score: 2.0	memory length: 100000	epsilon: 0.699479560006524
steps: 180	lr: 4e-05	evaluation reward: 2.0	
episode: 1355	score: 2.0	memory length: 100000	epsilon: 0.6990459400065334

steps: 219	lr: 4e-05	evaluation reward: 1.99	
episode: 1356	score: 5.0	memory length: 100000	epsilon: 0.6983984800065475
steps: 327	lr: 4e-05	evaluation reward: 2.02	
episode: 1357	score: 2.0	memory length: 100000	epsilon: 0.698006440006556
steps: 198	lr: 4e-05	evaluation reward: 2.01	
episode: 1358	score: 2.0	memory length: 100000	epsilon: 0.6976144000065645
steps: 198	lr: 4e-05	evaluation reward: 2.01	
episode: 1359	score: 2.0	memory length: 100000	epsilon: 0.6972580000065722
steps: 180	lr: 4e-05	evaluation reward: 2.02	
episode: 1360	score: 3.0	memory length: 100000	epsilon: 0.6967689400065828
steps: 247	lr: 4e-05	evaluation reward: 2.05	
episode: 1361	score: 0.0	memory length: 100000	epsilon: 0.6965254000065881
steps: 123	lr: 4e-05	evaluation reward: 2.02	
episode: 1362	score: 0.0	memory length: 100000	epsilon: 0.6962838400065934
steps: 122	lr: 4e-05	evaluation reward: 2.0	
episode: 1363	score: 2.0	memory length: 100000	epsilon: 0.6958522000066028
steps: 218	lr: 4e-05	evaluation reward: 2.0	
episode: 1364	score: 4.0	memory length: 100000	epsilon: 0.6952661200066155
steps: 296	lr: 4e-05	evaluation reward: 2.03	
episode: 1365	score: 1.0	memory length: 100000	epsilon: 0.6949315000066227
steps: 169	lr: 4e-05	evaluation reward: 2.03	
episode: 1366	score: 6.0	memory length: 100000	epsilon: 0.694226620006638
steps: 356	lr: 4e-05	evaluation reward: 2.08	
episode: 1367	score: 2.0	memory length: 100000	epsilon: 0.6938345800066466
steps: 198	lr: 4e-05	evaluation reward: 2.08	
episode: 1368	score: 5.0	memory length: 100000	epsilon: 0.6932663200066589
steps: 287	lr: 4e-05	evaluation reward: 2.13	
episode: 1369	score: 2.0	memory length: 100000	epsilon: 0.6928703200066675
steps: 200	lr: 4e-05	evaluation reward: 2.13	
episode: 1370	score: 5.0	memory length: 100000	epsilon: 0.6922842400066802
steps: 296	lr: 4e-05	evaluation reward: 2.14	
episode: 1371	score: 1.0	memory length: 100000	epsilon: 0.6919516000066874
steps: 168	lr: 4e-05	evaluation reward: 2.15	
episode: 1372	score: 4.0	memory length: 100000	epsilon: 0.6913912600066996
steps: 283	lr: 4e-05	evaluation reward: 2.18	
episode: 1373	score: 3.0	memory length: 100000	epsilon: 0.6909002200067103
steps: 248	lr: 4e-05	evaluation reward: 2.19	
episode: 1374	score: 2.0	memory length: 100000	epsilon: 0.6905081800067188
steps: 198	lr: 4e-05	evaluation reward: 2.18	
episode: 1375	score: 4.0	memory length: 100000	epsilon: 0.6899280400067314
steps: 293	lr: 4e-05	evaluation reward: 2.22	
episode: 1376	score: 1.0	memory length: 100000	epsilon: 0.6896310400067378
steps: 150	lr: 4e-05	evaluation reward: 2.19	
episode: 1377	score: 1.0	memory length: 100000	epsilon: 0.6892924600067452
steps: 171	lr: 4e-05	evaluation reward: 2.19	
episode: 1378	score: 3.0	memory length: 100000	epsilon: 0.6888449800067549
steps: 226	lr: 4e-05	evaluation reward: 2.21	
episode: 1379	score: 1.0	memory length: 100000	epsilon: 0.6885103600067621

steps: 169	lr: 4e-05	evaluation reward: 2.22	
episode: 1380	score: 4.0	memory length: 100000	epsilon: 0.687963880006774
steps: 276	lr: 4e-05	evaluation reward: 2.22	
episode: 1381	score: 0.0	memory length: 100000	epsilon: 0.6877203400067793
steps: 123	lr: 4e-05	evaluation reward: 2.22	
episode: 1382	score: 1.0	memory length: 100000	epsilon: 0.6874213600067858
steps: 151	lr: 4e-05	evaluation reward: 2.21	
episode: 1383	score: 4.0	memory length: 100000	epsilon: 0.6869085400067969
steps: 259	lr: 4e-05	evaluation reward: 2.23	
episode: 1384	score: 1.0	memory length: 100000	epsilon: 0.6865719400068042
steps: 170	lr: 4e-05	evaluation reward: 2.23	
episode: 1385	score: 6.0	memory length: 100000	epsilon: 0.6858650800068196
steps: 357	lr: 4e-05	evaluation reward: 2.28	
episode: 1386	score: 6.0	memory length: 100000	epsilon: 0.6851027800068361
steps: 385	lr: 4e-05	evaluation reward: 2.32	
episode: 1387	score: 5.0	memory length: 100000	epsilon: 0.68446324000685
steps: 323	lr: 4e-05	evaluation reward: 2.36	
episode: 1388	score: 5.0	memory length: 100000	epsilon: 0.683817760006864
steps: 326	lr: 4e-05	evaluation reward: 2.39	
episode: 1389	score: 2.0	memory length: 100000	epsilon: 0.6834257200068725
steps: 198	lr: 4e-05	evaluation reward: 2.38	
episode: 1390	score: 2.0	memory length: 100000	epsilon: 0.6830317000068811
steps: 199	lr: 4e-05	evaluation reward: 2.39	
episode: 1391	score: 3.0	memory length: 100000	epsilon: 0.6825485800068916
steps: 244	lr: 4e-05	evaluation reward: 2.41	
episode: 1392	score: 4.0	memory length: 100000	epsilon: 0.6819882400069037
steps: 283	lr: 4e-05	evaluation reward: 2.44	
episode: 1393	score: 2.0	memory length: 100000	epsilon: 0.681562540006913
steps: 215	lr: 4e-05	evaluation reward: 2.43	
episode: 1394	score: 2.0	memory length: 100000	epsilon: 0.6811705000069215
steps: 198	lr: 4e-05	evaluation reward: 2.44	
episode: 1395	score: 0.0	memory length: 100000	epsilon: 0.6809289400069267
steps: 122	lr: 4e-05	evaluation reward: 2.4	
episode: 1396	score: 6.0	memory length: 100000	epsilon: 0.6802300000069419
steps: 353	lr: 4e-05	evaluation reward: 2.44	
episode: 1397	score: 3.0	memory length: 100000	epsilon: 0.6797389600069526
steps: 248	lr: 4e-05	evaluation reward: 2.42	
episode: 1398	score: 1.0	memory length: 100000	epsilon: 0.6794043400069598
steps: 169	lr: 4e-05	evaluation reward: 2.42	
episode: 1399	score: 3.0	memory length: 100000	epsilon: 0.6789152800069704
steps: 247	lr: 4e-05	evaluation reward: 2.43	
episode: 1400	score: 2.0	memory length: 100000	epsilon: 0.678523240006979
steps: 198	lr: 4e-05	evaluation reward: 2.45	
episode: 1401	score: 0.0	memory length: 100000	epsilon: 0.6782797000069842
steps: 123	lr: 4e-05	evaluation reward: 2.43	
episode: 1402	score: 0.0	memory length: 100000	epsilon: 0.6780361600069895
steps: 123	lr: 4e-05	evaluation reward: 2.41	
episode: 1403	score: 2.0	memory length: 100000	epsilon: 0.6776025400069989

steps: 219	lr: 4e-05	evaluation reward: 2.4	
episode: 1404	score: 0.0	memory length: 100000	epsilon: 0.6773590000070042
steps: 123	lr: 4e-05	evaluation reward: 2.39	
episode: 1405	score: 0.0	memory length: 100000	epsilon: 0.6771154600070095
steps: 123	lr: 4e-05	evaluation reward: 2.38	
episode: 1406	score: 3.0	memory length: 100000	epsilon: 0.6766679800070192
steps: 226	lr: 4e-05	evaluation reward: 2.4	
episode: 1407	score: 4.0	memory length: 100000	epsilon: 0.676123480007031
steps: 275	lr: 4e-05	evaluation reward: 2.42	
episode: 1408	score: 2.0	memory length: 100000	epsilon: 0.6757314400070396
steps: 198	lr: 4e-05	evaluation reward: 2.44	
episode: 1409	score: 0.0	memory length: 100000	epsilon: 0.6754879000070448
steps: 123	lr: 4e-05	evaluation reward: 2.42	
episode: 1410	score: 1.0	memory length: 100000	epsilon: 0.6751889200070513
steps: 151	lr: 4e-05	evaluation reward: 2.39	
episode: 1411	score: 1.0	memory length: 100000	epsilon: 0.6748899400070578
steps: 151	lr: 4e-05	evaluation reward: 2.38	
episode: 1412	score: 4.0	memory length: 100000	epsilon: 0.6743098000070704
steps: 293	lr: 4e-05	evaluation reward: 2.4	
episode: 1413	score: 4.0	memory length: 100000	epsilon: 0.673684120007084
steps: 316	lr: 4e-05	evaluation reward: 2.43	
episode: 1414	score: 4.0	memory length: 100000	epsilon: 0.6731396200070958
steps: 275	lr: 4e-05	evaluation reward: 2.45	
episode: 1415	score: 4.0	memory length: 100000	epsilon: 0.6725931400071077
steps: 276	lr: 4e-05	evaluation reward: 2.46	
episode: 1416	score: 3.0	memory length: 100000	epsilon: 0.6721120000071181
steps: 243	lr: 4e-05	evaluation reward: 2.44	
episode: 1417	score: 6.0	memory length: 100000	epsilon: 0.6714031600071335
steps: 358	lr: 4e-05	evaluation reward: 2.47	
episode: 1418	score: 2.0	memory length: 100000	epsilon: 0.671011120007142
steps: 198	lr: 4e-05	evaluation reward: 2.47	
episode: 1419	score: 3.0	memory length: 100000	epsilon: 0.6705636400071517
steps: 226	lr: 4e-05	evaluation reward: 2.48	
episode: 1420	score: 0.0	memory length: 100000	epsilon: 0.670320100007157
steps: 123	lr: 4e-05	evaluation reward: 2.44	
episode: 1421	score: 4.0	memory length: 100000	epsilon: 0.6697657000071691
steps: 280	lr: 4e-05	evaluation reward: 2.45	
episode: 1422	score: 1.0	memory length: 100000	epsilon: 0.6694271200071764
steps: 171	lr: 4e-05	evaluation reward: 2.43	
episode: 1423	score: 4.0	memory length: 100000	epsilon: 0.668846980007189
steps: 293	lr: 4e-05	evaluation reward: 2.44	
episode: 1424	score: 0.0	memory length: 100000	epsilon: 0.6686034400071943
steps: 123	lr: 4e-05	evaluation reward: 2.43	
episode: 1425	score: 1.0	memory length: 100000	epsilon: 0.6682688200072016
steps: 169	lr: 4e-05	evaluation reward: 2.4	
episode: 1426	score: 3.0	memory length: 100000	epsilon: 0.6678154000072114
steps: 229	lr: 4e-05	evaluation reward: 2.42	
episode: 1427	score: 5.0	memory length: 100000	epsilon: 0.6672055600072246

steps: 308	lr: 4e-05	evaluation reward: 2.46	
episode: 1428	score: 2.0	memory length: 100000	epsilon: 0.6668135200072332
steps: 198	lr: 4e-05	evaluation reward: 2.47	
episode: 1429	score: 3.0	memory length: 100000	epsilon: 0.6663224800072438
steps: 248	lr: 4e-05	evaluation reward: 2.49	
episode: 1430	score: 1.0	memory length: 100000	epsilon: 0.6660235000072503
steps: 151	lr: 4e-05	evaluation reward: 2.49	
episode: 1431	score: 2.0	memory length: 100000	epsilon: 0.6655918600072597
steps: 218	lr: 4e-05	evaluation reward: 2.47	
episode: 1432	score: 3.0	memory length: 100000	epsilon: 0.6651047800072702
steps: 246	lr: 4e-05	evaluation reward: 2.47	
episode: 1433	score: 2.0	memory length: 100000	epsilon: 0.6647127400072788
steps: 198	lr: 4e-05	evaluation reward: 2.47	
episode: 1434	score: 4.0	memory length: 100000	epsilon: 0.6641741800072904
steps: 272	lr: 4e-05	evaluation reward: 2.49	
episode: 1435	score: 1.0	memory length: 100000	epsilon: 0.6638752000072969
steps: 151	lr: 4e-05	evaluation reward: 2.47	
episode: 1436	score: 4.0	memory length: 100000	epsilon: 0.6633307000073088
steps: 275	lr: 4e-05	evaluation reward: 2.51	
episode: 1437	score: 1.0	memory length: 100000	epsilon: 0.6630317200073153
steps: 151	lr: 4e-05	evaluation reward: 2.49	
episode: 1438	score: 3.0	memory length: 100000	epsilon: 0.6625862200073249
steps: 225	lr: 4e-05	evaluation reward: 2.49	
episode: 1439	score: 2.0	memory length: 100000	epsilon: 0.6621902200073335
steps: 200	lr: 4e-05	evaluation reward: 2.51	
episode: 1440	score: 5.0	memory length: 100000	epsilon: 0.6615645400073471
steps: 316	lr: 4e-05	evaluation reward: 2.54	
episode: 1441	score: 1.0	memory length: 100000	epsilon: 0.6612675400073535
steps: 150	lr: 4e-05	evaluation reward: 2.54	
episode: 1442	score: 2.0	memory length: 100000	epsilon: 0.660877480007362
steps: 197	lr: 4e-05	evaluation reward: 2.53	
episode: 1443	score: 3.0	memory length: 100000	epsilon: 0.6603844600073727
steps: 249	lr: 4e-05	evaluation reward: 2.48	
episode: 1444	score: 2.0	memory length: 100000	epsilon: 0.6599924200073812
steps: 198	lr: 4e-05	evaluation reward: 2.47	
episode: 1445	score: 1.0	memory length: 100000	epsilon: 0.6596934400073877
steps: 151	lr: 4e-05	evaluation reward: 2.48	
episode: 1446	score: 4.0	memory length: 100000	epsilon: 0.6591093400074004
steps: 295	lr: 4e-05	evaluation reward: 2.49	
episode: 1447	score: 2.0	memory length: 100000	epsilon: 0.658713340007409
steps: 200	lr: 4e-05	evaluation reward: 2.48	
episode: 1448	score: 5.0	memory length: 100000	epsilon: 0.6580322200074238
steps: 344	lr: 4e-05	evaluation reward: 2.53	
episode: 1449	score: 0.0	memory length: 100000	epsilon: 0.6577886800074291
steps: 123	lr: 4e-05	evaluation reward: 2.5	
episode: 1450	score: 3.0	memory length: 100000	epsilon: 0.6573016000074396
steps: 246	lr: 4e-05	evaluation reward: 2.51	
episode: 1451	score: 1.0	memory length: 100000	epsilon: 0.6570026200074461

steps: 151	lr: 4e-05	evaluation reward: 2.49	
episode: 1452	score: 1.0	memory length: 100000	epsilon: 0.6566640400074535
steps: 171	lr: 4e-05	evaluation reward: 2.5	
episode: 1453	score: 2.0	memory length: 100000	epsilon: 0.6563056600074613
steps: 181	lr: 4e-05	evaluation reward: 2.48	
episode: 1454	score: 0.0	memory length: 100000	epsilon: 0.6560621200074666
steps: 123	lr: 4e-05	evaluation reward: 2.46	
episode: 1455	score: 5.0	memory length: 100000	epsilon: 0.6553790200074814
steps: 345	lr: 4e-05	evaluation reward: 2.49	
episode: 1456	score: 3.0	memory length: 100000	epsilon: 0.654891940007492
steps: 246	lr: 4e-05	evaluation reward: 2.47	
episode: 1457	score: 1.0	memory length: 100000	epsilon: 0.6545929600074984
steps: 151	lr: 4e-05	evaluation reward: 2.46	
episode: 1458	score: 2.0	memory length: 100000	epsilon: 0.654196960007507
steps: 200	lr: 4e-05	evaluation reward: 2.46	
episode: 1459	score: 3.0	memory length: 100000	epsilon: 0.6537455200075168
steps: 228	lr: 4e-05	evaluation reward: 2.47	
episode: 1460	score: 2.0	memory length: 100000	epsilon: 0.6533554600075253
steps: 197	lr: 4e-05	evaluation reward: 2.46	
episode: 1461	score: 2.0	memory length: 100000	epsilon: 0.6529258000075346
steps: 217	lr: 4e-05	evaluation reward: 2.48	
episode: 1462	score: 2.0	memory length: 100000	epsilon: 0.652492180007544
steps: 219	lr: 4e-05	evaluation reward: 2.5	
episode: 1463	score: 0.0	memory length: 100000	epsilon: 0.6522506200075493
steps: 122	lr: 4e-05	evaluation reward: 2.48	
episode: 1464	score: 3.0	memory length: 100000	epsilon: 0.6517219600075608
steps: 267	lr: 4e-05	evaluation reward: 2.47	
episode: 1465	score: 2.0	memory length: 100000	epsilon: 0.6512923000075701
steps: 217	lr: 4e-05	evaluation reward: 2.48	
episode: 1466	score: 1.0	memory length: 100000	epsilon: 0.6509933200075766
steps: 151	lr: 4e-05	evaluation reward: 2.43	
episode: 1467	score: 1.0	memory length: 100000	epsilon: 0.6506606800075838
steps: 168	lr: 4e-05	evaluation reward: 2.42	
episode: 1468	score: 3.0	memory length: 100000	epsilon: 0.6501280600075954
steps: 269	lr: 4e-05	evaluation reward: 2.4	
episode: 1469	score: 5.0	memory length: 100000	epsilon: 0.6494330800076105
steps: 351	lr: 4e-05	evaluation reward: 2.43	
episode: 1470	score: 3.0	memory length: 100000	epsilon: 0.6489816400076203
steps: 228	lr: 4e-05	evaluation reward: 2.41	
episode: 1471	score: 3.0	memory length: 100000	epsilon: 0.6485361400076299
steps: 225	lr: 4e-05	evaluation reward: 2.43	
episode: 1472	score: 6.0	memory length: 100000	epsilon: 0.6478352200076452
steps: 354	lr: 4e-05	evaluation reward: 2.45	
episode: 1473	score: 1.0	memory length: 100000	epsilon: 0.6475362400076516
steps: 151	lr: 4e-05	evaluation reward: 2.43	
episode: 1474	score: 2.0	memory length: 100000	epsilon: 0.647104600007661
steps: 218	lr: 4e-05	evaluation reward: 2.43	
episode: 1475	score: 2.0	memory length: 100000	epsilon: 0.6467125600076695

steps: 198	lr: 4e-05	evaluation reward: 2.41	
episode: 1476	score: 2.0	memory length: 100000	epsilon: 0.6463185400076781
steps: 199	lr: 4e-05	evaluation reward: 2.42	
episode: 1477	score: 2.0	memory length: 100000	epsilon: 0.6459284800076865
steps: 197	lr: 4e-05	evaluation reward: 2.43	
episode: 1478	score: 4.0	memory length: 100000	epsilon: 0.6453404200076993
steps: 297	lr: 4e-05	evaluation reward: 2.44	
episode: 1479	score: 8.0	memory length: 100000	epsilon: 0.6444890200077178
steps: 430	lr: 4e-05	evaluation reward: 2.51	
episode: 1480	score: 3.0	memory length: 100000	epsilon: 0.6440039200077283
steps: 245	lr: 4e-05	evaluation reward: 2.5	
episode: 1481	score: 2.0	memory length: 100000	epsilon: 0.6436435600077361
steps: 182	lr: 4e-05	evaluation reward: 2.52	
episode: 1482	score: 2.0	memory length: 100000	epsilon: 0.6432515200077447
steps: 198	lr: 4e-05	evaluation reward: 2.53	
episode: 1483	score: 4.0	memory length: 100000	epsilon: 0.6427070200077565
steps: 275	lr: 4e-05	evaluation reward: 2.53	
episode: 1484	score: 2.0	memory length: 100000	epsilon: 0.642314980007765
steps: 198	lr: 4e-05	evaluation reward: 2.54	
episode: 1485	score: 2.0	memory length: 100000	epsilon: 0.6419229400077735
steps: 198	lr: 4e-05	evaluation reward: 2.5	
episode: 1486	score: 5.0	memory length: 100000	epsilon: 0.6412774600077875
steps: 326	lr: 4e-05	evaluation reward: 2.49	
episode: 1487	score: 0.0	memory length: 100000	epsilon: 0.6410339200077928
steps: 123	lr: 4e-05	evaluation reward: 2.44	
episode: 1488	score: 5.0	memory length: 100000	epsilon: 0.6404082400078064
steps: 316	lr: 4e-05	evaluation reward: 2.44	
episode: 1489	score: 3.0	memory length: 100000	epsilon: 0.6399568000078162
steps: 228	lr: 4e-05	evaluation reward: 2.45	
episode: 1490	score: 3.0	memory length: 100000	epsilon: 0.6394697200078268
steps: 246	lr: 4e-05	evaluation reward: 2.46	
episode: 1491	score: 2.0	memory length: 100000	epsilon: 0.6390400600078361
steps: 217	lr: 4e-05	evaluation reward: 2.45	
episode: 1492	score: 2.0	memory length: 100000	epsilon: 0.6386500000078446
steps: 197	lr: 4e-05	evaluation reward: 2.43	
episode: 1493	score: 2.0	memory length: 100000	epsilon: 0.6382223200078538
steps: 216	lr: 4e-05	evaluation reward: 2.43	
episode: 1494	score: 3.0	memory length: 100000	epsilon: 0.6376916800078654
steps: 268	lr: 4e-05	evaluation reward: 2.44	
episode: 1495	score: 3.0	memory length: 100000	epsilon: 0.6371630200078768
steps: 267	lr: 4e-05	evaluation reward: 2.47	
episode: 1496	score: 3.0	memory length: 100000	epsilon: 0.6367135600078866
steps: 227	lr: 4e-05	evaluation reward: 2.44	
episode: 1497	score: 4.0	memory length: 100000	epsilon: 0.6362007400078977
steps: 259	lr: 4e-05	evaluation reward: 2.45	
episode: 1498	score: 4.0	memory length: 100000	epsilon: 0.6356522800079096
steps: 277	lr: 4e-05	evaluation reward: 2.48	
episode: 1499	score: 2.0	memory length: 100000	epsilon: 0.635222620007919

steps: 217	lr: 4e-05	evaluation reward: 2.47	
episode: 1500	score: 3.0	memory length: 100000	epsilon: 0.6347355400079295
steps: 246	lr: 4e-05	evaluation reward: 2.48	
episode: 1501	score: 2.0	memory length: 100000	epsilon: 0.634343500007938
steps: 198	lr: 4e-05	evaluation reward: 2.5	
episode: 1502	score: 2.0	memory length: 100000	epsilon: 0.6339534400079465
steps: 197	lr: 4e-05	evaluation reward: 2.52	
episode: 1503	score: 0.0	memory length: 100000	epsilon: 0.6337099000079518
steps: 123	lr: 4e-05	evaluation reward: 2.5	
episode: 1504	score: 1.0	memory length: 100000	epsilon: 0.6333713200079591
steps: 171	lr: 4e-05	evaluation reward: 2.51	
episode: 1505	score: 5.0	memory length: 100000	epsilon: 0.6326961400079738
steps: 341	lr: 4e-05	evaluation reward: 2.56	
episode: 1506	score: 3.0	memory length: 100000	epsilon: 0.6322486600079835
steps: 226	lr: 4e-05	evaluation reward: 2.56	
episode: 1507	score: 4.0	memory length: 100000	epsilon: 0.6316625800079962
steps: 296	lr: 4e-05	evaluation reward: 2.56	
episode: 1508	score: 2.0	memory length: 100000	epsilon: 0.6312725200080047
steps: 197	lr: 4e-05	evaluation reward: 2.56	
episode: 1509	score: 4.0	memory length: 100000	epsilon: 0.630797320008015
steps: 240	lr: 4e-05	evaluation reward: 2.6	
episode: 1510	score: 3.0	memory length: 100000	epsilon: 0.6303498400080247
steps: 226	lr: 4e-05	evaluation reward: 2.62	
episode: 1511	score: 2.0	memory length: 100000	epsilon: 0.6299578000080333
steps: 198	lr: 4e-05	evaluation reward: 2.63	
episode: 1512	score: 4.0	memory length: 100000	epsilon: 0.6294449800080444
steps: 259	lr: 4e-05	evaluation reward: 2.63	
episode: 1513	score: 3.0	memory length: 100000	epsilon: 0.6289876000080543
steps: 231	lr: 4e-05	evaluation reward: 2.62	
episode: 1514	score: 1.0	memory length: 100000	epsilon: 0.6286510000080616
steps: 170	lr: 4e-05	evaluation reward: 2.59	
episode: 1515	score: 3.0	memory length: 100000	epsilon: 0.6281560000080724
steps: 250	lr: 4e-05	evaluation reward: 2.58	
episode: 1516	score: 2.0	memory length: 100000	epsilon: 0.6277243600080817
steps: 218	lr: 4e-05	evaluation reward: 2.57	
episode: 1517	score: 4.0	memory length: 100000	epsilon: 0.6271798600080936
steps: 275	lr: 4e-05	evaluation reward: 2.55	
episode: 1518	score: 3.0	memory length: 100000	epsilon: 0.6266927800081041
steps: 246	lr: 4e-05	evaluation reward: 2.56	
episode: 1519	score: 2.0	memory length: 100000	epsilon: 0.6262967800081127
steps: 200	lr: 4e-05	evaluation reward: 2.55	
episode: 1520	score: 3.0	memory length: 100000	epsilon: 0.6258136600081232
steps: 244	lr: 4e-05	evaluation reward: 2.58	
episode: 1521	score: 3.0	memory length: 100000	epsilon: 0.6253602400081331
steps: 229	lr: 4e-05	evaluation reward: 2.57	
episode: 1522	score: 0.0	memory length: 100000	epsilon: 0.6251167000081383
steps: 123	lr: 4e-05	evaluation reward: 2.56	
episode: 1523	score: 1.0	memory length: 100000	epsilon: 0.6248177200081448

steps: 151	lr: 4e-05	evaluation reward: 2.53	
episode: 1524	score: 0.0	memory length: 100000	epsilon: 0.6245741800081501
steps: 123	lr: 4e-05	evaluation reward: 2.53	
episode: 1525	score: 4.0	memory length: 100000	epsilon: 0.6240930400081606
steps: 243	lr: 4e-05	evaluation reward: 2.56	
episode: 1526	score: 2.0	memory length: 100000	epsilon: 0.62365942000817
steps: 219	lr: 4e-05	evaluation reward: 2.55	
episode: 1527	score: 2.0	memory length: 100000	epsilon: 0.6232673800081785
steps: 198	lr: 4e-05	evaluation reward: 2.52	
episode: 1528	score: 4.0	memory length: 100000	epsilon: 0.6226773400081913
steps: 298	lr: 4e-05	evaluation reward: 2.54	
episode: 1529	score: 2.0	memory length: 100000	epsilon: 0.622322920008199
steps: 179	lr: 4e-05	evaluation reward: 2.53	
episode: 1530	score: 4.0	memory length: 100000	epsilon: 0.6217804000082108
steps: 274	lr: 4e-05	evaluation reward: 2.56	
episode: 1531	score: 2.0	memory length: 100000	epsilon: 0.6213844000082194
steps: 200	lr: 4e-05	evaluation reward: 2.56	
episode: 1532	score: 2.0	memory length: 100000	epsilon: 0.6209923600082279
steps: 198	lr: 4e-05	evaluation reward: 2.55	
episode: 1533	score: 3.0	memory length: 100000	epsilon: 0.6205033000082385
steps: 247	lr: 4e-05	evaluation reward: 2.56	
episode: 1534	score: 5.0	memory length: 100000	epsilon: 0.6198617800082524
steps: 324	lr: 4e-05	evaluation reward: 2.57	
episode: 1535	score: 3.0	memory length: 100000	epsilon: 0.6194143000082621
steps: 226	lr: 4e-05	evaluation reward: 2.59	
episode: 1536	score: 2.0	memory length: 100000	epsilon: 0.6190242400082706
steps: 197	lr: 4e-05	evaluation reward: 2.57	
episode: 1537	score: 4.0	memory length: 100000	epsilon: 0.6184718200082826
steps: 279	lr: 4e-05	evaluation reward: 2.6	
episode: 1538	score: 4.0	memory length: 100000	epsilon: 0.6178857400082953
steps: 296	lr: 4e-05	evaluation reward: 2.61	
episode: 1539	score: 5.0	memory length: 100000	epsilon: 0.6172838200083084
steps: 304	lr: 4e-05	evaluation reward: 2.64	
episode: 1540	score: 4.0	memory length: 100000	epsilon: 0.6167373400083203
steps: 276	lr: 4e-05	evaluation reward: 2.63	
episode: 1541	score: 2.0	memory length: 100000	epsilon: 0.6163453000083288
steps: 198	lr: 4e-05	evaluation reward: 2.64	
episode: 1542	score: 3.0	memory length: 100000	epsilon: 0.6158899000083387
steps: 230	lr: 4e-05	evaluation reward: 2.65	
episode: 1543	score: 3.0	memory length: 100000	epsilon: 0.6154384600083485
steps: 228	lr: 4e-05	evaluation reward: 2.65	
episode: 1544	score: 3.0	memory length: 100000	epsilon: 0.6149474200083591
steps: 248	lr: 4e-05	evaluation reward: 2.66	
episode: 1545	score: 4.0	memory length: 100000	epsilon: 0.6144643000083696
steps: 244	lr: 4e-05	evaluation reward: 2.69	
episode: 1546	score: 10.0	memory length: 100000	epsilon:
0.6136742800083868	steps: 399	lr: 4e-05	evaluation reward: 2.75
episode: 1547	score: 1.0	memory length: 100000	epsilon: 0.6133357000083941

steps: 171	lr: 4e-05	evaluation reward: 2.74	
episode: 1548	score: 2.0	memory length: 100000	epsilon: 0.6129357400084028
steps: 202	lr: 4e-05	evaluation reward: 2.71	
episode: 1549	score: 3.0	memory length: 100000	epsilon: 0.6125199400084118
steps: 210	lr: 4e-05	evaluation reward: 2.74	
episode: 1550	score: 3.0	memory length: 100000	epsilon: 0.6120605800084218
steps: 232	lr: 4e-05	evaluation reward: 2.74	
episode: 1551	score: 2.0	memory length: 100000	epsilon: 0.6116269600084312
steps: 219	lr: 4e-05	evaluation reward: 2.75	
episode: 1552	score: 4.0	memory length: 100000	epsilon: 0.6110804800084431
steps: 276	lr: 4e-05	evaluation reward: 2.78	
episode: 1553	score: 4.0	memory length: 100000	epsilon: 0.610530040008455
steps: 278	lr: 4e-05	evaluation reward: 2.8	
episode: 1554	score: 4.0	memory length: 100000	epsilon: 0.6099855400084668
steps: 275	lr: 4e-05	evaluation reward: 2.84	
episode: 1555	score: 6.0	memory length: 100000	epsilon: 0.6092489800084828
steps: 372	lr: 4e-05	evaluation reward: 2.85	
episode: 1556	score: 1.0	memory length: 100000	epsilon: 0.6089500000084893
steps: 151	lr: 4e-05	evaluation reward: 2.83	
episode: 1557	score: 2.0	memory length: 100000	epsilon: 0.6085599400084978
steps: 197	lr: 4e-05	evaluation reward: 2.84	
episode: 1558	score: 1.0	memory length: 100000	epsilon: 0.6082213600085051
steps: 171	lr: 4e-05	evaluation reward: 2.83	
episode: 1559	score: 3.0	memory length: 100000	epsilon: 0.607767940008515
steps: 229	lr: 4e-05	evaluation reward: 2.83	
episode: 1560	score: 4.0	memory length: 100000	epsilon: 0.6072630400085259
steps: 255	lr: 4e-05	evaluation reward: 2.85	
episode: 1561	score: 6.0	memory length: 100000	epsilon: 0.6065918200085405
steps: 339	lr: 4e-05	evaluation reward: 2.89	
episode: 1562	score: 3.0	memory length: 100000	epsilon: 0.6061403800085503
steps: 228	lr: 4e-05	evaluation reward: 2.9	
episode: 1563	score: 3.0	memory length: 100000	epsilon: 0.6056137000085617
steps: 266	lr: 4e-05	evaluation reward: 2.93	
episode: 1564	score: 2.0	memory length: 100000	epsilon: 0.6052216600085703
steps: 198	lr: 4e-05	evaluation reward: 2.92	
episode: 1565	score: 3.0	memory length: 100000	epsilon: 0.6047761600085799
steps: 225	lr: 4e-05	evaluation reward: 2.93	
episode: 1566	score: 3.0	memory length: 100000	epsilon: 0.6043286800085896
steps: 226	lr: 4e-05	evaluation reward: 2.95	
episode: 1567	score: 1.0	memory length: 100000	epsilon: 0.6040297000085961
steps: 151	lr: 4e-05	evaluation reward: 2.95	
episode: 1568	score: 4.0	memory length: 100000	epsilon: 0.603485200008608
steps: 275	lr: 1.6000000000000003e-05	evaluation reward: 2.96	
episode: 1569	score: 1.0	memory length: 100000	epsilon: 0.6031505800086152
steps: 169	lr: 1.6000000000000003e-05	evaluation reward: 2.92	
episode: 1570	score: 4.0	memory length: 100000	epsilon: 0.6025664800086279
steps: 295	lr: 1.6000000000000003e-05	evaluation reward: 2.93	
episode: 1571	score: 3.0	memory length: 100000	epsilon: 0.6020754400086386

```

steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 2.93
episode: 1572  score: 2.0    memory length: 100000    epsilon: 0.6016834000086471
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 2.89
episode: 1573  score: 3.0    memory length: 100000    epsilon: 0.6012319600086569
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 2.91
episode: 1574  score: 5.0    memory length: 100000    epsilon: 0.6005409400086719
steps: 349    lr: 1.6000000000000003e-05    evaluation reward: 2.94
episode: 1575  score: 3.0    memory length: 100000    epsilon: 0.6000122800086833
steps: 267    lr: 1.6000000000000003e-05    evaluation reward: 2.95
episode: 1576  score: 4.0    memory length: 100000    epsilon: 0.5994242200086961
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 2.97
episode: 1577  score: 1.0    memory length: 100000    epsilon: 0.5991252400087026
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1578  score: 3.0    memory length: 100000    epsilon: 0.5986777600087123
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 2.95
episode: 1579  score: 4.0    memory length: 100000    epsilon: 0.5981332600087241
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 2.91
episode: 1580  score: 6.0    memory length: 100000    epsilon: 0.5974264000087395
steps: 357    lr: 1.6000000000000003e-05    evaluation reward: 2.94
episode: 1581  score: 1.0    memory length: 100000    epsilon: 0.597127420008746
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 2.93
episode: 1582  score: 0.0    memory length: 100000    epsilon: 0.5968838800087513
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 2.91
episode: 1583  score: 5.0    memory length: 100000    epsilon: 0.5961671200087668
steps: 362    lr: 1.6000000000000003e-05    evaluation reward: 2.92
episode: 1584  score: 6.0    memory length: 100000    epsilon: 0.5954879800087816
steps: 343    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1585  score: 2.0    memory length: 100000    epsilon: 0.5950563400087909
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1586  score: 5.0    memory length: 100000    epsilon: 0.5944009600088052
steps: 331    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1587  score: 3.0    memory length: 100000    epsilon: 0.593947540008815
steps: 229    lr: 1.6000000000000003e-05    evaluation reward: 2.99
episode: 1588  score: 5.0    memory length: 100000    epsilon: 0.5932703800088297
steps: 342    lr: 1.6000000000000003e-05    evaluation reward: 2.99
episode: 1589  score: 5.0    memory length: 100000    epsilon: 0.5926645000088429
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1590  score: 1.0    memory length: 100000    epsilon: 0.5923298800088501
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 2.99
episode: 1591  score: 4.0    memory length: 100000    epsilon: 0.5918190400088612
steps: 258    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1592  score: 3.0    memory length: 100000    epsilon: 0.591365620008871
steps: 229    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1593  score: 6.0    memory length: 100000    epsilon: 0.5905518400088887
steps: 411    lr: 1.6000000000000003e-05    evaluation reward: 3.06
episode: 1594  score: 2.0    memory length: 100000    epsilon: 0.5901578200088973
steps: 199    lr: 1.6000000000000003e-05    evaluation reward: 3.05
episode: 1595  score: 0.0    memory length: 100000    epsilon: 0.5899142800089026

```

```

steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1596  score: 3.0    memory length: 100000    epsilon: 0.5894668000089123
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1597  score: 2.0    memory length: 100000    epsilon: 0.5890767400089207
steps: 197    lr: 1.6000000000000003e-05    evaluation reward: 3.0
episode: 1598  score: 3.0    memory length: 100000    epsilon: 0.5885896600089313
steps: 246    lr: 1.6000000000000003e-05    evaluation reward: 2.99
episode: 1599  score: 4.0    memory length: 100000    epsilon: 0.5880412000089432
steps: 277    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1600  score: 3.0    memory length: 100000    epsilon: 0.5875541200089538
steps: 246    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1601  score: 3.0    memory length: 100000    epsilon: 0.5870650600089644
steps: 247    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1602  score: 4.0    memory length: 100000    epsilon: 0.5865205600089762
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.04
episode: 1603  score: 3.0    memory length: 100000    epsilon: 0.5860315000089868
steps: 247    lr: 1.6000000000000003e-05    evaluation reward: 3.07
episode: 1604  score: 3.0    memory length: 100000    epsilon: 0.5856137200089959
steps: 211    lr: 1.6000000000000003e-05    evaluation reward: 3.09
episode: 1605  score: 2.0    memory length: 100000    epsilon: 0.5851820800090053
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.06
episode: 1606  score: 6.0    memory length: 100000    epsilon: 0.5844415600090214
steps: 374    lr: 1.6000000000000003e-05    evaluation reward: 3.09
episode: 1607  score: 1.0    memory length: 100000    epsilon: 0.5841029800090287
steps: 171    lr: 1.6000000000000003e-05    evaluation reward: 3.06
episode: 1608  score: 0.0    memory length: 100000    epsilon: 0.583859440009034
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.04
episode: 1609  score: 3.0    memory length: 100000    epsilon: 0.5833723600090446
steps: 246    lr: 1.6000000000000003e-05    evaluation reward: 3.03
episode: 1610  score: 3.0    memory length: 100000    epsilon: 0.5829248800090543
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.03
episode: 1611  score: 1.0    memory length: 100000    epsilon: 0.5826259000090608
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1612  score: 2.0    memory length: 100000    epsilon: 0.5822695000090685
steps: 180    lr: 1.6000000000000003e-05    evaluation reward: 3.0
episode: 1613  score: 4.0    memory length: 100000    epsilon: 0.5817151000090806
steps: 280    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1614  score: 2.0    memory length: 100000    epsilon: 0.5813230600090891
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1615  score: 3.0    memory length: 100000    epsilon: 0.5808755800090988
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1616  score: 2.0    memory length: 100000    epsilon: 0.5804835400091073
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1617  score: 4.0    memory length: 100000    epsilon: 0.57989746000912
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1618  score: 2.0    memory length: 100000    epsilon: 0.5794697800091293
steps: 216    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1619  score: 3.0    memory length: 100000    epsilon: 0.5790143800091392

```

```

steps: 230    lr: 1.6000000000000003e-05    evaluation reward: 3.02
episode: 1620  score: 9.0    memory length: 100000    epsilon: 0.5782243600091563
steps: 399    lr: 1.6000000000000003e-05    evaluation reward: 3.08
episode: 1621  score: 4.0    memory length: 100000    epsilon: 0.577642240009169
steps: 294    lr: 1.6000000000000003e-05    evaluation reward: 3.09
episode: 1622  score: 1.0    memory length: 100000    epsilon: 0.5773412800091755
steps: 152    lr: 1.6000000000000003e-05    evaluation reward: 3.1
episode: 1623  score: 6.0    memory length: 100000    epsilon: 0.5766364000091908
steps: 356    lr: 1.6000000000000003e-05    evaluation reward: 3.15
episode: 1624  score: 1.0    memory length: 100000    epsilon: 0.5763017800091981
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 3.16
episode: 1625  score: 6.0    memory length: 100000    epsilon: 0.5756048200092132
steps: 352    lr: 1.6000000000000003e-05    evaluation reward: 3.18
episode: 1626  score: 2.0    memory length: 100000    epsilon: 0.5752147600092217
steps: 197    lr: 1.6000000000000003e-05    evaluation reward: 3.18
episode: 1627  score: 4.0    memory length: 100000    epsilon: 0.5746306600092344
steps: 295    lr: 1.6000000000000003e-05    evaluation reward: 3.2
episode: 1628  score: 3.0    memory length: 100000    epsilon: 0.5741020000092458
steps: 267    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1629  score: 2.0    memory length: 100000    epsilon: 0.5736703600092552
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1630  score: 4.0    memory length: 100000    epsilon: 0.573127840009267
steps: 274    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1631  score: 2.0    memory length: 100000    epsilon: 0.5726942200092764
steps: 219    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1632  score: 3.0    memory length: 100000    epsilon: 0.5722467400092861
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.2
episode: 1633  score: 1.0    memory length: 100000    epsilon: 0.5719081600092935
steps: 171    lr: 1.6000000000000003e-05    evaluation reward: 3.18
episode: 1634  score: 4.0    memory length: 100000    epsilon: 0.5714349400093037
steps: 239    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1635  score: 5.0    memory length: 100000    epsilon: 0.5708488600093165
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1636  score: 2.0    memory length: 100000    epsilon: 0.5704172200093258
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1637  score: 4.0    memory length: 100000    epsilon: 0.5698707400093377
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1638  score: 5.0    memory length: 100000    epsilon: 0.5692648600093508
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 3.2
episode: 1639  score: 2.0    memory length: 100000    epsilon: 0.5689084600093586
steps: 180    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1640  score: 12.0   memory length: 100000    epsilon:
0.5679976600093783    steps: 460    lr: 1.6000000000000003e-05    evaluation
reward: 3.25
episode: 1641  score: 3.0    memory length: 100000    epsilon: 0.5675501800093881
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.26
episode: 1642  score: 7.0    memory length: 100000    epsilon: 0.5667502600094054
steps: 404    lr: 1.6000000000000003e-05    evaluation reward: 3.3

```


episode: 1643 score: 3.0 memory length: 100000 epsilon: 0.5663047600094151
 steps: 225 lr: 1.6000000000000003e-05 evaluation reward: 3.3
 episode: 1644 score: 3.0 memory length: 100000 epsilon: 0.5658157000094257
 steps: 247 lr: 1.6000000000000003e-05 evaluation reward: 3.3
 episode: 1645 score: 4.0 memory length: 100000 epsilon: 0.5653385200094361
 steps: 241 lr: 1.6000000000000003e-05 evaluation reward: 3.3
 episode: 1646 score: 1.0 memory length: 100000 epsilon: 0.5650395400094426
 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 3.21
 episode: 1647 score: 0.0 memory length: 100000 epsilon: 0.5647960000094479
 steps: 123 lr: 1.6000000000000003e-05 evaluation reward: 3.2
 episode: 1648 score: 5.0 memory length: 100000 epsilon: 0.5641980400094608
 steps: 302 lr: 1.6000000000000003e-05 evaluation reward: 3.23
 episode: 1649 score: 2.0 memory length: 100000 epsilon: 0.5638020400094694
 steps: 200 lr: 1.6000000000000003e-05 evaluation reward: 3.22
 episode: 1650 score: 4.0 memory length: 100000 epsilon: 0.5632179400094821
 steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 3.23
 episode: 1651 score: 3.0 memory length: 100000 epsilon: 0.5626793800094938
 steps: 272 lr: 1.6000000000000003e-05 evaluation reward: 3.24
 episode: 1652 score: 7.0 memory length: 100000 epsilon: 0.5618774800095112
 steps: 405 lr: 1.6000000000000003e-05 evaluation reward: 3.27
 episode: 1653 score: 3.0 memory length: 100000 epsilon: 0.5613943600095217
 steps: 244 lr: 1.6000000000000003e-05 evaluation reward: 3.26
 episode: 1654 score: 2.0 memory length: 100000 epsilon: 0.5610023200095302
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.24
 episode: 1655 score: 5.0 memory length: 100000 epsilon: 0.5603964400095434
 steps: 306 lr: 1.6000000000000003e-05 evaluation reward: 3.23
 episode: 1656 score: 6.0 memory length: 100000 epsilon: 0.5596935400095586
 steps: 355 lr: 1.6000000000000003e-05 evaluation reward: 3.28
 episode: 1657 score: 4.0 memory length: 100000 epsilon: 0.5591470600095705
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.3
 episode: 1658 score: 9.0 memory length: 100000 epsilon: 0.5584065400095866
 steps: 374 lr: 1.6000000000000003e-05 evaluation reward: 3.38
 episode: 1659 score: 3.0 memory length: 100000 epsilon: 0.5579194600095971
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 3.38
 episode: 1660 score: 1.0 memory length: 100000 epsilon: 0.5576165200096037
 steps: 153 lr: 1.6000000000000003e-05 evaluation reward: 3.35
 episode: 1661 score: 3.0 memory length: 100000 epsilon: 0.5571690400096134
 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 3.32
 episode: 1662 score: 5.0 memory length: 100000 epsilon: 0.5565592000096267
 steps: 308 lr: 1.6000000000000003e-05 evaluation reward: 3.34
 episode: 1663 score: 3.0 memory length: 100000 epsilon: 0.5560780600096371
 steps: 243 lr: 1.6000000000000003e-05 evaluation reward: 3.34
 episode: 1664 score: 6.0 memory length: 100000 epsilon: 0.5552959600096541
 steps: 395 lr: 1.6000000000000003e-05 evaluation reward: 3.38
 episode: 1665 score: 4.0 memory length: 100000 epsilon: 0.5547851200096652
 steps: 258 lr: 1.6000000000000003e-05 evaluation reward: 3.39
 episode: 1666 score: 3.0 memory length: 100000 epsilon: 0.5542525000096767
 steps: 269 lr: 1.6000000000000003e-05 evaluation reward: 3.39

episode: 1667 score: 4.0 memory length: 100000 epsilon: 0.5537476000096877
 steps: 255 lr: 1.6000000000000003e-05 evaluation reward: 3.42
 episode: 1668 score: 5.0 memory length: 100000 epsilon: 0.5531377600097009
 steps: 308 lr: 1.6000000000000003e-05 evaluation reward: 3.43
 episode: 1669 score: 2.0 memory length: 100000 epsilon: 0.5527457200097095
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.44
 episode: 1670 score: 2.0 memory length: 100000 epsilon: 0.552353680009718
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.42
 episode: 1671 score: 4.0 memory length: 100000 epsilon: 0.55179928000973
 steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 3.43
 episode: 1672 score: 4.0 memory length: 100000 epsilon: 0.5512547800097418
 steps: 275 lr: 1.6000000000000003e-05 evaluation reward: 3.45
 episode: 1673 score: 1.0 memory length: 100000 epsilon: 0.5509558000097483
 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 3.43
 episode: 1674 score: 4.0 memory length: 100000 epsilon: 0.5504113000097601
 steps: 275 lr: 1.6000000000000003e-05 evaluation reward: 3.42
 episode: 1675 score: 3.0 memory length: 100000 epsilon: 0.5499638200097698
 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 3.42
 episode: 1676 score: 5.0 memory length: 100000 epsilon: 0.5493975400097821
 steps: 286 lr: 1.6000000000000003e-05 evaluation reward: 3.43
 episode: 1677 score: 2.0 memory length: 100000 epsilon: 0.5490055000097906
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.44
 episode: 1678 score: 1.0 memory length: 100000 epsilon: 0.5487045400097972
 steps: 152 lr: 1.6000000000000003e-05 evaluation reward: 3.42
 episode: 1679 score: 4.0 memory length: 100000 epsilon: 0.548158060009809
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.42
 episode: 1680 score: 2.0 memory length: 100000 epsilon: 0.5477660200098176
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.38
 episode: 1681 score: 3.0 memory length: 100000 epsilon: 0.5473522000098265
 steps: 209 lr: 1.6000000000000003e-05 evaluation reward: 3.4
 episode: 1682 score: 4.0 memory length: 100000 epsilon: 0.5468077000098384
 steps: 275 lr: 1.6000000000000003e-05 evaluation reward: 3.44
 episode: 1683 score: 0.0 memory length: 100000 epsilon: 0.5465641600098436
 steps: 123 lr: 1.6000000000000003e-05 evaluation reward: 3.39
 episode: 1684 score: 3.0 memory length: 100000 epsilon: 0.5461107400098535
 steps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.36
 episode: 1685 score: 6.0 memory length: 100000 epsilon: 0.5454058600098688
 steps: 356 lr: 1.6000000000000003e-05 evaluation reward: 3.4
 episode: 1686 score: 4.0 memory length: 100000 epsilon: 0.5448554200098807
 steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 3.39
 episode: 1687 score: 3.0 memory length: 100000 epsilon: 0.5443287400098922
 steps: 266 lr: 1.6000000000000003e-05 evaluation reward: 3.39
 episode: 1688 score: 4.0 memory length: 100000 epsilon: 0.5437406800099049
 steps: 297 lr: 1.6000000000000003e-05 evaluation reward: 3.38
 episode: 1689 score: 2.0 memory length: 100000 epsilon: 0.5433446800099135
 steps: 200 lr: 1.6000000000000003e-05 evaluation reward: 3.35
 episode: 1690 score: 2.0 memory length: 100000 epsilon: 0.542954620009922
 steps: 197 lr: 1.6000000000000003e-05 evaluation reward: 3.36

episode: 1691 score: 4.0 memory length: 100000 epsilon: 0.5424774400099324
 steps: 241 lr: 1.6000000000000003e-05 evaluation reward: 3.36
 episode: 1692 score: 2.0 memory length: 100000 epsilon: 0.5420854000099409
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.35
 episode: 1693 score: 5.0 memory length: 100000 epsilon: 0.5414854600099539
 steps: 303 lr: 1.6000000000000003e-05 evaluation reward: 3.34
 episode: 1694 score: 3.0 memory length: 100000 epsilon: 0.5409587800099653
 steps: 266 lr: 1.6000000000000003e-05 evaluation reward: 3.35
 episode: 1695 score: 4.0 memory length: 100000 epsilon: 0.5403707200099781
 steps: 297 lr: 1.6000000000000003e-05 evaluation reward: 3.39
 episode: 1696 score: 3.0 memory length: 100000 epsilon: 0.5399173000099879
 steps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.39
 episode: 1697 score: 4.0 memory length: 100000 epsilon: 0.5393767600099997
 steps: 273 lr: 1.6000000000000003e-05 evaluation reward: 3.41
 episode: 1698 score: 3.0 memory length: 100000 epsilon: 0.5389609600100087
 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 3.41
 episode: 1699 score: 5.0 memory length: 100000 epsilon: 0.5383214200100226
 steps: 323 lr: 1.6000000000000003e-05 evaluation reward: 3.42
 episode: 1700 score: 7.0 memory length: 100000 epsilon: 0.537567040010039
 steps: 381 lr: 1.6000000000000003e-05 evaluation reward: 3.46
 episode: 1701 score: 2.0 memory length: 100000 epsilon: 0.5371334200100484
 steps: 219 lr: 1.6000000000000003e-05 evaluation reward: 3.45
 episode: 1702 score: 7.0 memory length: 100000 epsilon: 0.5363077600100663
 steps: 417 lr: 1.6000000000000003e-05 evaluation reward: 3.48
 episode: 1703 score: 5.0 memory length: 100000 epsilon: 0.535630600010081
 steps: 342 lr: 1.6000000000000003e-05 evaluation reward: 3.5
 episode: 1704 score: 3.0 memory length: 100000 epsilon: 0.5351435200100916
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 3.5
 episode: 1705 score: 1.0 memory length: 100000 epsilon: 0.5348445400100981
 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 3.49
 episode: 1706 score: 5.0 memory length: 100000 epsilon: 0.5342743000101104
 steps: 288 lr: 1.6000000000000003e-05 evaluation reward: 3.48
 episode: 1707 score: 2.0 memory length: 100000 epsilon: 0.533878300010119
 steps: 200 lr: 1.6000000000000003e-05 evaluation reward: 3.49
 episode: 1708 score: 5.0 memory length: 100000 epsilon: 0.5331971800101338
 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 3.54
 episode: 1709 score: 5.0 memory length: 100000 epsilon: 0.5326309000101461
 steps: 286 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1710 score: 3.0 memory length: 100000 epsilon: 0.5321438200101567
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1711 score: 1.0 memory length: 100000 epsilon: 0.531809200010164
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1712 score: 5.0 memory length: 100000 epsilon: 0.531116200010179
 steps: 350 lr: 1.6000000000000003e-05 evaluation reward: 3.59
 episode: 1713 score: 2.0 memory length: 100000 epsilon: 0.5307598000101867
 steps: 180 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1714 score: 4.0 memory length: 100000 epsilon: 0.5302113400101987
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 3.59

episode: 1715 score: 2.0 memory length: 100000 epsilon: 0.5298193000102072
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.58
 episode: 1716 score: 6.0 memory length: 100000 epsilon: 0.5290708600102234
 steps: 378 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1717 score: 4.0 memory length: 100000 epsilon: 0.5284788400102363
 steps: 299 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1718 score: 5.0 memory length: 100000 epsilon: 0.5278333600102503
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 3.65
 episode: 1719 score: 5.0 memory length: 100000 epsilon: 0.5271958000102641
 steps: 322 lr: 1.6000000000000003e-05 evaluation reward: 3.67
 episode: 1720 score: 1.0 memory length: 100000 epsilon: 0.5268948400102706
 steps: 152 lr: 1.6000000000000003e-05 evaluation reward: 3.59
 episode: 1721 score: 3.0 memory length: 100000 epsilon: 0.5264374600102806
 steps: 231 lr: 1.6000000000000003e-05 evaluation reward: 3.58
 episode: 1722 score: 1.0 memory length: 100000 epsilon: 0.526140460010287
 steps: 150 lr: 1.6000000000000003e-05 evaluation reward: 3.58
 episode: 1723 score: 5.0 memory length: 100000 epsilon: 0.525498940010301
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1724 score: 5.0 memory length: 100000 epsilon: 0.5248891000103142
 steps: 308 lr: 1.6000000000000003e-05 evaluation reward: 3.61
 episode: 1725 score: 1.0 memory length: 100000 epsilon: 0.5245544800103215
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1726 score: 3.0 memory length: 100000 epsilon: 0.5240634400103321
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1727 score: 2.0 memory length: 100000 epsilon: 0.5237030800103399
 steps: 182 lr: 1.6000000000000003e-05 evaluation reward: 3.55
 episode: 1728 score: 4.0 memory length: 100000 epsilon: 0.5230912600103532
 steps: 309 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1729 score: 3.0 memory length: 100000 epsilon: 0.5226437800103629
 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1730 score: 4.0 memory length: 100000 epsilon: 0.522134920010374
 steps: 257 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1731 score: 7.0 memory length: 100000 epsilon: 0.5214201400103895
 steps: 361 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1732 score: 5.0 memory length: 100000 epsilon: 0.5207746600104035
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 3.64
 episode: 1733 score: 1.0 memory length: 100000 epsilon: 0.52047568001041
 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 3.64
 episode: 1734 score: 3.0 memory length: 100000 epsilon: 0.5200579000104191
 steps: 211 lr: 1.6000000000000003e-05 evaluation reward: 3.63
 episode: 1735 score: 4.0 memory length: 100000 epsilon: 0.5194718200104318
 steps: 296 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1736 score: 3.0 memory length: 100000 epsilon: 0.5189847400104424
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 3.63
 episode: 1737 score: 5.0 memory length: 100000 epsilon: 0.518309560010457
 steps: 341 lr: 1.6000000000000003e-05 evaluation reward: 3.64
 episode: 1738 score: 2.0 memory length: 100000 epsilon: 0.5178779200104664
 steps: 218 lr: 1.6000000000000003e-05 evaluation reward: 3.61

episode: 1739 score: 5.0 memory length: 100000 epsilon: 0.5172344200104804
 steps: 325 lr: 1.6000000000000003e-05 evaluation reward: 3.64
 episode: 1740 score: 4.0 memory length: 100000 epsilon: 0.5166800200104924
 steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1741 score: 4.0 memory length: 100000 epsilon: 0.5160959200105051
 steps: 295 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1742 score: 2.0 memory length: 100000 epsilon: 0.5156682400105144
 steps: 216 lr: 1.6000000000000003e-05 evaluation reward: 3.52
 episode: 1743 score: 3.0 memory length: 100000 epsilon: 0.5152128400105243
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 3.52
 episode: 1744 score: 7.0 memory length: 100000 epsilon: 0.5144188600105415
 steps: 401 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1745 score: 3.0 memory length: 100000 epsilon: 0.5138842600105531
 steps: 270 lr: 1.6000000000000003e-05 evaluation reward: 3.55
 episode: 1746 score: 2.0 memory length: 100000 epsilon: 0.5134922200105616
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1747 score: 2.0 memory length: 100000 epsilon: 0.5131338400105694
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 3.58
 episode: 1748 score: 4.0 memory length: 100000 epsilon: 0.5125913200105812
 steps: 274 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1749 score: 7.0 memory length: 100000 epsilon: 0.5117478400105995
 steps: 426 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1750 score: 4.0 memory length: 100000 epsilon: 0.511169680010612
 steps: 292 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1751 score: 2.0 memory length: 100000 epsilon: 0.5107776400106205
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.61
 episode: 1752 score: 3.0 memory length: 100000 epsilon: 0.510294520010631
 steps: 244 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1753 score: 1.0 memory length: 100000 epsilon: 0.5099599000106383
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 3.55
 episode: 1754 score: 2.0 memory length: 100000 epsilon: 0.5095282600106477
 steps: 218 lr: 1.6000000000000003e-05 evaluation reward: 3.55
 episode: 1755 score: 1.0 memory length: 100000 epsilon: 0.5091936400106549
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 3.51
 episode: 1756 score: 8.0 memory length: 100000 epsilon: 0.5083838200106725
 steps: 409 lr: 1.6000000000000003e-05 evaluation reward: 3.53
 episode: 1757 score: 3.0 memory length: 100000 epsilon: 0.5078591200106839
 steps: 265 lr: 1.6000000000000003e-05 evaluation reward: 3.52
 episode: 1758 score: 4.0 memory length: 100000 epsilon: 0.5073126400106958
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 3.47
 episode: 1759 score: 4.0 memory length: 100000 epsilon: 0.5067582400107078
 steps: 280 lr: 1.6000000000000003e-05 evaluation reward: 3.48
 episode: 1760 score: 6.0 memory length: 100000 epsilon: 0.5060632600107229
 steps: 351 lr: 1.6000000000000003e-05 evaluation reward: 3.53
 episode: 1761 score: 6.0 memory length: 100000 epsilon: 0.505320760010739
 steps: 375 lr: 1.6000000000000003e-05 evaluation reward: 3.56
 episode: 1762 score: 4.0 memory length: 100000 epsilon: 0.5046990400107525
 steps: 314 lr: 1.6000000000000003e-05 evaluation reward: 3.55

episode: 1763 score: 7.0 memory length: 100000 epsilon: 0.5039446600107689
 steps: 381 lr: 1.6000000000000003e-05 evaluation reward: 3.59
 episode: 1764 score: 5.0 memory length: 100000 epsilon: 0.503293240010783
 steps: 329 lr: 1.6000000000000003e-05 evaluation reward: 3.58
 episode: 1765 score: 3.0 memory length: 100000 epsilon: 0.5027626000107945
 steps: 268 lr: 1.6000000000000003e-05 evaluation reward: 3.57
 episode: 1766 score: 5.0 memory length: 100000 epsilon: 0.5021191000108085
 steps: 325 lr: 1.6000000000000003e-05 evaluation reward: 3.59
 episode: 1767 score: 7.0 memory length: 100000 epsilon: 0.5012340400108277
 steps: 447 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1768 score: 5.0 memory length: 100000 epsilon: 0.5005925200108416
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 3.62
 episode: 1769 score: 3.0 memory length: 100000 epsilon: 0.5001014800108523
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.63
 episode: 1770 score: 5.0 memory length: 100000 epsilon: 0.4995292600108515
 steps: 289 lr: 1.6000000000000003e-05 evaluation reward: 3.66
 episode: 1771 score: 6.0 memory length: 100000 epsilon:
 0.49887190001084736 steps: 332 lr: 1.6000000000000003e-05 evaluation
 reward: 3.68
 episode: 1772 score: 4.0 memory length: 100000 epsilon:
 0.49828384001084364 steps: 297 lr: 1.6000000000000003e-05 evaluation
 reward: 3.68
 episode: 1773 score: 6.0 memory length: 100000 epsilon:
 0.49754332001083895 steps: 374 lr: 1.6000000000000003e-05 evaluation
 reward: 3.73
 episode: 1774 score: 5.0 memory length: 100000 epsilon: 0.4969018000108349
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 3.74
 episode: 1775 score: 6.0 memory length: 100000 epsilon:
 0.49620286001083047 steps: 353 lr: 1.6000000000000003e-05 evaluation
 reward: 3.77
 episode: 1776 score: 3.0 memory length: 100000 epsilon: 0.4957534000108276
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.75
 episode: 1777 score: 3.0 memory length: 100000 epsilon: 0.4953079000108248
 steps: 225 lr: 1.6000000000000003e-05 evaluation reward: 3.76
 episode: 1778 score: 3.0 memory length: 100000 epsilon:
 0.49485448001082194 steps: 229 lr: 1.6000000000000003e-05 evaluation
 reward: 3.78
 episode: 1779 score: 5.0 memory length: 100000 epsilon: 0.4942129600108179
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 3.79
 episode: 1780 score: 3.0 memory length: 100000 epsilon:
 0.49376548001081505 steps: 226 lr: 1.6000000000000003e-05 evaluation
 reward: 3.8
 episode: 1781 score: 2.0 memory length: 100000 epsilon: 0.4934071000108128
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 3.79
 episode: 1782 score: 2.0 memory length: 100000 epsilon: 0.4930507000108105
 steps: 180 lr: 1.6000000000000003e-05 evaluation reward: 3.77
 episode: 1783 score: 2.0 memory length: 100000 epsilon: 0.4926230200108078
 steps: 216 lr: 1.6000000000000003e-05 evaluation reward: 3.79

episode: 1784 score: 5.0 memory length: 100000 epsilon:
 0.49197754001080374 steps: 326 lr: 1.6000000000000003e-05 evaluation
 reward: 3.81
 episode: 1785 score: 2.0 memory length: 100000 epsilon:
 0.49161718001080146 steps: 182 lr: 1.6000000000000003e-05 evaluation
 reward: 3.77
 episode: 1786 score: 0.0 memory length: 100000 epsilon: 0.4913736400107999
 steps: 123 lr: 1.6000000000000003e-05 evaluation reward: 3.73
 episode: 1787 score: 3.0 memory length: 100000 epsilon: 0.4909459600107972
 steps: 216 lr: 1.6000000000000003e-05 evaluation reward: 3.73
 episode: 1788 score: 8.0 memory length: 100000 epsilon: 0.4900906000107918
 steps: 432 lr: 1.6000000000000003e-05 evaluation reward: 3.77
 episode: 1789 score: 11.0 memory length: 100000 epsilon:
 0.4892966200107868 steps: 401 lr: 1.6000000000000003e-05 evaluation
 reward: 3.86
 episode: 1790 score: 3.0 memory length: 100000 epsilon:
 0.48884914001078394 steps: 226 lr: 1.6000000000000003e-05 evaluation
 reward: 3.87
 episode: 1791 score: 6.0 memory length: 100000 epsilon: 0.4881026800107792
 steps: 377 lr: 1.6000000000000003e-05 evaluation reward: 3.89
 episode: 1792 score: 3.0 memory length: 100000 epsilon: 0.4876532200107764
 steps: 227 lr: 1.6000000000000003e-05 evaluation reward: 3.9
 episode: 1793 score: 1.0 memory length: 100000 epsilon: 0.4873562200107745
 steps: 150 lr: 1.6000000000000003e-05 evaluation reward: 3.86
 episode: 1794 score: 2.0 memory length: 100000 epsilon:
 0.48692458001077177 steps: 218 lr: 1.6000000000000003e-05 evaluation
 reward: 3.85
 episode: 1795 score: 7.0 memory length: 100000 epsilon:
 0.48613456001076677 steps: 399 lr: 1.6000000000000003e-05 evaluation
 reward: 3.88
 episode: 1796 score: 2.0 memory length: 100000 epsilon: 0.4857425200107643
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1797 score: 6.0 memory length: 100000 epsilon:
 0.48491092001075903 steps: 420 lr: 1.6000000000000003e-05 evaluation
 reward: 3.89
 episode: 1798 score: 5.0 memory length: 100000 epsilon:
 0.48432682001075533 steps: 295 lr: 1.6000000000000003e-05 evaluation
 reward: 3.91
 episode: 1799 score: 6.0 memory length: 100000 epsilon: 0.4835783800107506
 steps: 378 lr: 1.6000000000000003e-05 evaluation reward: 3.92
 episode: 1800 score: 5.0 memory length: 100000 epsilon: 0.4828952800107463
 steps: 345 lr: 1.6000000000000003e-05 evaluation reward: 3.9
 episode: 1801 score: 6.0 memory length: 100000 epsilon:
 0.48213100001074144 steps: 386 lr: 1.6000000000000003e-05 evaluation
 reward: 3.94
 episode: 1802 score: 2.0 memory length: 100000 epsilon:
 0.48173896001073896 steps: 198 lr: 1.6000000000000003e-05 evaluation
 reward: 3.89

episode: 1803 score: 3.0 memory length: 100000 epsilon: 0.4812855400107361
 steps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1804 score: 2.0 memory length: 100000 epsilon: 0.4809251800107338
 steps: 182 lr: 1.6000000000000003e-05 evaluation reward: 3.86
 episode: 1805 score: 5.0 memory length: 100000 epsilon: 0.4802797000107297
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 3.9
 episode: 1806 score: 4.0 memory length: 100000 epsilon: 0.479693620010726
 steps: 296 lr: 1.6000000000000003e-05 evaluation reward: 3.89
 episode: 1807 score: 4.0 memory length: 100000 epsilon:
 0.47917486001072274 steps: 262 lr: 1.6000000000000003e-05 evaluation
 reward: 3.91
 episode: 1808 score: 3.0 memory length: 100000 epsilon: 0.4787273800107199
 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 3.89
 episode: 1809 score: 3.0 memory length: 100000 epsilon: 0.4782363400107168
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1810 score: 3.0 memory length: 100000 epsilon:
 0.47778886001071397 steps: 226 lr: 1.6000000000000003e-05 evaluation
 reward: 3.87
 episode: 1811 score: 5.0 memory length: 100000 epsilon: 0.4771117000107097
 steps: 342 lr: 1.6000000000000003e-05 evaluation reward: 3.91
 episode: 1812 score: 3.0 memory length: 100000 epsilon:
 0.47658304001070634 steps: 267 lr: 1.6000000000000003e-05 evaluation
 reward: 3.89
 episode: 1813 score: 4.0 memory length: 100000 epsilon: 0.4759890400107026
 steps: 300 lr: 1.6000000000000003e-05 evaluation reward: 3.91
 episode: 1814 score: 6.0 memory length: 100000 epsilon: 0.4753475200106985
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 3.93
 episode: 1815 score: 3.0 memory length: 100000 epsilon: 0.4748564800106954
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 3.94
 episode: 1816 score: 3.0 memory length: 100000 epsilon:
 0.47437336001069236 steps: 244 lr: 1.6000000000000003e-05 evaluation
 reward: 3.91
 episode: 1817 score: 4.0 memory length: 100000 epsilon: 0.4738229200106889
 steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 3.91
 episode: 1818 score: 4.0 memory length: 100000 epsilon: 0.4732744600106854
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 3.9
 episode: 1819 score: 2.0 memory length: 100000 epsilon: 0.4728448000106827
 steps: 217 lr: 1.6000000000000003e-05 evaluation reward: 3.87
 episode: 1820 score: 3.0 memory length: 100000 epsilon: 0.4723557400106796
 steps: 247 lr: 1.6000000000000003e-05 evaluation reward: 3.89
 episode: 1821 score: 5.0 memory length: 100000 epsilon:
 0.47164096001067507 steps: 361 lr: 1.6000000000000003e-05 evaluation
 reward: 3.91
 episode: 1822 score: 3.0 memory length: 100000 epsilon:
 0.47114794001067195 steps: 249 lr: 1.6000000000000003e-05 evaluation
 reward: 3.93
 episode: 1823 score: 3.0 memory length: 100000 epsilon: 0.4706945200106691
 steps: 229 lr: 1.6000000000000003e-05 evaluation reward: 3.91

episode: 1824 score: 3.0 memory length: 100000 epsilon: 0.470205460010666
 steps: 247 lr: 1.6000000000000003e-05 evaluation reward: 3.89
 episode: 1825 score: 5.0 memory length: 100000 epsilon: 0.4696550200106625
 steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 3.93
 episode: 1826 score: 7.0 memory length: 100000 epsilon: 0.4688135200106572
 steps: 425 lr: 1.6000000000000003e-05 evaluation reward: 3.97
 episode: 1827 score: 7.0 memory length: 100000 epsilon: 0.4680274600106522
 steps: 397 lr: 1.6000000000000003e-05 evaluation reward: 4.02
 episode: 1828 score: 5.0 memory length: 100000 epsilon: 0.4674255400106484
 steps: 304 lr: 1.6000000000000003e-05 evaluation reward: 4.03
 episode: 1829 score: 2.0 memory length: 100000 epsilon:
 0.46706716001064613 steps: 181 lr: 1.6000000000000003e-05 evaluation
 reward: 4.02
 episode: 1830 score: 2.0 memory length: 100000 epsilon: 0.4667107600106439
 steps: 180 lr: 1.6000000000000003e-05 evaluation reward: 4.0
 episode: 1831 score: 2.0 memory length: 100000 epsilon:
 0.46627714001064113 steps: 219 lr: 1.6000000000000003e-05 evaluation
 reward: 3.95
 episode: 1832 score: 1.0 memory length: 100000 epsilon: 0.465942520010639
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 3.91
 episode: 1833 score: 3.0 memory length: 100000 epsilon:
 0.46552078001063635 steps: 213 lr: 1.6000000000000003e-05 evaluation
 reward: 3.93
 episode: 1834 score: 3.0 memory length: 100000 epsilon: 0.4650693400106335
 steps: 228 lr: 1.6000000000000003e-05 evaluation reward: 3.93
 episode: 1835 score: 8.0 memory length: 100000 epsilon: 0.4640912200106273
 steps: 494 lr: 1.6000000000000003e-05 evaluation reward: 3.97
 episode: 1836 score: 6.0 memory length: 100000 epsilon: 0.4633447600106226
 steps: 377 lr: 1.6000000000000003e-05 evaluation reward: 4.0
 episode: 1837 score: 11.0 memory length: 100000 epsilon:
 0.4622597200106157 steps: 548 lr: 1.6000000000000003e-05 evaluation
 reward: 4.06
 episode: 1838 score: 3.0 memory length: 100000 epsilon: 0.4617686800106126
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 4.07
 episode: 1839 score: 3.0 memory length: 100000 epsilon: 0.4612816000106095
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.05
 episode: 1840 score: 3.0 memory length: 100000 epsilon:
 0.46079254001060643 steps: 247 lr: 1.6000000000000003e-05 evaluation
 reward: 4.04
 episode: 1841 score: 4.0 memory length: 100000 epsilon:
 0.46020844001060274 steps: 295 lr: 1.6000000000000003e-05 evaluation
 reward: 4.04
 episode: 1842 score: 5.0 memory length: 100000 epsilon: 0.4595669200105987
 steps: 324 lr: 1.6000000000000003e-05 evaluation reward: 4.07
 episode: 1843 score: 5.0 memory length: 100000 epsilon:
 0.45896302001059486 steps: 305 lr: 1.6000000000000003e-05 evaluation
 reward: 4.09
 episode: 1844 score: 3.0 memory length: 100000 epsilon: 0.458513560010592

```

steps: 227    lr: 1.6000000000000003e-05    evaluation reward: 4.05
episode: 1845    score: 3.0    memory length: 100000    epsilon:
0.45805816001058913    steps: 230    lr: 1.6000000000000003e-05    evaluation
reward: 4.05
episode: 1846    score: 4.0    memory length: 100000    epsilon: 0.4574681200105854
steps: 298    lr: 1.6000000000000003e-05    evaluation reward: 4.07
episode: 1847    score: 3.0    memory length: 100000    epsilon: 0.4569770800105823
steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 4.08
episode: 1848    score: 2.0    memory length: 100000    epsilon: 0.4565494000105796
steps: 216    lr: 1.6000000000000003e-05    evaluation reward: 4.06
episode: 1849    score: 10.0    memory length: 100000    epsilon:
0.45579898001057484    steps: 379    lr: 1.6000000000000003e-05    evaluation
reward: 4.09
episode: 1850    score: 3.0    memory length: 100000    epsilon:
0.45530992001057174    steps: 247    lr: 1.6000000000000003e-05    evaluation
reward: 4.08
episode: 1851    score: 2.0    memory length: 100000    epsilon: 0.4549515400105695
steps: 181    lr: 1.6000000000000003e-05    evaluation reward: 4.08
episode: 1852    score: 1.0    memory length: 100000    epsilon:
0.45461692001056736    steps: 169    lr: 1.6000000000000003e-05    evaluation
reward: 4.06
episode: 1853    score: 3.0    memory length: 100000    epsilon: 0.4541357800105643
steps: 243    lr: 1.6000000000000003e-05    evaluation reward: 4.08
episode: 1854    score: 7.0    memory length: 100000    epsilon:
0.45353980001056055    steps: 301    lr: 1.6000000000000003e-05    evaluation
reward: 4.13
episode: 1855    score: 7.0    memory length: 100000    epsilon: 0.4527874000105558
steps: 380    lr: 1.6000000000000003e-05    evaluation reward: 4.19
episode: 1856    score: 5.0    memory length: 100000    epsilon:
0.45217954001055194    steps: 307    lr: 1.6000000000000003e-05    evaluation
reward: 4.16
episode: 1857    score: 3.0    memory length: 100000    epsilon: 0.4517281000105491
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 4.16
episode: 1858    score: 5.0    memory length: 100000    epsilon: 0.4511281600105453
steps: 303    lr: 1.6000000000000003e-05    evaluation reward: 4.17
episode: 1859    score: 8.0    memory length: 100000    epsilon:
0.45028468001053995    steps: 426    lr: 1.6000000000000003e-05    evaluation
reward: 4.21
episode: 1860    score: 6.0    memory length: 100000    epsilon:
0.44958574001053553    steps: 353    lr: 1.6000000000000003e-05    evaluation
reward: 4.21
episode: 1861    score: 3.0    memory length: 100000    epsilon:
0.44909668001053243    steps: 247    lr: 1.6000000000000003e-05    evaluation
reward: 4.18
episode: 1862    score: 4.0    memory length: 100000    epsilon:
0.44851654001052876    steps: 293    lr: 1.6000000000000003e-05    evaluation
reward: 4.18
episode: 1863    score: 2.0    memory length: 100000    epsilon: 0.4481245000105263

```

```

steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 4.13
episode: 1864  score: 5.0  memory length: 100000  epsilon: 0.447447340010522
steps: 342    lr: 1.6000000000000003e-05    evaluation reward: 4.13
episode: 1865  score: 5.0  memory length: 100000  epsilon: 0.4467563200105176
steps: 349    lr: 1.6000000000000003e-05    evaluation reward: 4.15
episode: 1866  score: 1.0  memory length: 100000  epsilon:
0.44645932001051575  steps: 150  lr: 1.6000000000000003e-05    evaluation
reward: 4.11
episode: 1867  score: 3.0  memory length: 100000  epsilon: 0.4460435200105131
steps: 210    lr: 1.6000000000000003e-05    evaluation reward: 4.07
episode: 1868  score: 3.0  memory length: 100000  epsilon: 0.4456257400105105
steps: 211    lr: 1.6000000000000003e-05    evaluation reward: 4.05
episode: 1869  score: 4.0  memory length: 100000  epsilon: 0.445079260010507
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 4.06
episode: 1870  score: 7.0  memory length: 100000  epsilon:
0.44427538001050193  steps: 406  lr: 1.6000000000000003e-05    evaluation
reward: 4.08
episode: 1871  score: 4.0  memory length: 100000  epsilon:
0.44369524001049826  steps: 293  lr: 1.6000000000000003e-05    evaluation
reward: 4.06
episode: 1872  score: 6.0  memory length: 100000  epsilon: 0.4429586800104936
steps: 372    lr: 1.6000000000000003e-05    evaluation reward: 4.08
episode: 1873  score: 5.0  memory length: 100000  epsilon:
0.44235280001048977  steps: 306  lr: 1.6000000000000003e-05    evaluation
reward: 4.07
episode: 1874  score: 5.0  memory length: 100000  epsilon: 0.4417132600104857
steps: 323    lr: 1.6000000000000003e-05    evaluation reward: 4.07
episode: 1875  score: 5.0  memory length: 100000  epsilon: 0.4410737200104817
steps: 323    lr: 1.6000000000000003e-05    evaluation reward: 4.06
episode: 1876  score: 8.0  memory length: 100000  epsilon: 0.4402104400104762
steps: 436    lr: 1.6000000000000003e-05    evaluation reward: 4.11
episode: 1877  score: 6.0  memory length: 100000  epsilon:
0.43950754001047176  steps: 355  lr: 1.6000000000000003e-05    evaluation
reward: 4.14
episode: 1878  score: 4.0  memory length: 100000  epsilon:
0.43891948001046804  steps: 297  lr: 1.6000000000000003e-05    evaluation
reward: 4.15
episode: 1879  score: 3.0  memory length: 100000  epsilon:
0.43843240001046496  steps: 246  lr: 1.6000000000000003e-05    evaluation
reward: 4.13
episode: 1880  score: 7.0  memory length: 100000  epsilon:
0.43759486001045966  steps: 423  lr: 1.6000000000000003e-05    evaluation
reward: 4.17
episode: 1881  score: 3.0  memory length: 100000  epsilon: 0.4371434200104568
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 4.18
episode: 1882  score: 5.0  memory length: 100000  epsilon: 0.4365771400104532
steps: 286    lr: 1.6000000000000003e-05    evaluation reward: 4.21
episode: 1883  score: 4.0  memory length: 100000  epsilon: 0.4360366000104498

```

steps: 273 lr: 1.6000000000000003e-05 evaluation reward: 4.23
 episode: 1884 score: 3.0 memory length: 100000 epsilon:
 0.43550794001044646 steps: 267 lr: 1.6000000000000003e-05 evaluation
 reward: 4.21
 episode: 1885 score: 3.0 memory length: 100000 epsilon: 0.4350248200104434
 steps: 244 lr: 1.6000000000000003e-05 evaluation reward: 4.22
 episode: 1886 score: 4.0 memory length: 100000 epsilon:
 0.43447834001043995 steps: 276 lr: 1.6000000000000003e-05 evaluation
 reward: 4.26
 episode: 1887 score: 4.0 memory length: 100000 epsilon: 0.4339318600104365
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.27
 episode: 1888 score: 2.0 memory length: 100000 epsilon:
 0.43357546001043423 steps: 180 lr: 1.6000000000000003e-05 evaluation
 reward: 4.21
 episode: 1889 score: 6.0 memory length: 100000 epsilon:
 0.43292998001043015 steps: 326 lr: 1.6000000000000003e-05 evaluation
 reward: 4.16
 episode: 1890 score: 5.0 memory length: 100000 epsilon:
 0.43235776001042653 steps: 289 lr: 1.6000000000000003e-05 evaluation
 reward: 4.18
 episode: 1891 score: 5.0 memory length: 100000 epsilon:
 0.43171624001042247 steps: 324 lr: 1.6000000000000003e-05 evaluation
 reward: 4.17
 episode: 1892 score: 2.0 memory length: 100000 epsilon: 0.43132618001042
 steps: 197 lr: 1.6000000000000003e-05 evaluation reward: 4.16
 episode: 1893 score: 4.0 memory length: 100000 epsilon: 0.4307282200104162
 steps: 302 lr: 1.6000000000000003e-05 evaluation reward: 4.19
 episode: 1894 score: 4.0 memory length: 100000 epsilon: 0.4301401600104125
 steps: 297 lr: 1.6000000000000003e-05 evaluation reward: 4.21
 episode: 1895 score: 6.0 memory length: 100000 epsilon:
 0.42946696001040824 steps: 340 lr: 1.6000000000000003e-05 evaluation
 reward: 4.2
 episode: 1896 score: 4.0 memory length: 100000 epsilon: 0.4289204800104048
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.22
 episode: 1897 score: 4.0 memory length: 100000 epsilon:
 0.42837796001040135 steps: 274 lr: 1.6000000000000003e-05 evaluation
 reward: 4.2
 episode: 1898 score: 3.0 memory length: 100000 epsilon: 0.4279304800103985
 steps: 226 lr: 1.6000000000000003e-05 evaluation reward: 4.18
 episode: 1899 score: 5.0 memory length: 100000 epsilon:
 0.42729094001039447 steps: 323 lr: 1.6000000000000003e-05 evaluation
 reward: 4.17
 episode: 1900 score: 3.0 memory length: 100000 epsilon:
 0.42687712001039185 steps: 209 lr: 1.6000000000000003e-05 evaluation
 reward: 4.15
 episode: 1901 score: 2.0 memory length: 100000 epsilon: 0.4265207200103896
 steps: 180 lr: 1.6000000000000003e-05 evaluation reward: 4.11
 episode: 1902 score: 3.0 memory length: 100000 epsilon: 0.4260316600103865

```

steps: 247    lr: 1.6000000000000003e-05    evaluation reward: 4.12
episode: 1903  score: 3.0  memory length: 100000  epsilon: 0.4255386400103834
steps: 249    lr: 1.6000000000000003e-05    evaluation reward: 4.12
episode: 1904  score: 3.0  memory length: 100000  epsilon: 0.4250535400103803
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 4.13
episode: 1905  score: 5.0  memory length: 100000  epsilon: 0.4244060800103762
steps: 327    lr: 1.6000000000000003e-05    evaluation reward: 4.13
episode: 1906  score: 3.0  memory length: 100000  epsilon: 0.4239566200103734
steps: 227    lr: 1.6000000000000003e-05    evaluation reward: 4.12
episode: 1907  score: 4.0  memory length: 100000  epsilon:
0.42337054001036967  steps: 296  lr: 1.6000000000000003e-05    evaluation
reward: 4.12
episode: 1908  score: 3.0  memory length: 100000  epsilon:
0.42284584001036635  steps: 265  lr: 1.6000000000000003e-05    evaluation
reward: 4.12
episode: 1909  score: 7.0  memory length: 100000  epsilon: 0.4220795800103615
steps: 387    lr: 1.6000000000000003e-05    evaluation reward: 4.16
episode: 1910  score: 2.0  memory length: 100000  epsilon:
0.42164794001035877  steps: 218  lr: 1.6000000000000003e-05    evaluation
reward: 4.15
episode: 1911  score: 9.0  memory length: 100000  epsilon:
0.42074110001035303  steps: 458  lr: 1.6000000000000003e-05    evaluation
reward: 4.19
episode: 1912  score: 5.0  memory length: 100000  epsilon:
0.42004810001034865  steps: 350  lr: 1.6000000000000003e-05    evaluation
reward: 4.21
episode: 1913  score: 6.0  memory length: 100000  epsilon: 0.4193491600103442
steps: 353    lr: 1.6000000000000003e-05    evaluation reward: 4.23
episode: 1914  score: 3.0  memory length: 100000  epsilon:
0.41886010001034113  steps: 247  lr: 1.6000000000000003e-05    evaluation
reward: 4.2
episode: 1915  score: 4.0  memory length: 100000  epsilon:
0.41831164001033766  steps: 277  lr: 1.6000000000000003e-05    evaluation
reward: 4.21
episode: 1916  score: 4.0  memory length: 100000  epsilon: 0.4177968400103344
steps: 260    lr: 1.6000000000000003e-05    evaluation reward: 4.22
episode: 1917  score: 5.0  memory length: 100000  epsilon: 0.4172266000103308
steps: 288    lr: 1.6000000000000003e-05    evaluation reward: 4.23
episode: 1918  score: 4.0  memory length: 100000  epsilon: 0.4166741800103273
steps: 279    lr: 1.6000000000000003e-05    evaluation reward: 4.23
episode: 1919  score: 9.0  memory length: 100000  epsilon: 0.4156584400103209
steps: 513    lr: 1.6000000000000003e-05    evaluation reward: 4.3
episode: 1920  score: 5.0  memory length: 100000  epsilon: 0.4150129600103168
steps: 326    lr: 1.6000000000000003e-05    evaluation reward: 4.32
episode: 1921  score: 8.0  memory length: 100000  epsilon:
0.41407444001031085  steps: 474  lr: 1.6000000000000003e-05    evaluation
reward: 4.35
episode: 1922  score: 4.0  memory length: 100000  epsilon:

```

0.41356756001030764 steps: 256 lr: 1.6000000000000003e-05 evaluation
 reward: 4.36
 episode: 1923 score: 5.0 memory length: 100000 epsilon: 0.4129715800103039
 steps: 301 lr: 1.6000000000000003e-05 evaluation reward: 4.38
 episode: 1924 score: 4.0 memory length: 100000 epsilon:
 0.41241520001030035 steps: 281 lr: 1.6000000000000003e-05 evaluation
 reward: 4.39
 episode: 1925 score: 2.0 memory length: 100000 epsilon: 0.4120231600102979
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 4.36
 episode: 1926 score: 6.0 memory length: 100000 epsilon:
 0.41124700001029296 steps: 392 lr: 1.6000000000000003e-05 evaluation
 reward: 4.35
 episode: 1927 score: 3.0 memory length: 100000 epsilon:
 0.41079952001029013 steps: 226 lr: 1.6000000000000003e-05 evaluation
 reward: 4.31
 episode: 1928 score: 7.0 memory length: 100000 epsilon: 0.4100708800102855
 steps: 368 lr: 1.6000000000000003e-05 evaluation reward: 4.33
 episode: 1929 score: 5.0 memory length: 100000 epsilon: 0.4094313400102815
 steps: 323 lr: 1.6000000000000003e-05 evaluation reward: 4.36
 episode: 1930 score: 4.0 memory length: 100000 epsilon:
 0.40891852001027823 steps: 259 lr: 1.6000000000000003e-05 evaluation
 reward: 4.38
 episode: 1931 score: 3.0 memory length: 100000 epsilon:
 0.40843342001027516 steps: 245 lr: 1.6000000000000003e-05 evaluation
 reward: 4.39
 episode: 1932 score: 1.0 memory length: 100000 epsilon:
 0.40810078001027306 steps: 168 lr: 1.6000000000000003e-05 evaluation
 reward: 4.39
 episode: 1933 score: 12.0 memory length: 100000 epsilon:
 0.40713652001026696 steps: 487 lr: 1.6000000000000003e-05 evaluation
 reward: 4.48
 episode: 1934 score: 4.0 memory length: 100000 epsilon: 0.4065900400102635
 steps: 276 lr: 1.6000000000000003e-05 evaluation reward: 4.49
 episode: 1935 score: 3.0 memory length: 100000 epsilon: 0.4060990000102604
 steps: 248 lr: 1.6000000000000003e-05 evaluation reward: 4.44
 episode: 1936 score: 6.0 memory length: 100000 epsilon: 0.4053228400102555
 steps: 392 lr: 6.400000000000001e-06 evaluation reward: 4.44
 episode: 1937 score: 3.0 memory length: 100000 epsilon: 0.4048337800102524
 steps: 247 lr: 6.400000000000001e-06 evaluation reward: 4.36
 episode: 1938 score: 2.0 memory length: 100000 epsilon: 0.4044377800102499
 steps: 200 lr: 6.400000000000001e-06 evaluation reward: 4.35
 episode: 1939 score: 2.0 memory length: 100000 epsilon: 0.4040477200102474
 steps: 197 lr: 6.400000000000001e-06 evaluation reward: 4.34
 episode: 1940 score: 8.0 memory length: 100000 epsilon: 0.4031468200102417
 steps: 455 lr: 6.400000000000001e-06 evaluation reward: 4.39
 episode: 1941 score: 5.0 memory length: 100000 epsilon: 0.4026082600102383
 steps: 272 lr: 6.400000000000001e-06 evaluation reward: 4.4
 episode: 1942 score: 4.0 memory length: 100000 epsilon:

0.40205980001023484 steps: 277 lr: 6.4000000000000001e-06 evaluation
 reward: 4.39
 episode: 1943 score: 6.0 memory length: 100000 epsilon:
 0.40131730001023014 steps: 375 lr: 6.4000000000000001e-06 evaluation
 reward: 4.4
 episode: 1944 score: 6.0 memory length: 100000 epsilon:
 0.40063816001022584 steps: 343 lr: 6.4000000000000001e-06 evaluation
 reward: 4.43
 episode: 1945 score: 3.0 memory length: 100000 epsilon: 0.4002203800102232
 steps: 211 lr: 6.4000000000000001e-06 evaluation reward: 4.43
 episode: 1946 score: 8.0 memory length: 100000 epsilon: 0.3993214600102175
 steps: 454 lr: 6.4000000000000001e-06 evaluation reward: 4.47
 episode: 1947 score: 4.0 memory length: 100000 epsilon: 0.3987353800102138
 steps: 296 lr: 6.4000000000000001e-06 evaluation reward: 4.48
 episode: 1948 score: 3.0 memory length: 100000 epsilon: 0.398289880010211
 steps: 225 lr: 6.4000000000000001e-06 evaluation reward: 4.49
 episode: 1949 score: 6.0 memory length: 100000 epsilon: 0.3975374800102062
 steps: 380 lr: 6.4000000000000001e-06 evaluation reward: 4.45
 episode: 1950 score: 7.0 memory length: 100000 epsilon: 0.3967454800102012
 steps: 400 lr: 6.4000000000000001e-06 evaluation reward: 4.49
 episode: 1951 score: 3.0 memory length: 100000 epsilon:
 0.39632176001019853 steps: 214 lr: 6.4000000000000001e-06 evaluation
 reward: 4.5
 episode: 1952 score: 5.0 memory length: 100000 epsilon: 0.3957198400101947
 steps: 304 lr: 6.4000000000000001e-06 evaluation reward: 4.54
 episode: 1953 score: 5.0 memory length: 100000 epsilon:
 0.39515356001019114 steps: 286 lr: 6.4000000000000001e-06 evaluation
 reward: 4.56
 episode: 1954 score: 3.0 memory length: 100000 epsilon: 0.3946704400101881
 steps: 244 lr: 6.4000000000000001e-06 evaluation reward: 4.52
 episode: 1955 score: 2.0 memory length: 100000 epsilon: 0.3942744400101856
 steps: 200 lr: 6.4000000000000001e-06 evaluation reward: 4.47
 episode: 1956 score: 4.0 memory length: 100000 epsilon: 0.3937240000101821
 steps: 278 lr: 6.4000000000000001e-06 evaluation reward: 4.46
 episode: 1957 score: 4.0 memory length: 100000 epsilon: 0.3931359400101784
 steps: 297 lr: 6.4000000000000001e-06 evaluation reward: 4.47
 episode: 1958 score: 3.0 memory length: 100000 epsilon:
 0.39268846001017554 steps: 226 lr: 6.4000000000000001e-06 evaluation
 reward: 4.45
 episode: 1959 score: 7.0 memory length: 100000 epsilon: 0.3918944800101705
 steps: 401 lr: 6.4000000000000001e-06 evaluation reward: 4.44
 episode: 1960 score: 3.0 memory length: 100000 epsilon:
 0.39147076001016784 steps: 214 lr: 6.4000000000000001e-06 evaluation
 reward: 4.41
 episode: 1961 score: 8.0 memory length: 100000 epsilon: 0.3906074800101624
 steps: 436 lr: 6.4000000000000001e-06 evaluation reward: 4.46
 episode: 1962 score: 6.0 memory length: 100000 epsilon: 0.389912500010158
 steps: 351 lr: 6.4000000000000001e-06 evaluation reward: 4.48

episode: 1963 score: 3.0 memory length: 100000 epsilon:
 0.38946304001015514 steps: 227 lr: 6.4000000000000001e-06 evaluation
 reward: 4.49
 episode: 1964 score: 6.0 memory length: 100000 epsilon: 0.3887957800101509
 steps: 337 lr: 6.4000000000000001e-06 evaluation reward: 4.5
 episode: 1965 score: 6.0 memory length: 100000 epsilon:
 0.38806120001014627 steps: 371 lr: 6.4000000000000001e-06 evaluation
 reward: 4.51
 episode: 1966 score: 3.0 memory length: 100000 epsilon: 0.3876077800101434
 steps: 229 lr: 6.4000000000000001e-06 evaluation reward: 4.53
 episode: 1967 score: 5.0 memory length: 100000 epsilon: 0.3869286400101391
 steps: 343 lr: 6.4000000000000001e-06 evaluation reward: 4.55
 episode: 1968 score: 6.0 memory length: 100000 epsilon: 0.3861881200101344
 steps: 374 lr: 6.4000000000000001e-06 evaluation reward: 4.58
 episode: 1969 score: 3.0 memory length: 100000 epsilon: 0.3857406400101316
 steps: 226 lr: 6.4000000000000001e-06 evaluation reward: 4.57
 episode: 1970 score: 5.0 memory length: 100000 epsilon:
 0.38505556001012725 steps: 346 lr: 6.4000000000000001e-06 evaluation
 reward: 4.55
 episode: 1971 score: 5.0 memory length: 100000 epsilon:
 0.38437642001012295 steps: 343 lr: 6.4000000000000001e-06 evaluation
 reward: 4.56
 episode: 1972 score: 4.0 memory length: 100000 epsilon: 0.3838319200101195
 steps: 275 lr: 6.4000000000000001e-06 evaluation reward: 4.54
 episode: 1973 score: 2.0 memory length: 100000 epsilon: 0.383435920010117
 steps: 200 lr: 6.4000000000000001e-06 evaluation reward: 4.51
 episode: 1974 score: 5.0 memory length: 100000 epsilon: 0.3827567800101127
 steps: 343 lr: 6.4000000000000001e-06 evaluation reward: 4.51
 episode: 1975 score: 4.0 memory length: 100000 epsilon: 0.3822162400101093
 steps: 273 lr: 6.4000000000000001e-06 evaluation reward: 4.5
 episode: 1976 score: 5.0 memory length: 100000 epsilon: 0.3815707600101052
 steps: 326 lr: 6.4000000000000001e-06 evaluation reward: 4.47
 episode: 1977 score: 6.0 memory length: 100000 epsilon: 0.380903500010101
 steps: 337 lr: 6.4000000000000001e-06 evaluation reward: 4.47
 episode: 1978 score: 4.0 memory length: 100000 epsilon: 0.3803550400100975
 steps: 277 lr: 6.4000000000000001e-06 evaluation reward: 4.47
 episode: 1979 score: 7.0 memory length: 100000 epsilon: 0.3796145200100928
 steps: 374 lr: 6.4000000000000001e-06 evaluation reward: 4.51
 episode: 1980 score: 4.0 memory length: 100000 epsilon:
 0.37903240001008914 steps: 294 lr: 6.4000000000000001e-06 evaluation
 reward: 4.48
 episode: 1981 score: 6.0 memory length: 100000 epsilon:
 0.37832554001008467 steps: 357 lr: 6.4000000000000001e-06 evaluation
 reward: 4.51
 episode: 1982 score: 2.0 memory length: 100000 epsilon: 0.3779651800100824
 steps: 182 lr: 6.4000000000000001e-06 evaluation reward: 4.48
 episode: 1983 score: 2.0 memory length: 100000 epsilon:
 0.37753354001007966 steps: 218 lr: 6.4000000000000001e-06 evaluation


```

reward: 4.46
episode: 1984   score: 3.0   memory length: 100000   epsilon: 0.377111800010077
steps: 213     lr: 6.4000000000000001e-06   evaluation reward: 4.46
episode: 1985   score: 6.0   memory length: 100000   epsilon: 0.3763673200100723
steps: 376     lr: 6.4000000000000001e-06   evaluation reward: 4.49
episode: 1986   score: 2.0   memory length: 100000   epsilon: 0.3759733000100698
steps: 199     lr: 6.4000000000000001e-06   evaluation reward: 4.47
episode: 1987   score: 6.0   memory length: 100000   epsilon: 0.3752327800100651
steps: 374     lr: 6.4000000000000001e-06   evaluation reward: 4.49
episode: 1988   score: 6.0   memory length: 100000   epsilon:
0.37454176001006073   steps: 349   lr: 6.4000000000000001e-06   evaluation
reward: 4.53
episode: 1989   score: 4.0   memory length: 100000   epsilon: 0.3739655800100571
steps: 291     lr: 6.4000000000000001e-06   evaluation reward: 4.51
episode: 1990   score: 7.0   memory length: 100000   epsilon:
0.37323100001005244   steps: 371   lr: 6.4000000000000001e-06   evaluation
reward: 4.53
episode: 1991   score: 1.0   memory length: 100000   epsilon: 0.3728944000100503
steps: 170     lr: 6.4000000000000001e-06   evaluation reward: 4.49
episode: 1992   score: 6.0   memory length: 100000   epsilon: 0.3721519000100456
steps: 375     lr: 6.4000000000000001e-06   evaluation reward: 4.53
episode: 1993   score: 4.0   memory length: 100000   epsilon:
0.37160740001004217   steps: 275   lr: 6.4000000000000001e-06   evaluation
reward: 4.53
episode: 1994   score: 5.0   memory length: 100000   epsilon: 0.3710114200100384
steps: 301     lr: 6.4000000000000001e-06   evaluation reward: 4.54
episode: 1995   score: 5.0   memory length: 100000   epsilon: 0.3703619800100343
steps: 328     lr: 6.4000000000000001e-06   evaluation reward: 4.53
episode: 1996   score: 5.0   memory length: 100000   epsilon:
0.36972244001003024   steps: 323   lr: 6.4000000000000001e-06   evaluation
reward: 4.54
episode: 1997   score: 5.0   memory length: 100000   epsilon:
0.36907300001002613   steps: 328   lr: 6.4000000000000001e-06   evaluation
reward: 4.55
episode: 1998   score: 5.0   memory length: 100000   epsilon: 0.3684710800100223
steps: 304     lr: 6.4000000000000001e-06   evaluation reward: 4.57
episode: 1999   score: 3.0   memory length: 100000   epsilon: 0.3680117200100194
steps: 232     lr: 6.4000000000000001e-06   evaluation reward: 4.55
episode: 2000   score: 4.0   memory length: 100000   epsilon: 0.3674216800100157
steps: 298     lr: 6.4000000000000001e-06   evaluation reward: 4.56
episode: 2001   score: 4.0   memory length: 100000   epsilon:
0.36687718001001224   steps: 275   lr: 6.4000000000000001e-06   evaluation
reward: 4.58
episode: 2002   score: 5.0   memory length: 100000   epsilon: 0.3662713000100084
steps: 306     lr: 6.4000000000000001e-06   evaluation reward: 4.6
episode: 2003   score: 8.0   memory length: 100000   epsilon: 0.3653704000100027
steps: 455     lr: 6.4000000000000001e-06   evaluation reward: 4.65
episode: 2004   score: 4.0   memory length: 100000   epsilon:

```

0.36482194000999923 steps: 277 lr: 6.4000000000000001e-06 evaluation
 reward: 4.66
 episode: 2005 score: 7.0 memory length: 100000 epsilon:
 0.36398440000999394 steps: 423 lr: 6.4000000000000001e-06 evaluation
 reward: 4.68
 episode: 2006 score: 5.0 memory length: 100000 epsilon:
 0.36342010000999037 steps: 285 lr: 6.4000000000000001e-06 evaluation
 reward: 4.7
 episode: 2007 score: 5.0 memory length: 100000 epsilon: 0.3627746200099863
 steps: 326 lr: 6.4000000000000001e-06 evaluation reward: 4.71
 episode: 2008 score: 7.0 memory length: 100000 epsilon: 0.3619568800099811
 steps: 413 lr: 6.4000000000000001e-06 evaluation reward: 4.75
 episode: 2009 score: 3.0 memory length: 100000 epsilon: 0.3615430600099785
 steps: 209 lr: 6.4000000000000001e-06 evaluation reward: 4.71
 episode: 2010 score: 7.0 memory length: 100000 epsilon:
 0.36074908000997347 steps: 401 lr: 6.4000000000000001e-06 evaluation
 reward: 4.76
 episode: 2011 score: 5.0 memory length: 100000 epsilon: 0.3600620200099691
 steps: 347 lr: 6.4000000000000001e-06 evaluation reward: 4.72
 episode: 2012 score: 6.0 memory length: 100000 epsilon: 0.3593611000099647
 steps: 354 lr: 6.4000000000000001e-06 evaluation reward: 4.73
 episode: 2013 score: 6.0 memory length: 100000 epsilon: 0.3586641400099603
 steps: 352 lr: 6.4000000000000001e-06 evaluation reward: 4.73
 episode: 2014 score: 6.0 memory length: 100000 epsilon: 0.3579236200099556
 steps: 374 lr: 6.4000000000000001e-06 evaluation reward: 4.76
 episode: 2015 score: 6.0 memory length: 100000 epsilon:
 0.35718904000995094 steps: 371 lr: 6.4000000000000001e-06 evaluation
 reward: 4.78
 episode: 2016 score: 5.0 memory length: 100000 epsilon: 0.3565495000099469
 steps: 323 lr: 6.4000000000000001e-06 evaluation reward: 4.79
 episode: 2017 score: 4.0 memory length: 100000 epsilon: 0.3560446000099437
 steps: 255 lr: 6.4000000000000001e-06 evaluation reward: 4.78
 episode: 2018 score: 4.0 memory length: 100000 epsilon:
 0.35549812000994024 steps: 276 lr: 6.4000000000000001e-06 evaluation
 reward: 4.78
 episode: 2019 score: 7.0 memory length: 100000 epsilon: 0.3547972000099358
 steps: 354 lr: 6.4000000000000001e-06 evaluation reward: 4.76
 episode: 2020 score: 7.0 memory length: 100000 epsilon: 0.3540032200099308
 steps: 401 lr: 6.4000000000000001e-06 evaluation reward: 4.78
 episode: 2021 score: 8.0 memory length: 100000 epsilon: 0.3531221200099252
 steps: 445 lr: 6.4000000000000001e-06 evaluation reward: 4.78
 episode: 2022 score: 2.0 memory length: 100000 epsilon:
 0.35273008000992273 steps: 198 lr: 6.4000000000000001e-06 evaluation
 reward: 4.76
 episode: 2023 score: 6.0 memory length: 100000 epsilon: 0.3519479800099178
 steps: 395 lr: 6.4000000000000001e-06 evaluation reward: 4.77
 episode: 2024 score: 3.0 memory length: 100000 epsilon:
 0.35153020000991514 steps: 211 lr: 6.4000000000000001e-06 evaluation

```

reward: 4.76
episode: 2025   score: 8.0   memory length: 100000   epsilon: 0.3506550400099096
steps: 442     lr: 6.4000000000000001e-06   evaluation reward: 4.82
episode: 2026   score: 7.0   memory length: 100000   epsilon: 0.3498630400099046
steps: 400     lr: 6.4000000000000001e-06   evaluation reward: 4.83
episode: 2027   score: 5.0   memory length: 100000   epsilon: 0.349292800009901
steps: 288     lr: 6.4000000000000001e-06   evaluation reward: 4.85
episode: 2028   score: 2.0   memory length: 100000   epsilon:
0.34889482000989847   steps: 201   lr: 6.4000000000000001e-06   evaluation
reward: 4.8
episode: 2029   score: 5.0   memory length: 100000   epsilon: 0.3482493400098944
steps: 326     lr: 6.4000000000000001e-06   evaluation reward: 4.8
episode: 2030   score: 6.0   memory length: 100000   epsilon:
0.34750486000988967   steps: 376   lr: 6.4000000000000001e-06   evaluation
reward: 4.82
episode: 2031   score: 6.0   memory length: 100000   epsilon: 0.346766320009885
steps: 373     lr: 6.4000000000000001e-06   evaluation reward: 4.85
episode: 2032   score: 4.0   memory length: 100000   epsilon:
0.34625152000988174   steps: 260   lr: 6.4000000000000001e-06   evaluation
reward: 4.88
episode: 2033   score: 5.0   memory length: 100000   epsilon:
0.34560802000987767   steps: 325   lr: 6.4000000000000001e-06   evaluation
reward: 4.81
episode: 2034   score: 8.0   memory length: 100000   epsilon:
0.34478236000987245   steps: 417   lr: 6.4000000000000001e-06   evaluation
reward: 4.85
episode: 2035   score: 5.0   memory length: 100000   epsilon: 0.3441725200098686
steps: 308     lr: 6.4000000000000001e-06   evaluation reward: 4.87
episode: 2036   score: 6.0   memory length: 100000   epsilon: 0.3433983400098637
steps: 391     lr: 6.4000000000000001e-06   evaluation reward: 4.87
episode: 2037   score: 3.0   memory length: 100000   epsilon: 0.342976600009861
steps: 213     lr: 6.4000000000000001e-06   evaluation reward: 4.87
episode: 2038   score: 2.0   memory length: 100000   epsilon:
0.34258456000985854   steps: 198   lr: 6.4000000000000001e-06   evaluation
reward: 4.87
episode: 2039   score: 6.0   memory length: 100000   epsilon: 0.3418163200098537
steps: 388     lr: 6.4000000000000001e-06   evaluation reward: 4.91
episode: 2040   score: 5.0   memory length: 100000   epsilon:
0.34117678000984963   steps: 323   lr: 6.4000000000000001e-06   evaluation
reward: 4.88
episode: 2041   score: 3.0   memory length: 100000   epsilon: 0.3407293000098468
steps: 226     lr: 6.4000000000000001e-06   evaluation reward: 4.86
episode: 2042   score: 9.0   memory length: 100000   epsilon: 0.3397492000098406
steps: 495     lr: 6.4000000000000001e-06   evaluation reward: 4.91
episode: 2043   score: 4.0   memory length: 100000   epsilon: 0.3391670800098369
steps: 294     lr: 6.4000000000000001e-06   evaluation reward: 4.89
episode: 2044   score: 5.0   memory length: 100000   epsilon: 0.3385275400098329
steps: 323     lr: 6.4000000000000001e-06   evaluation reward: 4.88

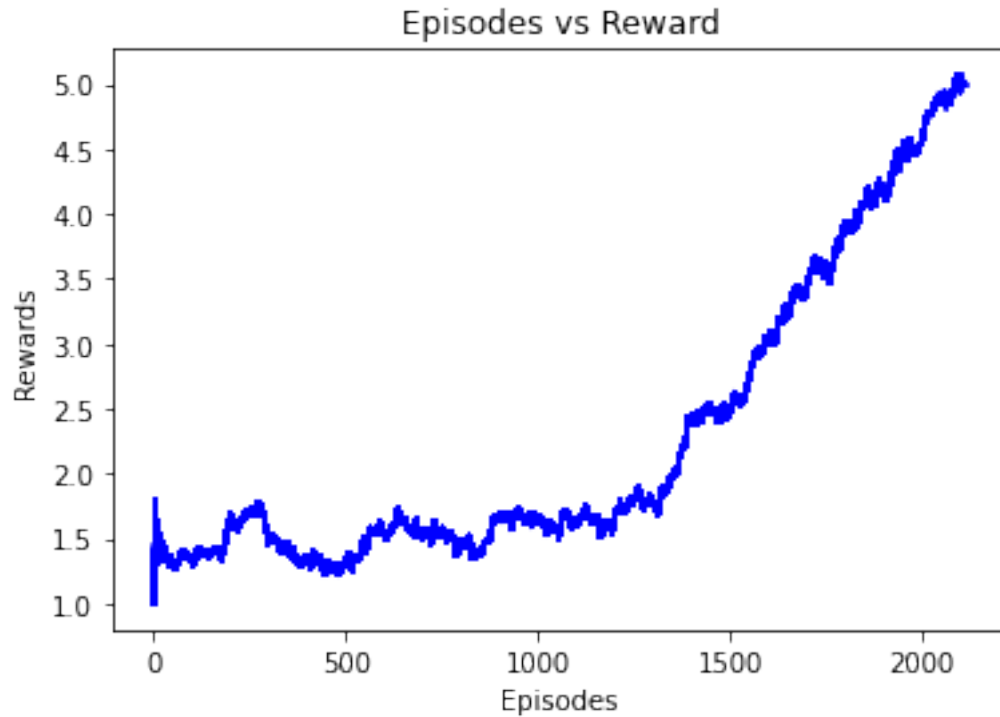
```

episode: 2045 score: 8.0 memory length: 100000 epsilon: 0.3376603000098274
 steps: 438 lr: 6.4000000000000001e-06 evaluation reward: 4.93
 episode: 2046 score: 3.0 memory length: 100000 epsilon:
 0.33721282000982455 steps: 226 lr: 6.4000000000000001e-06 evaluation
 reward: 4.88
 episode: 2047 score: 5.0 memory length: 100000 epsilon:
 0.33667228000982113 steps: 273 lr: 6.4000000000000001e-06 evaluation
 reward: 4.89
 episode: 2048 score: 4.0 memory length: 100000 epsilon: 0.3361258000098177
 steps: 276 lr: 6.4000000000000001e-06 evaluation reward: 4.9
 episode: 2049 score: 6.0 memory length: 100000 epsilon: 0.3354625000098135
 steps: 335 lr: 6.4000000000000001e-06 evaluation reward: 4.9
 episode: 2050 score: 6.0 memory length: 100000 epsilon: 0.334753660009809
 steps: 358 lr: 6.4000000000000001e-06 evaluation reward: 4.89
 episode: 2051 score: 2.0 memory length: 100000 epsilon: 0.3343576600098065
 steps: 200 lr: 6.4000000000000001e-06 evaluation reward: 4.88
 episode: 2052 score: 3.0 memory length: 100000 epsilon: 0.3339022600098036
 steps: 230 lr: 6.4000000000000001e-06 evaluation reward: 4.86
 episode: 2053 score: 6.0 memory length: 100000 epsilon:
 0.33316768000979896 steps: 371 lr: 6.4000000000000001e-06 evaluation
 reward: 4.87
 episode: 2054 score: 2.0 memory length: 100000 epsilon:
 0.33277168000979646 steps: 200 lr: 6.4000000000000001e-06 evaluation
 reward: 4.86
 episode: 2055 score: 6.0 memory length: 100000 epsilon:
 0.33207274000979203 steps: 353 lr: 6.4000000000000001e-06 evaluation
 reward: 4.9
 episode: 2056 score: 7.0 memory length: 100000 epsilon: 0.331276780009787
 steps: 402 lr: 6.4000000000000001e-06 evaluation reward: 4.93
 episode: 2057 score: 5.0 memory length: 100000 epsilon: 0.3306768400097832
 steps: 303 lr: 6.4000000000000001e-06 evaluation reward: 4.94
 episode: 2058 score: 3.0 memory length: 100000 epsilon:
 0.33025906000978056 steps: 211 lr: 6.4000000000000001e-06 evaluation
 reward: 4.94
 episode: 2059 score: 3.0 memory length: 100000 epsilon:
 0.32981356000977774 steps: 225 lr: 6.4000000000000001e-06 evaluation
 reward: 4.9
 episode: 2060 score: 1.0 memory length: 100000 epsilon:
 0.32951656000977586 steps: 150 lr: 6.4000000000000001e-06 evaluation
 reward: 4.88
 episode: 2061 score: 4.0 memory length: 100000 epsilon: 0.3289681000097724
 steps: 277 lr: 6.4000000000000001e-06 evaluation reward: 4.84
 episode: 2062 score: 4.0 memory length: 100000 epsilon:
 0.32834638000976846 steps: 314 lr: 6.4000000000000001e-06 evaluation
 reward: 4.82
 episode: 2063 score: 7.0 memory length: 100000 epsilon:
 0.32757022000976355 steps: 392 lr: 6.4000000000000001e-06 evaluation
 reward: 4.86

episode: 2064 score: 5.0 memory length: 100000 epsilon: 0.3269623600097597
 steps: 307 lr: 6.400000000000001e-06 evaluation reward: 4.85
 episode: 2065 score: 10.0 memory length: 100000 epsilon:
 0.32591296000975306 steps: 530 lr: 6.400000000000001e-06 evaluation
 reward: 4.89
 episode: 2066 score: 4.0 memory length: 100000 epsilon:
 0.32532886000974937 steps: 295 lr: 6.400000000000001e-06 evaluation
 reward: 4.9
 episode: 2067 score: 2.0 memory length: 100000 epsilon: 0.3249368200097469
 steps: 198 lr: 6.400000000000001e-06 evaluation reward: 4.87
 episode: 2068 score: 4.0 memory length: 100000 epsilon:
 0.32439628000974347 steps: 273 lr: 6.400000000000001e-06 evaluation
 reward: 4.85
 episode: 2069 score: 5.0 memory length: 100000 epsilon:
 0.32379040000973963 steps: 306 lr: 6.400000000000001e-06 evaluation
 reward: 4.87
 episode: 2070 score: 4.0 memory length: 100000 epsilon: 0.3232459000097362
 steps: 275 lr: 6.400000000000001e-06 evaluation reward: 4.86
 episode: 2071 score: 4.0 memory length: 100000 epsilon:
 0.32273506000973295 steps: 258 lr: 6.400000000000001e-06 evaluation
 reward: 4.85
 episode: 2072 score: 4.0 memory length: 100000 epsilon:
 0.32222818000972975 steps: 256 lr: 6.400000000000001e-06 evaluation
 reward: 4.85
 episode: 2073 score: 11.0 memory length: 100000 epsilon:
 0.3212381800097235 steps: 500 lr: 6.400000000000001e-06 evaluation
 reward: 4.94
 episode: 2074 score: 5.0 memory length: 100000 epsilon:
 0.32063032000971964 steps: 307 lr: 6.400000000000001e-06 evaluation
 reward: 4.94
 episode: 2075 score: 3.0 memory length: 100000 epsilon: 0.320214520009717
 steps: 210 lr: 6.400000000000001e-06 evaluation reward: 4.93
 episode: 2076 score: 4.0 memory length: 100000 epsilon: 0.3196284400097133
 steps: 296 lr: 6.400000000000001e-06 evaluation reward: 4.92
 episode: 2077 score: 4.0 memory length: 100000 epsilon: 0.3190443400097096
 steps: 295 lr: 6.400000000000001e-06 evaluation reward: 4.9
 episode: 2078 score: 9.0 memory length: 100000 epsilon:
 0.31808404000970353 steps: 485 lr: 6.400000000000001e-06 evaluation
 reward: 4.95
 episode: 2079 score: 5.0 memory length: 100000 epsilon: 0.3173969800096992
 steps: 347 lr: 6.400000000000001e-06 evaluation reward: 4.93
 episode: 2080 score: 5.0 memory length: 100000 epsilon:
 0.31678912000969534 steps: 307 lr: 6.400000000000001e-06 evaluation
 reward: 4.94
 episode: 2081 score: 6.0 memory length: 100000 epsilon: 0.3160921600096909
 steps: 352 lr: 6.400000000000001e-06 evaluation reward: 4.94
 episode: 2082 score: 7.0 memory length: 100000 epsilon: 0.3152962000096859
 steps: 402 lr: 6.400000000000001e-06 evaluation reward: 4.99

episode: 2083 score: 3.0 memory length: 100000 epsilon:
 0.31487842000968325 steps: 211 lr: 6.4000000000000001e-06 evaluation
 reward: 5.0
 episode: 2084 score: 6.0 memory length: 100000 epsilon: 0.314205220009679
 steps: 340 lr: 6.4000000000000001e-06 evaluation reward: 5.03
 episode: 2085 score: 3.0 memory length: 100000 epsilon: 0.3137201200096759
 steps: 245 lr: 6.4000000000000001e-06 evaluation reward: 5.0
 episode: 2086 score: 3.0 memory length: 100000 epsilon: 0.313262740009673
 steps: 231 lr: 6.4000000000000001e-06 evaluation reward: 5.01
 episode: 2087 score: 9.0 memory length: 100000 epsilon:
 0.31222324000966645 steps: 525 lr: 6.4000000000000001e-06 evaluation
 reward: 5.04
 episode: 2088 score: 9.0 memory length: 100000 epsilon: 0.3114708400096617
 steps: 380 lr: 6.4000000000000001e-06 evaluation reward: 5.07
 episode: 2089 score: 2.0 memory length: 100000 epsilon: 0.3110788000096592
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 5.05
 episode: 2090 score: 3.0 memory length: 100000 epsilon:
 0.31062934000965636 steps: 227 lr: 6.4000000000000001e-06 evaluation
 reward: 5.01
 episode: 2091 score: 6.0 memory length: 100000 epsilon:
 0.30988684000965167 steps: 375 lr: 6.4000000000000001e-06 evaluation
 reward: 5.06
 episode: 2092 score: 3.0 memory length: 100000 epsilon: 0.309465100009649
 steps: 213 lr: 6.4000000000000001e-06 evaluation reward: 5.03
 episode: 2093 score: 6.0 memory length: 100000 epsilon: 0.3087206200096443
 steps: 376 lr: 6.4000000000000001e-06 evaluation reward: 5.05
 episode: 2094 score: 3.0 memory length: 100000 epsilon:
 0.30830482000964166 steps: 210 lr: 6.4000000000000001e-06 evaluation
 reward: 5.03
 episode: 2095 score: 2.0 memory length: 100000 epsilon: 0.3079444600096394
 steps: 182 lr: 6.4000000000000001e-06 evaluation reward: 5.0
 episode: 2096 score: 3.0 memory length: 100000 epsilon:
 0.30749500000963653 steps: 227 lr: 6.4000000000000001e-06 evaluation
 reward: 4.98
 episode: 2097 score: 2.0 memory length: 100000 epsilon: 0.307099000009634
 steps: 200 lr: 6.4000000000000001e-06 evaluation reward: 4.95
 episode: 2098 score: 4.0 memory length: 100000 epsilon: 0.3065129200096303
 steps: 296 lr: 6.4000000000000001e-06 evaluation reward: 4.94
 episode: 2099 score: 7.0 memory length: 100000 epsilon: 0.3057506200096255
 steps: 385 lr: 6.4000000000000001e-06 evaluation reward: 4.98
 episode: 2100 score: 10.0 memory length: 100000 epsilon:
 0.3047705200096193 steps: 495 lr: 6.4000000000000001e-06 evaluation
 reward: 5.04
 episode: 2101 score: 7.0 memory length: 100000 epsilon:
 0.30392704000961396 steps: 426 lr: 6.4000000000000001e-06 evaluation
 reward: 5.07
 episode: 2102 score: 3.0 memory length: 100000 epsilon: 0.3035112400096113
 steps: 210 lr: 6.4000000000000001e-06 evaluation reward: 5.05

episode: 2103 score: 4.0 memory length: 100000 epsilon:
 0.30296476000960787 steps: 276 lr: 6.4000000000000001e-06 evaluation
 reward: 5.01
 episode: 2104 score: 5.0 memory length: 100000 epsilon:
 0.30239056000960424 steps: 290 lr: 6.4000000000000001e-06 evaluation
 reward: 5.02
 episode: 2105 score: 7.0 memory length: 100000 epsilon:
 0.30155302000959894 steps: 423 lr: 6.4000000000000001e-06 evaluation
 reward: 5.02
 episode: 2106 score: 4.0 memory length: 100000 epsilon: 0.3010421800095957
 steps: 258 lr: 6.4000000000000001e-06 evaluation reward: 5.01
 episode: 2107 score: 6.0 memory length: 100000 epsilon:
 0.30030562000959105 steps: 372 lr: 6.4000000000000001e-06 evaluation
 reward: 5.02
 episode: 2108 score: 6.0 memory length: 100000 epsilon: 0.2995987600095866
 steps: 357 lr: 6.4000000000000001e-06 evaluation reward: 5.01
 episode: 2109 score: 3.0 memory length: 100000 epsilon: 0.2991116800095835
 steps: 246 lr: 6.4000000000000001e-06 evaluation reward: 5.01
 episode: 2110 score: 6.0 memory length: 100000 epsilon: 0.2984365000095792
 steps: 341 lr: 6.4000000000000001e-06 evaluation reward: 5.0
 episode: 2111 score: 7.0 memory length: 100000 epsilon:
 0.29766628000957435 steps: 389 lr: 6.4000000000000001e-06 evaluation
 reward: 5.02
 episode: 2112 score: 5.0 memory length: 100000 epsilon:
 0.29710000000957076 steps: 286 lr: 6.4000000000000001e-06 evaluation
 reward: 5.01
 episode: 2113 score: 4.0 memory length: 100000 epsilon: 0.2965535200095673
 steps: 276 lr: 6.4000000000000001e-06 evaluation reward: 4.99
 episode: 2114 score: 6.0 memory length: 100000 epsilon: 0.2959199200095633
 steps: 320 lr: 6.4000000000000001e-06 evaluation reward: 4.99
 episode: 2115 score: 6.0 memory length: 100000 epsilon:
 0.29517148000955856 steps: 378 lr: 6.4000000000000001e-06 evaluation
 reward: 4.99



```
[ ]: ### double DQN

rewards, episodes = [], []
best_eval_reward = 0
start=time.time()
for e in range(EPISODES):
    done = False
    score = 0

    history = np.zeros([5, 84, 84], dtype=np.uint8)
    step = 0
    d = False
    state = env.reset()
    next_state = state
    life = number_lives

    get_init_state(history, state)

    while not done:
        step += 1
        frame += 1

        # Perform a fire action if ball is no longer on screen to continue onto
        →next life
```



```

if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
    action = 0
else:
    action = agent.get_action(np.float32(history[None,:4, :, :]) / 255.)
state = next_state
next_state, reward, done, info = env.step(action + 1)

frame_next_state = get_frame(next_state)
history[4, :, :] = frame_next_state
terminal_state = check_live(life, info['ale.lives'])

life = info['ale.lives']
r = np.clip(reward, -1, 1)
r = reward

# Store the transition in memory
agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
# Start training after random sample generation
if (frame >= train_frame):
    agent.train_policy_net(frame)
    # Update the target network only for Double DQN only
    if double_dqn and (frame % update_target_network_frequency) == 0:
        agent.update_target_net()
    score += reward
    history[:4, :, :] = history[1:, :, :]

if done:
    evaluation_reward.append(score)
    rewards.append(np.mean(evaluation_reward))
    episodes.append(e)
    pylab.plot(episodes, rewards, 'b')
    pylab.xlabel('Episodes')
    pylab.ylabel('Rewards')
    pylab.title('Episodes vs Reward')
    pylab.savefig("./save_graph/breakout_dqn.png") # save graph for
→ training visualization

    # every episode, plot the play time
    print("episode:", e, " score:", score, " memory length:",
          len(agent.memory), " epsilon:", agent.epsilon, " steps:",
→ step,
          " lr:", agent.optimizer.param_groups[0]['lr'], "
→ evaluation reward:", np.mean(evaluation_reward))

    # if the mean of scores of last 100 episode is bigger than 5 save
→ model

    ### Change this save condition to whatever you prefer ###

```

```

        if np.mean(evaluation_reward) > 5 and np.mean(evaluation_reward) >
→best_eval_reward:
            torch.save(agent.policy_net, "./save_model/breakout_dqn.pth")
            best_eval_reward = np.mean(evaluation_reward)

print(time.time()-start)

```

```

episode: 0    score: 0.0    memory length: 122    epsilon: 1.0    steps: 122    lr:
0.0001    evaluation reward: 0.0
episode: 1    score: 5.0    memory length: 465    epsilon: 1.0    steps: 343    lr:
0.0001    evaluation reward: 2.5
episode: 2    score: 1.0    memory length: 615    epsilon: 1.0    steps: 150    lr:
0.0001    evaluation reward: 2.0
episode: 3    score: 0.0    memory length: 738    epsilon: 1.0    steps: 123    lr:
0.0001    evaluation reward: 1.5
episode: 4    score: 4.0    memory length: 1015    epsilon: 1.0    steps: 277
lr: 0.0001    evaluation reward: 2.0
episode: 5    score: 2.0    memory length: 1212    epsilon: 1.0    steps: 197
lr: 0.0001    evaluation reward: 2.0
episode: 6    score: 0.0    memory length: 1335    epsilon: 1.0    steps: 123
lr: 0.0001    evaluation reward: 1.7142857142857142
episode: 7    score: 0.0    memory length: 1457    epsilon: 1.0    steps: 122
lr: 0.0001    evaluation reward: 1.5
episode: 8    score: 1.0    memory length: 1626    epsilon: 1.0    steps: 169
lr: 0.0001    evaluation reward: 1.4444444444444444
episode: 9    score: 0.0    memory length: 1749    epsilon: 1.0    steps: 123
lr: 0.0001    evaluation reward: 1.3
episode: 10   score: 3.0    memory length: 1978    epsilon: 1.0    steps: 229
lr: 0.0001    evaluation reward: 1.4545454545454546
episode: 11   score: 0.0    memory length: 2100    epsilon: 1.0    steps: 122
lr: 0.0001    evaluation reward: 1.3333333333333333
episode: 12   score: 0.0    memory length: 2223    epsilon: 1.0    steps: 123
lr: 0.0001    evaluation reward: 1.2307692307692308
episode: 13   score: 3.0    memory length: 2471    epsilon: 1.0    steps: 248
lr: 0.0001    evaluation reward: 1.3571428571428572
episode: 14   score: 1.0    memory length: 2622    epsilon: 1.0    steps: 151
lr: 0.0001    evaluation reward: 1.3333333333333333
episode: 15   score: 3.0    memory length: 2869    epsilon: 1.0    steps: 247
lr: 0.0001    evaluation reward: 1.4375
episode: 16   score: 1.0    memory length: 3019    epsilon: 1.0    steps: 150
lr: 0.0001    evaluation reward: 1.411764705882353
episode: 17   score: 1.0    memory length: 3170    epsilon: 1.0    steps: 151
lr: 0.0001    evaluation reward: 1.3888888888888888
episode: 18   score: 1.0    memory length: 3321    epsilon: 1.0    steps: 151
lr: 0.0001    evaluation reward: 1.368421052631579
episode: 19   score: 2.0    memory length: 3539    epsilon: 1.0    steps: 218

```

lr: 0.0001	evaluation reward: 1.4		
episode: 20	score: 3.0	memory length: 3790	epsilon: 1.0 steps: 251
lr: 0.0001	evaluation reward: 1.4761904761904763		
episode: 21	score: 0.0	memory length: 3913	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4090909090909092		
episode: 22	score: 1.0	memory length: 4063	epsilon: 1.0 steps: 150
lr: 0.0001	evaluation reward: 1.391304347826087		
episode: 23	score: 4.0	memory length: 4361	epsilon: 1.0 steps: 298
lr: 0.0001	evaluation reward: 1.5		
episode: 24	score: 1.0	memory length: 4512	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.48		
episode: 25	score: 0.0	memory length: 4634	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.4230769230769231		
episode: 26	score: 0.0	memory length: 4757	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.3703703703703705		
episode: 27	score: 3.0	memory length: 5003	epsilon: 1.0 steps: 246
lr: 0.0001	evaluation reward: 1.4285714285714286		
episode: 28	score: 3.0	memory length: 5212	epsilon: 1.0 steps: 209
lr: 0.0001	evaluation reward: 1.4827586206896552		
episode: 29	score: 0.0	memory length: 5335	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4333333333333333		
episode: 30	score: 0.0	memory length: 5458	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.3870967741935485		
episode: 31	score: 3.0	memory length: 5683	epsilon: 1.0 steps: 225
lr: 0.0001	evaluation reward: 1.4375		
episode: 32	score: 2.0	memory length: 5901	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.4545454545454546		
episode: 33	score: 6.0	memory length: 6220	epsilon: 1.0 steps: 319
lr: 0.0001	evaluation reward: 1.588235294117647		
episode: 34	score: 1.0	memory length: 6388	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.5714285714285714		
episode: 35	score: 5.0	memory length: 6733	epsilon: 1.0 steps: 345
lr: 0.0001	evaluation reward: 1.6666666666666667		
episode: 36	score: 2.0	memory length: 6951	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.6756756756756757		
episode: 37	score: 2.0	memory length: 7149	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.6842105263157894		
episode: 38	score: 1.0	memory length: 7300	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.6666666666666667		
episode: 39	score: 0.0	memory length: 7423	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.625		
episode: 40	score: 1.0	memory length: 7574	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.6097560975609757		
episode: 41	score: 0.0	memory length: 7697	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5714285714285714		
episode: 42	score: 0.0	memory length: 7820	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5348837209302326		
episode: 43	score: 2.0	memory length: 8020	epsilon: 1.0 steps: 200

lr: 0.0001	evaluation reward: 1.5454545454545454			
episode: 44	score: 0.0	memory length: 8143	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5111111111111111			
episode: 45	score: 2.0	memory length: 8340	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.5217391304347827			
episode: 46	score: 1.0	memory length: 8509	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5106382978723405			
episode: 47	score: 0.0	memory length: 8632	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4791666666666667			
episode: 48	score: 2.0	memory length: 8832	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.489795918367347			
episode: 49	score: 1.0	memory length: 9001	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.48			
episode: 50	score: 0.0	memory length: 9124	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4509803921568627			
episode: 51	score: 1.0	memory length: 9293	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.4423076923076923			
episode: 52	score: 1.0	memory length: 9462	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.4339622641509433			
episode: 53	score: 0.0	memory length: 9585	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4074074074074074			
episode: 54	score: 0.0	memory length: 9707	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3818181818181818			
episode: 55	score: 2.0	memory length: 9928	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.3928571428571428			
episode: 56	score: 2.0	memory length: 10145	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.4035087719298245			
episode: 57	score: 0.0	memory length: 10268	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3793103448275863			
episode: 58	score: 2.0	memory length: 10466	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.3898305084745763			
episode: 59	score: 0.0	memory length: 10589	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3666666666666667			
episode: 60	score: 1.0	memory length: 10759	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.360655737704918			
episode: 61	score: 1.0	memory length: 10910	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3548387096774193			
episode: 62	score: 1.0	memory length: 11081	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.3492063492063493			
episode: 63	score: 3.0	memory length: 11346	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.375			
episode: 64	score: 1.0	memory length: 11516	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.3692307692307693			
episode: 65	score: 2.0	memory length: 11734	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.378787878787879			
episode: 66	score: 2.0	memory length: 11953	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.3880597014925373			
episode: 67	score: 3.0	memory length: 12197	epsilon: 1.0	steps: 244

lr: 0.0001	evaluation reward: 1.411764705882353		
episode: 68	score: 0.0	memory length: 12319	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.391304347826087		
episode: 69	score: 0.0	memory length: 12442	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.3714285714285714		
episode: 70	score: 2.0	memory length: 12640	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.380281690140845		
episode: 71	score: 0.0	memory length: 12762	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.3611111111111112		
episode: 72	score: 3.0	memory length: 13031	epsilon: 1.0 steps: 269
lr: 0.0001	evaluation reward: 1.3835616438356164		
episode: 73	score: 1.0	memory length: 13203	epsilon: 1.0 steps: 172
lr: 0.0001	evaluation reward: 1.3783783783783783		
episode: 74	score: 3.0	memory length: 13448	epsilon: 1.0 steps: 245
lr: 0.0001	evaluation reward: 1.4		
episode: 75	score: 3.0	memory length: 13676	epsilon: 1.0 steps: 228
lr: 0.0001	evaluation reward: 1.4210526315789473		
episode: 76	score: 2.0	memory length: 13893	epsilon: 1.0 steps: 217
lr: 0.0001	evaluation reward: 1.4285714285714286		
episode: 77	score: 0.0	memory length: 14016	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4102564102564104		
episode: 78	score: 3.0	memory length: 14265	epsilon: 1.0 steps: 249
lr: 0.0001	evaluation reward: 1.4303797468354431		
episode: 79	score: 0.0	memory length: 14388	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4125		
episode: 80	score: 1.0	memory length: 14539	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.4074074074074074		
episode: 81	score: 0.0	memory length: 14661	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.3902439024390243		
episode: 82	score: 0.0	memory length: 14783	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.3734939759036144		
episode: 83	score: 1.0	memory length: 14951	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.369047619047619		
episode: 84	score: 1.0	memory length: 15119	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.3647058823529412		
episode: 85	score: 1.0	memory length: 15270	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.3604651162790697		
episode: 86	score: 2.0	memory length: 15469	epsilon: 1.0 steps: 199
lr: 0.0001	evaluation reward: 1.367816091954023		
episode: 87	score: 0.0	memory length: 15592	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.3522727272727273		
episode: 88	score: 2.0	memory length: 15811	epsilon: 1.0 steps: 219
lr: 0.0001	evaluation reward: 1.3595505617977528		
episode: 89	score: 4.0	memory length: 16125	epsilon: 1.0 steps: 314
lr: 0.0001	evaluation reward: 1.3888888888888888		
episode: 90	score: 0.0	memory length: 16248	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.3736263736263736		
episode: 91	score: 2.0	memory length: 16466	epsilon: 1.0 steps: 218

lr: 0.0001	evaluation reward: 1.3804347826086956		
episode: 92	score: 2.0 memory length: 16663 epsilon: 1.0 steps: 197		
lr: 0.0001	evaluation reward: 1.3870967741935485		
episode: 93	score: 2.0 memory length: 16862 epsilon: 1.0 steps: 199		
lr: 0.0001	evaluation reward: 1.3936170212765957		
episode: 94	score: 1.0 memory length: 17031 epsilon: 1.0 steps: 169		
lr: 0.0001	evaluation reward: 1.3894736842105264		
episode: 95	score: 1.0 memory length: 17181 epsilon: 1.0 steps: 150		
lr: 0.0001	evaluation reward: 1.3854166666666667		
episode: 96	score: 1.0 memory length: 17332 epsilon: 1.0 steps: 151		
lr: 0.0001	evaluation reward: 1.3814432989690721		
episode: 97	score: 0.0 memory length: 17454 epsilon: 1.0 steps: 122		
lr: 0.0001	evaluation reward: 1.3673469387755102		
episode: 98	score: 1.0 memory length: 17622 epsilon: 1.0 steps: 168		
lr: 0.0001	evaluation reward: 1.3636363636363635		
episode: 99	score: 3.0 memory length: 17872 epsilon: 1.0 steps: 250		
lr: 0.0001	evaluation reward: 1.38		
episode: 100	score: 3.0 memory length: 18097 epsilon: 1.0 steps: 225		
lr: 0.0001	evaluation reward: 1.41		
episode: 101	score: 10.0 memory length: 18514 epsilon: 1.0 steps: 417		
lr: 0.0001	evaluation reward: 1.46		
episode: 102	score: 4.0 memory length: 18800 epsilon: 1.0 steps: 286		
lr: 0.0001	evaluation reward: 1.49		
episode: 103	score: 0.0 memory length: 18922 epsilon: 1.0 steps: 122		
lr: 0.0001	evaluation reward: 1.49		
episode: 104	score: 3.0 memory length: 19152 epsilon: 1.0 steps: 230		
lr: 0.0001	evaluation reward: 1.48		
episode: 105	score: 0.0 memory length: 19274 epsilon: 1.0 steps: 122		
lr: 0.0001	evaluation reward: 1.46		
episode: 106	score: 2.0 memory length: 19472 epsilon: 1.0 steps: 198		
lr: 0.0001	evaluation reward: 1.48		
episode: 107	score: 0.0 memory length: 19595 epsilon: 1.0 steps: 123		
lr: 0.0001	evaluation reward: 1.48		
episode: 108	score: 1.0 memory length: 19767 epsilon: 1.0 steps: 172		
lr: 0.0001	evaluation reward: 1.48		
episode: 109	score: 1.0 memory length: 19918 epsilon: 1.0 steps: 151		
lr: 0.0001	evaluation reward: 1.49		
episode: 110	score: 2.0 memory length: 20115 epsilon: 1.0 steps: 197		
lr: 0.0001	evaluation reward: 1.48		
episode: 111	score: 3.0 memory length: 20365 epsilon: 1.0 steps: 250		
lr: 0.0001	evaluation reward: 1.51		
episode: 112	score: 0.0 memory length: 20488 epsilon: 1.0 steps: 123		
lr: 0.0001	evaluation reward: 1.51		
episode: 113	score: 0.0 memory length: 20610 epsilon: 1.0 steps: 122		
lr: 0.0001	evaluation reward: 1.48		
episode: 114	score: 2.0 memory length: 20808 epsilon: 1.0 steps: 198		
lr: 0.0001	evaluation reward: 1.49		
episode: 115	score: 1.0 memory length: 20978 epsilon: 1.0 steps: 170		

lr: 0.0001	evaluation reward: 1.47		
episode: 116	score: 2.0 memory length: 21195	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.48		
episode: 117	score: 1.0 memory length: 21364	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.48		
episode: 118	score: 2.0 memory length: 21562	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49		
episode: 119	score: 0.0 memory length: 21685	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 120	score: 0.0 memory length: 21808	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 121	score: 0.0 memory length: 21931	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 122	score: 2.0 memory length: 22130	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.45		
episode: 123	score: 4.0 memory length: 22405	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.45		
episode: 124	score: 0.0 memory length: 22528	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 125	score: 1.0 memory length: 22679	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.45		
episode: 126	score: 2.0 memory length: 22896	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.47		
episode: 127	score: 1.0 memory length: 23066	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.45		
episode: 128	score: 3.0 memory length: 23313	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.45		
episode: 129	score: 2.0 memory length: 23533	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.47		
episode: 130	score: 1.0 memory length: 23684	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48		
episode: 131	score: 2.0 memory length: 23882	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 132	score: 1.0 memory length: 24051	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.46		
episode: 133	score: 0.0 memory length: 24173	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4		
episode: 134	score: 1.0 memory length: 24324	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.4		
episode: 135	score: 0.0 memory length: 24447	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 136	score: 1.0 memory length: 24618	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.34		
episode: 137	score: 0.0 memory length: 24741	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 138	score: 1.0 memory length: 24911	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.32		
episode: 139	score: 0.0 memory length: 25034	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.32		
episode: 140	score: 2.0 memory length: 25233	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.33		
episode: 141	score: 2.0 memory length: 25431	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.35		
episode: 142	score: 0.0 memory length: 25553	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.35		
episode: 143	score: 0.0 memory length: 25675	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.33		
episode: 144	score: 2.0 memory length: 25873	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.35		
episode: 145	score: 0.0 memory length: 25996	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 146	score: 0.0 memory length: 26119	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 147	score: 3.0 memory length: 26366	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.35		
episode: 148	score: 1.0 memory length: 26537	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.34		
episode: 149	score: 0.0 memory length: 26660	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 150	score: 2.0 memory length: 26881	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.35		
episode: 151	score: 2.0 memory length: 27079	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.36		
episode: 152	score: 2.0 memory length: 27278	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.37		
episode: 153	score: 3.0 memory length: 27507	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.4		
episode: 154	score: 2.0 memory length: 27707	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.42		
episode: 155	score: 2.0 memory length: 27925	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.42		
episode: 156	score: 0.0 memory length: 28048	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 157	score: 0.0 memory length: 28170	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4		
episode: 158	score: 0.0 memory length: 28292	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38		
episode: 159	score: 1.0 memory length: 28462	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.39		
episode: 160	score: 0.0 memory length: 28584	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38		
episode: 161	score: 2.0 memory length: 28787	epsilon: 1.0	steps: 203
lr: 0.0001	evaluation reward: 1.39		
episode: 162	score: 0.0 memory length: 28909	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38		
episode: 163	score: 4.0 memory length: 29223	epsilon: 1.0	steps: 314

lr: 0.0001	evaluation reward: 1.39		
episode: 164	score: 4.0 memory length: 29499	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.42		
episode: 165	score: 2.0 memory length: 29699	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.42		
episode: 166	score: 0.0 memory length: 29822	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 167	score: 0.0 memory length: 29944	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37		
episode: 168	score: 2.0 memory length: 30143	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.39		
episode: 169	score: 2.0 memory length: 30341	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41		
episode: 170	score: 0.0 memory length: 30464	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 171	score: 1.0 memory length: 30635	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.4		
episode: 172	score: 2.0 memory length: 30851	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.39		
episode: 173	score: 2.0 memory length: 31051	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.4		
episode: 174	score: 0.0 memory length: 31174	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37		
episode: 175	score: 2.0 memory length: 31391	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.36		
episode: 176	score: 1.0 memory length: 31559	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.35		
episode: 177	score: 2.0 memory length: 31757	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.37		
episode: 178	score: 0.0 memory length: 31880	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.34		
episode: 179	score: 1.0 memory length: 32050	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.35		
episode: 180	score: 3.0 memory length: 32297	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.37		
episode: 181	score: 0.0 memory length: 32420	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37		
episode: 182	score: 0.0 memory length: 32542	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37		
episode: 183	score: 2.0 memory length: 32740	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.38		
episode: 184	score: 2.0 memory length: 32937	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.39		
episode: 185	score: 4.0 memory length: 33230	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.42		
episode: 186	score: 1.0 memory length: 33399	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.41		
episode: 187	score: 6.0 memory length: 33769	epsilon: 1.0	steps: 370

lr: 0.0001	evaluation reward: 1.47		
episode: 188	score: 2.0 memory length: 33967	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 189	score: 0.0 memory length: 34090	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43		
episode: 190	score: 2.0 memory length: 34309	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.45		
episode: 191	score: 2.0 memory length: 34506	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.45		
episode: 192	score: 1.0 memory length: 34678	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.44		
episode: 193	score: 0.0 memory length: 34801	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 194	score: 0.0 memory length: 34923	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41		
episode: 195	score: 1.0 memory length: 35073	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.41		
episode: 196	score: 3.0 memory length: 35322	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.43		
episode: 197	score: 0.0 memory length: 35444	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43		
episode: 198	score: 0.0 memory length: 35567	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 199	score: 0.0 memory length: 35689	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39		
episode: 200	score: 1.0 memory length: 35858	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.37		
episode: 201	score: 1.0 memory length: 36027	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.28		
episode: 202	score: 0.0 memory length: 36149	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.24		
episode: 203	score: 2.0 memory length: 36350	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.26		
episode: 204	score: 2.0 memory length: 36548	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.25		
episode: 205	score: 0.0 memory length: 36671	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.25		
episode: 206	score: 0.0 memory length: 36793	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.23		
episode: 207	score: 0.0 memory length: 36916	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.23		
episode: 208	score: 1.0 memory length: 37087	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.23		
episode: 209	score: 1.0 memory length: 37255	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.23		
episode: 210	score: 4.0 memory length: 37532	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.25		
episode: 211	score: 1.0 memory length: 37701	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.23		
episode: 212	score: 0.0 memory length: 37824	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.23		
episode: 213	score: 2.0 memory length: 38024	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.25		
episode: 214	score: 5.0 memory length: 38305	epsilon: 1.0	steps: 281
lr: 0.0001	evaluation reward: 1.28		
episode: 215	score: 1.0 memory length: 38456	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.28		
episode: 216	score: 0.0 memory length: 38579	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.26		
episode: 217	score: 2.0 memory length: 38800	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.27		
episode: 218	score: 4.0 memory length: 39077	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.29		
episode: 219	score: 1.0 memory length: 39246	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.3		
episode: 220	score: 1.0 memory length: 39417	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.31		
episode: 221	score: 2.0 memory length: 39615	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.33		
episode: 222	score: 2.0 memory length: 39813	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.33		
episode: 223	score: 3.0 memory length: 40058	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.32		
episode: 224	score: 0.0 memory length: 40181	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 225	score: 2.0 memory length: 40378	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.33		
episode: 226	score: 3.0 memory length: 40624	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.34		
episode: 227	score: 0.0 memory length: 40747	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 228	score: 2.0 memory length: 40967	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.32		
episode: 229	score: 4.0 memory length: 41243	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.34		
episode: 230	score: 2.0 memory length: 41444	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.35		
episode: 231	score: 1.0 memory length: 41615	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.34		
episode: 232	score: 1.0 memory length: 41766	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.34		
episode: 233	score: 1.0 memory length: 41938	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.35		
episode: 234	score: 3.0 memory length: 42205	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.37		
episode: 235	score: 1.0 memory length: 42373	epsilon: 1.0	steps: 168

lr: 0.0001	evaluation reward: 1.38		
episode: 236	score: 2.0 memory length: 42593	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.39		
episode: 237	score: 2.0 memory length: 42793	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.41		
episode: 238	score: 5.0 memory length: 43118	epsilon: 1.0	steps: 325
lr: 0.0001	evaluation reward: 1.45		
episode: 239	score: 2.0 memory length: 43316	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 240	score: 1.0 memory length: 43484	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.46		
episode: 241	score: 0.0 memory length: 43606	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44		
episode: 242	score: 2.0 memory length: 43804	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.46		
episode: 243	score: 0.0 memory length: 43927	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 244	score: 1.0 memory length: 44096	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.45		
episode: 245	score: 0.0 memory length: 44218	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45		
episode: 246	score: 0.0 memory length: 44341	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45		
episode: 247	score: 1.0 memory length: 44510	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.43		
episode: 248	score: 0.0 memory length: 44632	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.42		
episode: 249	score: 2.0 memory length: 44829	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.44		
episode: 250	score: 1.0 memory length: 45000	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.43		
episode: 251	score: 5.0 memory length: 45319	epsilon: 1.0	steps: 319
lr: 0.0001	evaluation reward: 1.46		
episode: 252	score: 3.0 memory length: 45584	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.47		
episode: 253	score: 0.0 memory length: 45707	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 254	score: 3.0 memory length: 45977	epsilon: 1.0	steps: 270
lr: 0.0001	evaluation reward: 1.45		
episode: 255	score: 0.0 memory length: 46099	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43		
episode: 256	score: 4.0 memory length: 46396	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.47		
episode: 257	score: 2.0 memory length: 46594	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49		
episode: 258	score: 0.0 memory length: 46717	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 259	score: 0.0 memory length: 46839	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.48		
episode: 260	score: 0.0 memory length: 46962	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 261	score: 1.0 memory length: 47133	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.47		
episode: 262	score: 1.0 memory length: 47302	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.48		
episode: 263	score: 2.0 memory length: 47500	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.46		
episode: 264	score: 2.0 memory length: 47720	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.44		
episode: 265	score: 1.0 memory length: 47891	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.43		
episode: 266	score: 1.0 memory length: 48059	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.44		
episode: 267	score: 0.0 memory length: 48181	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44		
episode: 268	score: 1.0 memory length: 48331	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.43		
episode: 269	score: 1.0 memory length: 48482	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.42		
episode: 270	score: 1.0 memory length: 48651	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.43		
episode: 271	score: 2.0 memory length: 48848	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.44		
episode: 272	score: 1.0 memory length: 48999	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.43		
episode: 273	score: 1.0 memory length: 49149	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.42		
episode: 274	score: 2.0 memory length: 49347	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.44		
episode: 275	score: 2.0 memory length: 49526	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.44		
episode: 276	score: 0.0 memory length: 49649	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43		
episode: 277	score: 0.0 memory length: 49772	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 278	score: 1.0 memory length: 49941	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.42		
episode: 279	score: 0.0 memory length: 50064	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 280	score: 1.0 memory length: 50235	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.39		
episode: 281	score: 0.0 memory length: 50358	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 282	score: 2.0 memory length: 50555	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.41		
episode: 283	score: 4.0 memory length: 50833	epsilon: 1.0	steps: 278

lr: 0.0001	evaluation reward: 1.43		
episode: 284	score: 4.0 memory length: 51092	epsilon: 1.0	steps: 259
lr: 0.0001	evaluation reward: 1.45		
episode: 285	score: 2.0 memory length: 51290	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.43		
episode: 286	score: 4.0 memory length: 51583	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.46		
episode: 287	score: 3.0 memory length: 51826	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.43		
episode: 288	score: 2.0 memory length: 52043	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.43		
episode: 289	score: 0.0 memory length: 52165	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43		
episode: 290	score: 1.0 memory length: 52335	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.42		
episode: 291	score: 0.0 memory length: 52458	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 292	score: 4.0 memory length: 52751	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.43		
episode: 293	score: 0.0 memory length: 52874	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43		
episode: 294	score: 3.0 memory length: 53119	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.46		
episode: 295	score: 1.0 memory length: 53288	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.46		
episode: 296	score: 4.0 memory length: 53563	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.47		
episode: 297	score: 2.0 memory length: 53761	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49		
episode: 298	score: 0.0 memory length: 53884	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 299	score: 1.0 memory length: 54053	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5		
episode: 300	score: 3.0 memory length: 54285	epsilon: 1.0	steps: 232
lr: 0.0001	evaluation reward: 1.52		
episode: 301	score: 1.0 memory length: 54436	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.52		
episode: 302	score: 1.0 memory length: 54604	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.53		
episode: 303	score: 2.0 memory length: 54822	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.53		
episode: 304	score: 1.0 memory length: 54991	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.52		
episode: 305	score: 0.0 memory length: 55114	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 306	score: 0.0 memory length: 55237	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 307	score: 2.0 memory length: 55435	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.54		
episode: 308	score: 0.0 memory length: 55557	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.53		
episode: 309	score: 1.0 memory length: 55726	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.53		
episode: 310	score: 0.0 memory length: 55849	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 311	score: 0.0 memory length: 55971	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.48		
episode: 312	score: 2.0 memory length: 56150	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.5		
episode: 313	score: 1.0 memory length: 56303	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.49		
episode: 314	score: 5.0 memory length: 56648	epsilon: 1.0	steps: 345
lr: 0.0001	evaluation reward: 1.49		
episode: 315	score: 2.0 memory length: 56866	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.5		
episode: 316	score: 0.0 memory length: 56989	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 317	score: 2.0 memory length: 57187	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5		
episode: 318	score: 0.0 memory length: 57310	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 319	score: 2.0 memory length: 57512	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.47		
episode: 320	score: 0.0 memory length: 57634	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.46		
episode: 321	score: 5.0 memory length: 57944	epsilon: 1.0	steps: 310
lr: 0.0001	evaluation reward: 1.49		
episode: 322	score: 0.0 memory length: 58066	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.47		
episode: 323	score: 2.0 memory length: 58282	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.46		
episode: 324	score: 3.0 memory length: 58531	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.49		
episode: 325	score: 1.0 memory length: 58701	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.48		
episode: 326	score: 2.0 memory length: 58919	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.47		
episode: 327	score: 1.0 memory length: 59088	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.48		
episode: 328	score: 1.0 memory length: 59257	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.47		
episode: 329	score: 0.0 memory length: 59380	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43		
episode: 330	score: 0.0 memory length: 59503	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 331	score: 1.0 memory length: 59673	epsilon: 1.0	steps: 170

lr: 0.0001	evaluation reward: 1.41		
episode: 332	score: 1.0 memory length: 59843	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.41		
episode: 333	score: 2.0 memory length: 60059	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.42		
episode: 334	score: 2.0 memory length: 60257	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41		
episode: 335	score: 5.0 memory length: 60589	epsilon: 1.0	steps: 332
lr: 0.0001	evaluation reward: 1.45		
episode: 336	score: 2.0 memory length: 60789	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.45		
episode: 337	score: 2.0 memory length: 60974	epsilon: 1.0	steps: 185
lr: 0.0001	evaluation reward: 1.45		
episode: 338	score: 1.0 memory length: 61143	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.41		
episode: 339	score: 5.0 memory length: 61452	epsilon: 1.0	steps: 309
lr: 0.0001	evaluation reward: 1.44		
episode: 340	score: 2.0 memory length: 61650	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45		
episode: 341	score: 1.0 memory length: 61821	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.46		
episode: 342	score: 1.0 memory length: 61992	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.45		
episode: 343	score: 3.0 memory length: 62239	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.48		
episode: 344	score: 0.0 memory length: 62362	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 345	score: 2.0 memory length: 62560	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49		
episode: 346	score: 5.0 memory length: 62864	epsilon: 1.0	steps: 304
lr: 0.0001	evaluation reward: 1.54		
episode: 347	score: 1.0 memory length: 63033	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.54		
episode: 348	score: 2.0 memory length: 63232	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.56		
episode: 349	score: 5.0 memory length: 63553	epsilon: 1.0	steps: 321
lr: 0.0001	evaluation reward: 1.59		
episode: 350	score: 1.0 memory length: 63721	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.59		
episode: 351	score: 1.0 memory length: 63890	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.55		
episode: 352	score: 1.0 memory length: 64061	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.53		
episode: 353	score: 0.0 memory length: 64183	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.53		
episode: 354	score: 1.0 memory length: 64334	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.51		
episode: 355	score: 2.0 memory length: 64552	epsilon: 1.0	steps: 218

lr: 0.0001	evaluation reward: 1.53			
episode: 356	score: 0.0	memory length: 64675	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 357	score: 7.0	memory length: 65098	epsilon: 1.0	steps: 423
lr: 0.0001	evaluation reward: 1.54			
episode: 358	score: 0.0	memory length: 65221	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.54			
episode: 359	score: 1.0	memory length: 65390	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.55			
episode: 360	score: 1.0	memory length: 65560	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.56			
episode: 361	score: 2.0	memory length: 65758	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.57			
episode: 362	score: 1.0	memory length: 65909	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.57			
episode: 363	score: 1.0	memory length: 66078	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.56			
episode: 364	score: 4.0	memory length: 66373	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.58			
episode: 365	score: 0.0	memory length: 66495	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.57			
episode: 366	score: 3.0	memory length: 66746	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.59			
episode: 367	score: 1.0	memory length: 66915	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.6			
episode: 368	score: 2.0	memory length: 67138	epsilon: 1.0	steps: 223
lr: 0.0001	evaluation reward: 1.61			
episode: 369	score: 0.0	memory length: 67261	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6			
episode: 370	score: 3.0	memory length: 67507	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.62			
episode: 371	score: 2.0	memory length: 67704	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.62			
episode: 372	score: 5.0	memory length: 67987	epsilon: 1.0	steps: 283
lr: 0.0001	evaluation reward: 1.66			
episode: 373	score: 3.0	memory length: 68235	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.68			
episode: 374	score: 2.0	memory length: 68453	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.68			
episode: 375	score: 4.0	memory length: 68710	epsilon: 1.0	steps: 257
lr: 0.0001	evaluation reward: 1.7			
episode: 376	score: 2.0	memory length: 68907	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.72			
episode: 377	score: 0.0	memory length: 69030	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.72			
episode: 378	score: 4.0	memory length: 69346	epsilon: 1.0	steps: 316
lr: 0.0001	evaluation reward: 1.75			
episode: 379	score: 1.0	memory length: 69497	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.76		
episode: 380	score: 2.0 memory length: 69695	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.77		
episode: 381	score: 3.0 memory length: 69962	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.8		
episode: 382	score: 1.0 memory length: 70131	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.79		
episode: 383	score: 2.0 memory length: 70329	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.77		
episode: 384	score: 3.0 memory length: 70557	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.76		
episode: 385	score: 0.0 memory length: 70679	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.74		
episode: 386	score: 1.0 memory length: 70848	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.71		
episode: 387	score: 2.0 memory length: 71046	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.7		
episode: 388	score: 0.0 memory length: 71169	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.68		
episode: 389	score: 0.0 memory length: 71292	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.68		
episode: 390	score: 0.0 memory length: 71415	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67		
episode: 391	score: 0.0 memory length: 71537	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.67		
episode: 392	score: 3.0 memory length: 71801	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.66		
episode: 393	score: 1.0 memory length: 71972	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.67		
episode: 394	score: 1.0 memory length: 72141	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.65		
episode: 395	score: 0.0 memory length: 72263	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.64		
episode: 396	score: 2.0 memory length: 72479	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.62		
episode: 397	score: 0.0 memory length: 72602	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6		
episode: 398	score: 3.0 memory length: 72850	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.63		
episode: 399	score: 1.0 memory length: 73019	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.63		
episode: 400	score: 1.0 memory length: 73189	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.61		
episode: 401	score: 3.0 memory length: 73437	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.63		
episode: 402	score: 0.0 memory length: 73560	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 403	score: 1.0 memory length: 73729	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.61		
episode: 404	score: 0.0 memory length: 73852	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6		
episode: 405	score: 1.0 memory length: 74003	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.61		
episode: 406	score: 0.0 memory length: 74126	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61		
episode: 407	score: 0.0 memory length: 74249	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59		
episode: 408	score: 0.0 memory length: 74372	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59		
episode: 409	score: 3.0 memory length: 74606	epsilon: 1.0	steps: 234
lr: 0.0001	evaluation reward: 1.61		
episode: 410	score: 2.0 memory length: 74822	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.63		
episode: 411	score: 4.0 memory length: 75101	epsilon: 1.0	steps: 279
lr: 0.0001	evaluation reward: 1.67		
episode: 412	score: 0.0 memory length: 75224	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65		
episode: 413	score: 2.0 memory length: 75424	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.66		
episode: 414	score: 1.0 memory length: 75594	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62		
episode: 415	score: 2.0 memory length: 75794	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.62		
episode: 416	score: 1.0 memory length: 75944	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.63		
episode: 417	score: 0.0 memory length: 76066	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.61		
episode: 418	score: 1.0 memory length: 76237	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.62		
episode: 419	score: 2.0 memory length: 76455	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.62		
episode: 420	score: 1.0 memory length: 76624	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.63		
episode: 421	score: 1.0 memory length: 76775	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.59		
episode: 422	score: 2.0 memory length: 76975	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.61		
episode: 423	score: 4.0 memory length: 77250	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.63		
episode: 424	score: 2.0 memory length: 77448	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.62		
episode: 425	score: 1.0 memory length: 77598	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.62		
episode: 426	score: 1.0 memory length: 77766	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.61		
episode: 427	score: 2.0 memory length: 77964	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.62		
episode: 428	score: 1.0 memory length: 78133	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.62		
episode: 429	score: 5.0 memory length: 78461	epsilon: 1.0	steps: 328
lr: 0.0001	evaluation reward: 1.67		
episode: 430	score: 1.0 memory length: 78612	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.68		
episode: 431	score: 3.0 memory length: 78840	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.7		
episode: 432	score: 0.0 memory length: 78963	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.69		
episode: 433	score: 1.0 memory length: 79114	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.68		
episode: 434	score: 0.0 memory length: 79237	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66		
episode: 435	score: 1.0 memory length: 79406	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.62		
episode: 436	score: 3.0 memory length: 79634	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.63		
episode: 437	score: 2.0 memory length: 79856	epsilon: 1.0	steps: 222
lr: 0.0001	evaluation reward: 1.63		
episode: 438	score: 3.0 memory length: 80104	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.65		
episode: 439	score: 1.0 memory length: 80274	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.61		
episode: 440	score: 2.0 memory length: 80492	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.61		
episode: 441	score: 3.0 memory length: 80721	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.63		
episode: 442	score: 2.0 memory length: 80936	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.64		
episode: 443	score: 0.0 memory length: 81059	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61		
episode: 444	score: 4.0 memory length: 81334	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.65		
episode: 445	score: 3.0 memory length: 81601	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.66		
episode: 446	score: 1.0 memory length: 81751	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.62		
episode: 447	score: 1.0 memory length: 81920	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.62		
episode: 448	score: 3.0 memory length: 82168	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.63		
episode: 449	score: 1.0 memory length: 82337	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.59		
episode: 450	score: 1.0 memory length: 82506	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.59		
episode: 451	score: 2.0 memory length: 82704	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.6		
episode: 452	score: 1.0 memory length: 82857	epsilon: 1.0	steps: 153
lr: 0.0001	evaluation reward: 1.6		
episode: 453	score: 1.0 memory length: 83026	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.61		
episode: 454	score: 1.0 memory length: 83177	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.61		
episode: 455	score: 1.0 memory length: 83327	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.6		
episode: 456	score: 0.0 memory length: 83450	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6		
episode: 457	score: 2.0 memory length: 83647	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.55		
episode: 458	score: 4.0 memory length: 83945	epsilon: 1.0	steps: 298
lr: 0.0001	evaluation reward: 1.59		
episode: 459	score: 3.0 memory length: 84194	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.61		
episode: 460	score: 2.0 memory length: 84413	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.62		
episode: 461	score: 0.0 memory length: 84535	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.6		
episode: 462	score: 1.0 memory length: 84705	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.6		
episode: 463	score: 2.0 memory length: 84922	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.61		
episode: 464	score: 4.0 memory length: 85218	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.61		
episode: 465	score: 1.0 memory length: 85389	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.62		
episode: 466	score: 1.0 memory length: 85541	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.6		
episode: 467	score: 0.0 memory length: 85664	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59		
episode: 468	score: 3.0 memory length: 85911	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.6		
episode: 469	score: 0.0 memory length: 86034	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6		
episode: 470	score: 2.0 memory length: 86253	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.59		
episode: 471	score: 4.0 memory length: 86509	epsilon: 1.0	steps: 256
lr: 0.0001	evaluation reward: 1.61		
episode: 472	score: 4.0 memory length: 86774	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.6		
episode: 473	score: 0.0 memory length: 86897	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57		
episode: 474	score: 0.0 memory length: 87019	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.55		
episode: 475	score: 3.0 memory length: 87286	epsilon: 1.0	steps: 267

lr: 0.0001	evaluation reward: 1.54		
episode: 476	score: 3.0 memory length: 87537	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.55		
episode: 477	score: 0.0 memory length: 87660	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55		
episode: 478	score: 2.0 memory length: 87857	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.53		
episode: 479	score: 2.0 memory length: 88055	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.54		
episode: 480	score: 0.0 memory length: 88178	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 481	score: 4.0 memory length: 88438	epsilon: 1.0	steps: 260
lr: 0.0001	evaluation reward: 1.53		
episode: 482	score: 0.0 memory length: 88561	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 483	score: 1.0 memory length: 88711	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51		
episode: 484	score: 4.0 memory length: 89025	epsilon: 1.0	steps: 314
lr: 0.0001	evaluation reward: 1.52		
episode: 485	score: 2.0 memory length: 89225	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.54		
episode: 486	score: 0.0 memory length: 89348	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53		
episode: 487	score: 2.0 memory length: 89548	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.53		
episode: 488	score: 4.0 memory length: 89844	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.57		
episode: 489	score: 0.0 memory length: 89967	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57		
episode: 490	score: 7.0 memory length: 90327	epsilon: 1.0	steps: 360
lr: 0.0001	evaluation reward: 1.64		
episode: 491	score: 4.0 memory length: 90645	epsilon: 1.0	steps: 318
lr: 0.0001	evaluation reward: 1.68		
episode: 492	score: 2.0 memory length: 90866	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.67		
episode: 493	score: 4.0 memory length: 91159	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.7		
episode: 494	score: 2.0 memory length: 91356	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.71		
episode: 495	score: 1.0 memory length: 91525	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.72		
episode: 496	score: 0.0 memory length: 91648	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 497	score: 0.0 memory length: 91771	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 498	score: 3.0 memory length: 92016	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.7		
episode: 499	score: 4.0 memory length: 92308	epsilon: 1.0	steps: 292

lr: 0.0001	evaluation reward: 1.73		
episode: 500	score: 1.0 memory length: 92459	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.73		
episode: 501	score: 1.0 memory length: 92628	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.71		
episode: 502	score: 6.0 memory length: 93005	epsilon: 1.0	steps: 377
lr: 0.0001	evaluation reward: 1.77		
episode: 503	score: 2.0 memory length: 93224	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.78		
episode: 504	score: 4.0 memory length: 93520	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.82		
episode: 505	score: 3.0 memory length: 93763	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.84		
episode: 506	score: 1.0 memory length: 93913	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.85		
episode: 507	score: 0.0 memory length: 94036	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.85		
episode: 508	score: 2.0 memory length: 94252	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.87		
episode: 509	score: 2.0 memory length: 94451	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.86		
episode: 510	score: 4.0 memory length: 94770	epsilon: 1.0	steps: 319
lr: 0.0001	evaluation reward: 1.88		
episode: 511	score: 1.0 memory length: 94938	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.85		
episode: 512	score: 3.0 memory length: 95183	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.88		
episode: 513	score: 1.0 memory length: 95353	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.87		
episode: 514	score: 0.0 memory length: 95475	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.86		
episode: 515	score: 0.0 memory length: 95597	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.84		
episode: 516	score: 2.0 memory length: 95817	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.85		
episode: 517	score: 1.0 memory length: 95986	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.86		
episode: 518	score: 0.0 memory length: 96109	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.85		
episode: 519	score: 3.0 memory length: 96334	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.86		
episode: 520	score: 3.0 memory length: 96585	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.88		
episode: 521	score: 3.0 memory length: 96848	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.9		
episode: 522	score: 4.0 memory length: 97123	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.92		
episode: 523	score: 1.0 memory length: 97294	epsilon: 1.0	steps: 171

lr: 0.0001	evaluation reward: 1.89		
episode: 524	score: 1.0	memory length: 97445	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.88		
episode: 525	score: 0.0	memory length: 97567	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.87		
episode: 526	score: 0.0	memory length: 97690	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.86		
episode: 527	score: 0.0	memory length: 97813	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.84		
episode: 528	score: 0.0	memory length: 97936	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.83		
episode: 529	score: 0.0	memory length: 98058	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.78		
episode: 530	score: 0.0	memory length: 98180	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.77		
episode: 531	score: 3.0	memory length: 98449	epsilon: 1.0 steps: 269
lr: 0.0001	evaluation reward: 1.77		
episode: 532	score: 2.0	memory length: 98647	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.79		
episode: 533	score: 0.0	memory length: 98769	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.78		
episode: 534	score: 2.0	memory length: 98987	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.8		
episode: 535	score: 2.0	memory length: 99184	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.81		
episode: 536	score: 0.0	memory length: 99307	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.78		
episode: 537	score: 0.0	memory length: 99429	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.76		
episode: 538	score: 2.0	memory length: 99647	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.75		
episode: 539	score: 3.0	memory length: 99912	epsilon: 1.0 steps: 265
lr: 0.0001	evaluation reward: 1.77		
episode: 540	score: 5.0	memory length: 100000	epsilon: 0.9995703400000093
steps: 304	lr: 0.0001	evaluation reward: 1.8	
episode: 541	score: 1.0	memory length: 100000	epsilon: 0.9992713600000158
steps: 151	lr: 0.0001	evaluation reward: 1.78	
episode: 542	score: 4.0	memory length: 100000	epsilon: 0.9986833000000286
steps: 297	lr: 0.0001	evaluation reward: 1.8	
episode: 543	score: 0.0	memory length: 100000	epsilon: 0.9984397600000339
steps: 123	lr: 0.0001	evaluation reward: 1.8	
episode: 544	score: 2.0	memory length: 100000	epsilon: 0.9980081200000432
steps: 218	lr: 0.0001	evaluation reward: 1.78	
episode: 545	score: 1.0	memory length: 100000	epsilon: 0.9977111200000497
steps: 150	lr: 0.0001	evaluation reward: 1.76	
episode: 546	score: 0.0	memory length: 100000	epsilon: 0.9974675800000055
steps: 123	lr: 0.0001	evaluation reward: 1.75	
episode: 547	score: 1.0	memory length: 100000	epsilon: 0.99716860000000615

steps: 151	lr: 0.0001	evaluation reward: 1.75
episode: 548	score: 1.0	memory length: 100000 epsilon: 0.996869620000068
steps: 151	lr: 0.0001	evaluation reward: 1.73
episode: 549	score: 1.0	memory length: 100000 epsilon: 0.9965290600000754
steps: 172	lr: 0.0001	evaluation reward: 1.73
episode: 550	score: 2.0	memory length: 100000 epsilon: 0.9961350400000839
steps: 199	lr: 0.0001	evaluation reward: 1.74
episode: 551	score: 2.0	memory length: 100000 epsilon: 0.9957430000000924
steps: 198	lr: 0.0001	evaluation reward: 1.74
episode: 552	score: 0.0	memory length: 100000 epsilon: 0.9955014400000977
steps: 122	lr: 0.0001	evaluation reward: 1.73
episode: 553	score: 0.0	memory length: 100000 epsilon: 0.995257900000103
steps: 123	lr: 0.0001	evaluation reward: 1.72
episode: 554	score: 6.0	memory length: 100000 epsilon: 0.99447184000012
steps: 397	lr: 0.0001	evaluation reward: 1.77
episode: 555	score: 1.0	memory length: 100000 epsilon: 0.9941689000001266
steps: 153	lr: 0.0001	evaluation reward: 1.77
episode: 556	score: 0.0	memory length: 100000 epsilon: 0.9939253600001319
steps: 123	lr: 0.0001	evaluation reward: 1.77
episode: 557	score: 0.0	memory length: 100000 epsilon: 0.9936818200001372
steps: 123	lr: 0.0001	evaluation reward: 1.75
episode: 558	score: 2.0	memory length: 100000 epsilon: 0.9932917600001456
steps: 197	lr: 0.0001	evaluation reward: 1.73
episode: 559	score: 2.0	memory length: 100000 epsilon: 0.9928997200001541
steps: 198	lr: 0.0001	evaluation reward: 1.72
episode: 560	score: 1.0	memory length: 100000 epsilon: 0.9925987600001607
steps: 152	lr: 0.0001	evaluation reward: 1.71
episode: 561	score: 0.0	memory length: 100000 epsilon: 0.9923572000001659
steps: 122	lr: 0.0001	evaluation reward: 1.71
episode: 562	score: 0.0	memory length: 100000 epsilon: 0.9921156400001712
steps: 122	lr: 0.0001	evaluation reward: 1.7
episode: 563	score: 0.0	memory length: 100000 epsilon: 0.9918721000001764
steps: 123	lr: 0.0001	evaluation reward: 1.68
episode: 564	score: 1.0	memory length: 100000 epsilon: 0.9915335200001838
steps: 171	lr: 0.0001	evaluation reward: 1.65
episode: 565	score: 1.0	memory length: 100000 epsilon: 0.9911969200001911
steps: 170	lr: 0.0001	evaluation reward: 1.65
episode: 566	score: 1.0	memory length: 100000 epsilon: 0.9908583400001985
steps: 171	lr: 0.0001	evaluation reward: 1.65
episode: 567	score: 3.0	memory length: 100000 epsilon: 0.9903692800002091
steps: 247	lr: 0.0001	evaluation reward: 1.68
episode: 568	score: 0.0	memory length: 100000 epsilon: 0.9901257400002144
steps: 123	lr: 0.0001	evaluation reward: 1.65
episode: 569	score: 0.0	memory length: 100000 epsilon: 0.9898822000002196
steps: 123	lr: 0.0001	evaluation reward: 1.65
episode: 570	score: 4.0	memory length: 100000 epsilon: 0.9893753200002307
steps: 256	lr: 0.0001	evaluation reward: 1.67
episode: 571	score: 3.0	memory length: 100000 epsilon: 0.9889278400002404

steps: 226	lr: 0.0001	evaluation reward: 1.66
episode: 572	score: 0.0	memory length: 100000 epsilon: 0.9886843000002457
steps: 123	lr: 0.0001	evaluation reward: 1.62
episode: 573	score: 0.0	memory length: 100000 epsilon: 0.9884427400002509
steps: 122	lr: 0.0001	evaluation reward: 1.62
episode: 574	score: 1.0	memory length: 100000 epsilon: 0.9881101000002581
steps: 168	lr: 0.0001	evaluation reward: 1.63
episode: 575	score: 2.0	memory length: 100000 epsilon: 0.9877200400002666
steps: 197	lr: 0.0001	evaluation reward: 1.62
episode: 576	score: 1.0	memory length: 100000 epsilon: 0.9873854200002739
steps: 169	lr: 0.0001	evaluation reward: 1.6
episode: 577	score: 0.0	memory length: 100000 epsilon: 0.9871438600002791
steps: 122	lr: 0.0001	evaluation reward: 1.6
episode: 578	score: 2.0	memory length: 100000 epsilon: 0.9867518200002876
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 579	score: 0.0	memory length: 100000 epsilon: 0.9865082800002929
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 580	score: 2.0	memory length: 100000 epsilon: 0.9861162400003014
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 581	score: 2.0	memory length: 100000 epsilon: 0.9856865800003107
steps: 217	lr: 0.0001	evaluation reward: 1.58
episode: 582	score: 3.0	memory length: 100000 epsilon: 0.9851539600003223
steps: 269	lr: 0.0001	evaluation reward: 1.61
episode: 583	score: 3.0	memory length: 100000 epsilon: 0.9846312400003336
steps: 264	lr: 0.0001	evaluation reward: 1.63
episode: 584	score: 1.0	memory length: 100000 epsilon: 0.9842966200003409
steps: 169	lr: 0.0001	evaluation reward: 1.6
episode: 585	score: 0.0	memory length: 100000 epsilon: 0.9840530800003462
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 586	score: 2.0	memory length: 100000 epsilon: 0.9836610400003547
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 587	score: 2.0	memory length: 100000 epsilon: 0.9832709800003632
steps: 197	lr: 0.0001	evaluation reward: 1.6
episode: 588	score: 2.0	memory length: 100000 epsilon: 0.9828433000003725
steps: 216	lr: 0.0001	evaluation reward: 1.58
episode: 589	score: 1.0	memory length: 100000 epsilon: 0.9825086800003797
steps: 169	lr: 0.0001	evaluation reward: 1.59
episode: 590	score: 2.0	memory length: 100000 epsilon: 0.9820750600003891
steps: 219	lr: 0.0001	evaluation reward: 1.54
episode: 591	score: 2.0	memory length: 100000 epsilon: 0.9816830200003976
steps: 198	lr: 0.0001	evaluation reward: 1.52
episode: 592	score: 0.0	memory length: 100000 epsilon: 0.9814394800004029
steps: 123	lr: 0.0001	evaluation reward: 1.5
episode: 593	score: 1.0	memory length: 100000 epsilon: 0.9811424800004094
steps: 150	lr: 0.0001	evaluation reward: 1.47
episode: 594	score: 0.0	memory length: 100000 epsilon: 0.9808989400004147
steps: 123	lr: 0.0001	evaluation reward: 1.45
episode: 595	score: 3.0	memory length: 100000 epsilon: 0.9803742400004261

steps: 265	lr: 0.0001	evaluation reward: 1.47
episode: 596	score: 1.0	memory length: 100000 epsilon: 0.9800356600004334
steps: 171	lr: 0.0001	evaluation reward: 1.48
episode: 597	score: 1.0	memory length: 100000 epsilon: 0.9796970800004408
steps: 171	lr: 0.0001	evaluation reward: 1.49
episode: 598	score: 0.0	memory length: 100000 epsilon: 0.979453540000446
steps: 123	lr: 0.0001	evaluation reward: 1.46
episode: 599	score: 1.0	memory length: 100000 epsilon: 0.9791189200004533
steps: 169	lr: 0.0001	evaluation reward: 1.43
episode: 600	score: 0.0	memory length: 100000 epsilon: 0.9788753800004586
steps: 123	lr: 0.0001	evaluation reward: 1.42
episode: 601	score: 3.0	memory length: 100000 epsilon: 0.9783863200004692
steps: 247	lr: 0.0001	evaluation reward: 1.44
episode: 602	score: 0.0	memory length: 100000 epsilon: 0.9781427800004745
steps: 123	lr: 0.0001	evaluation reward: 1.38
episode: 603	score: 2.0	memory length: 100000 epsilon: 0.9777111400004839
steps: 218	lr: 0.0001	evaluation reward: 1.38
episode: 604	score: 2.0	memory length: 100000 epsilon: 0.9773191000004924
steps: 198	lr: 0.0001	evaluation reward: 1.36
episode: 605	score: 0.0	memory length: 100000 epsilon: 0.9770755600004977
steps: 123	lr: 0.0001	evaluation reward: 1.33
episode: 606	score: 3.0	memory length: 100000 epsilon: 0.9766300600005073
steps: 225	lr: 0.0001	evaluation reward: 1.35
episode: 607	score: 1.0	memory length: 100000 epsilon: 0.9762934600005146
steps: 170	lr: 0.0001	evaluation reward: 1.36
episode: 608	score: 3.0	memory length: 100000 epsilon: 0.9758004400005253
steps: 249	lr: 0.0001	evaluation reward: 1.37
episode: 609	score: 2.0	memory length: 100000 epsilon: 0.9754084000005339
steps: 198	lr: 0.0001	evaluation reward: 1.37
episode: 610	score: 3.0	memory length: 100000 epsilon: 0.974943100000544
steps: 235	lr: 0.0001	evaluation reward: 1.36
episode: 611	score: 0.0	memory length: 100000 epsilon: 0.9746995600005492
steps: 123	lr: 0.0001	evaluation reward: 1.35
episode: 612	score: 0.0	memory length: 100000 epsilon: 0.9744560200005545
steps: 123	lr: 0.0001	evaluation reward: 1.32
episode: 613	score: 1.0	memory length: 100000 epsilon: 0.9741174400005619
steps: 171	lr: 0.0001	evaluation reward: 1.32
episode: 614	score: 5.0	memory length: 100000 epsilon: 0.9733967200005775
steps: 364	lr: 0.0001	evaluation reward: 1.37
episode: 615	score: 2.0	memory length: 100000 epsilon: 0.973006660000586
steps: 197	lr: 0.0001	evaluation reward: 1.39
episode: 616	score: 1.0	memory length: 100000 epsilon: 0.9726661000005934
steps: 172	lr: 0.0001	evaluation reward: 1.38
episode: 617	score: 0.0	memory length: 100000 epsilon: 0.9724225600005987
steps: 123	lr: 0.0001	evaluation reward: 1.37
episode: 618	score: 2.0	memory length: 100000 epsilon: 0.9720265600006073
steps: 200	lr: 0.0001	evaluation reward: 1.39
episode: 619	score: 4.0	memory length: 100000 epsilon: 0.971486020000619

steps: 273	lr: 0.0001	evaluation reward: 1.4	
episode: 620	score: 3.0	memory length: 100000	epsilon: 0.9709573600006305
steps: 267	lr: 0.0001	evaluation reward: 1.4	
episode: 621	score: 4.0	memory length: 100000	epsilon: 0.9704524600006414
steps: 255	lr: 0.0001	evaluation reward: 1.41	
episode: 622	score: 2.0	memory length: 100000	epsilon: 0.9700208200006508
steps: 218	lr: 0.0001	evaluation reward: 1.39	
episode: 623	score: 0.0	memory length: 100000	epsilon: 0.9697792600006561
steps: 122	lr: 0.0001	evaluation reward: 1.38	
episode: 624	score: 3.0	memory length: 100000	epsilon: 0.9692981200006665
steps: 243	lr: 0.0001	evaluation reward: 1.4	
episode: 625	score: 4.0	memory length: 100000	epsilon: 0.9687853000006776
steps: 259	lr: 0.0001	evaluation reward: 1.44	
episode: 626	score: 1.0	memory length: 100000	epsilon: 0.9684863200006841
steps: 151	lr: 0.0001	evaluation reward: 1.45	
episode: 627	score: 5.0	memory length: 100000	epsilon: 0.9678091600006988
steps: 342	lr: 0.0001	evaluation reward: 1.5	
episode: 628	score: 0.0	memory length: 100000	epsilon: 0.9675676000007041
steps: 122	lr: 0.0001	evaluation reward: 1.5	
episode: 629	score: 1.0	memory length: 100000	epsilon: 0.9672646600007107
steps: 153	lr: 0.0001	evaluation reward: 1.51	
episode: 630	score: 1.0	memory length: 100000	epsilon: 0.9669300400007179
steps: 169	lr: 0.0001	evaluation reward: 1.52	
episode: 631	score: 3.0	memory length: 100000	epsilon: 0.9664013800007294
steps: 267	lr: 0.0001	evaluation reward: 1.52	
episode: 632	score: 3.0	memory length: 100000	epsilon: 0.9659202400007398
steps: 243	lr: 0.0001	evaluation reward: 1.53	
episode: 633	score: 3.0	memory length: 100000	epsilon: 0.9654292000007505
steps: 248	lr: 0.0001	evaluation reward: 1.56	
episode: 634	score: 0.0	memory length: 100000	epsilon: 0.9651876400007557
steps: 122	lr: 0.0001	evaluation reward: 1.54	
episode: 635	score: 0.0	memory length: 100000	epsilon: 0.964944100000761
steps: 123	lr: 0.0001	evaluation reward: 1.52	
episode: 636	score: 0.0	memory length: 100000	epsilon: 0.9647005600007663
steps: 123	lr: 0.0001	evaluation reward: 1.52	
episode: 637	score: 1.0	memory length: 100000	epsilon: 0.9644035600007728
steps: 150	lr: 0.0001	evaluation reward: 1.53	
episode: 638	score: 0.0	memory length: 100000	epsilon: 0.964160020000778
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 639	score: 2.0	memory length: 100000	epsilon: 0.9637303600007874
steps: 217	lr: 0.0001	evaluation reward: 1.5	
episode: 640	score: 3.0	memory length: 100000	epsilon: 0.9632056600007988
steps: 265	lr: 0.0001	evaluation reward: 1.48	
episode: 641	score: 0.0	memory length: 100000	epsilon: 0.9629621200008041
steps: 123	lr: 0.0001	evaluation reward: 1.47	
episode: 642	score: 1.0	memory length: 100000	epsilon: 0.9626275000008113
steps: 169	lr: 0.0001	evaluation reward: 1.44	
episode: 643	score: 4.0	memory length: 100000	epsilon: 0.9620711200008234

steps: 281	lr: 0.0001	evaluation reward: 1.48
episode: 644	score: 0.0	memory length: 100000 epsilon: 0.9618295600008286
steps: 122	lr: 0.0001	evaluation reward: 1.46
episode: 645	score: 2.0	memory length: 100000 epsilon: 0.9614375200008372
steps: 198	lr: 0.0001	evaluation reward: 1.47
episode: 646	score: 0.0	memory length: 100000 epsilon: 0.9611939800008424
steps: 123	lr: 0.0001	evaluation reward: 1.47
episode: 647	score: 0.0	memory length: 100000 epsilon: 0.9609504400008477
steps: 123	lr: 0.0001	evaluation reward: 1.46
episode: 648	score: 1.0	memory length: 100000 epsilon: 0.9606514600008542
steps: 151	lr: 0.0001	evaluation reward: 1.46
episode: 649	score: 2.0	memory length: 100000 epsilon: 0.9602218000008635
steps: 217	lr: 0.0001	evaluation reward: 1.47
episode: 650	score: 2.0	memory length: 100000 epsilon: 0.959831740000872
steps: 197	lr: 0.0001	evaluation reward: 1.47
episode: 651	score: 2.0	memory length: 100000 epsilon: 0.9594001000008814
steps: 218	lr: 0.0001	evaluation reward: 1.47
episode: 652	score: 0.0	memory length: 100000 epsilon: 0.9591565600008867
steps: 123	lr: 0.0001	evaluation reward: 1.47
episode: 653	score: 1.0	memory length: 100000 epsilon: 0.9588239200008939
steps: 168	lr: 0.0001	evaluation reward: 1.48
episode: 654	score: 0.0	memory length: 100000 epsilon: 0.9585803800008992
steps: 123	lr: 0.0001	evaluation reward: 1.42
episode: 655	score: 2.0	memory length: 100000 epsilon: 0.9581487400009085
steps: 218	lr: 0.0001	evaluation reward: 1.43
episode: 656	score: 2.0	memory length: 100000 epsilon: 0.9577567000009171
steps: 198	lr: 0.0001	evaluation reward: 1.45
episode: 657	score: 1.0	memory length: 100000 epsilon: 0.9574220800009243
steps: 169	lr: 0.0001	evaluation reward: 1.46
episode: 658	score: 4.0	memory length: 100000 epsilon: 0.9568716400009363
steps: 278	lr: 0.0001	evaluation reward: 1.48
episode: 659	score: 2.0	memory length: 100000 epsilon: 0.9564796000009448
steps: 198	lr: 0.0001	evaluation reward: 1.48
episode: 660	score: 0.0	memory length: 100000 epsilon: 0.9562360600009501
steps: 123	lr: 0.0001	evaluation reward: 1.47
episode: 661	score: 2.0	memory length: 100000 epsilon: 0.9558044200009594
steps: 218	lr: 0.0001	evaluation reward: 1.49
episode: 662	score: 1.0	memory length: 100000 epsilon: 0.9555054400009659
steps: 151	lr: 0.0001	evaluation reward: 1.5
episode: 663	score: 2.0	memory length: 100000 epsilon: 0.9550738000009753
steps: 218	lr: 0.0001	evaluation reward: 1.52
episode: 664	score: 2.0	memory length: 100000 epsilon: 0.9546778000009839
steps: 200	lr: 0.0001	evaluation reward: 1.53
episode: 665	score: 5.0	memory length: 100000 epsilon: 0.9540501400009975
steps: 317	lr: 0.0001	evaluation reward: 1.57
episode: 666	score: 2.0	memory length: 100000 epsilon: 0.9536937400010053
steps: 180	lr: 0.0001	evaluation reward: 1.58
episode: 667	score: 0.0	memory length: 100000 epsilon: 0.9534502000010106

steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 668	score: 1.0	memory length: 100000 epsilon: 0.9531155800010178
steps: 169	lr: 0.0001	evaluation reward: 1.56
episode: 669	score: 2.0	memory length: 100000 epsilon: 0.9526839400010272
steps: 218	lr: 0.0001	evaluation reward: 1.58
episode: 670	score: 0.0	memory length: 100000 epsilon: 0.9524423800010324
steps: 122	lr: 0.0001	evaluation reward: 1.54
episode: 671	score: 1.0	memory length: 100000 epsilon: 0.9521434000010389
steps: 151	lr: 0.0001	evaluation reward: 1.52
episode: 672	score: 4.0	memory length: 100000 epsilon: 0.9515989000010507
steps: 275	lr: 0.0001	evaluation reward: 1.56
episode: 673	score: 2.0	memory length: 100000 epsilon: 0.9512425000010585
steps: 180	lr: 0.0001	evaluation reward: 1.58
episode: 674	score: 1.0	memory length: 100000 epsilon: 0.9509078800010657
steps: 169	lr: 0.0001	evaluation reward: 1.58
episode: 675	score: 2.0	memory length: 100000 epsilon: 0.9505475200010736
steps: 182	lr: 0.0001	evaluation reward: 1.58
episode: 676	score: 2.0	memory length: 100000 epsilon: 0.950157460001082
steps: 197	lr: 0.0001	evaluation reward: 1.59
episode: 677	score: 6.0	memory length: 100000 epsilon: 0.9493496200010996
steps: 408	lr: 0.0001	evaluation reward: 1.65
episode: 678	score: 3.0	memory length: 100000 epsilon: 0.9489298600011087
steps: 212	lr: 0.0001	evaluation reward: 1.66
episode: 679	score: 1.0	memory length: 100000 epsilon: 0.948595240001116
steps: 169	lr: 0.0001	evaluation reward: 1.67
episode: 680	score: 0.0	memory length: 100000 epsilon: 0.9483517000011212
steps: 123	lr: 0.0001	evaluation reward: 1.65
episode: 681	score: 0.0	memory length: 100000 epsilon: 0.9481081600011265
steps: 123	lr: 0.0001	evaluation reward: 1.63
episode: 682	score: 1.0	memory length: 100000 epsilon: 0.9477695800011339
steps: 171	lr: 0.0001	evaluation reward: 1.61
episode: 683	score: 2.0	memory length: 100000 epsilon: 0.9473339800011433
steps: 220	lr: 0.0001	evaluation reward: 1.6
episode: 684	score: 0.0	memory length: 100000 epsilon: 0.9470904400011486
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 685	score: 3.0	memory length: 100000 epsilon: 0.9465558400011602
steps: 270	lr: 0.0001	evaluation reward: 1.62
episode: 686	score: 2.0	memory length: 100000 epsilon: 0.9461657800011687
steps: 197	lr: 0.0001	evaluation reward: 1.62
episode: 687	score: 1.0	memory length: 100000 epsilon: 0.9458648200011752
steps: 152	lr: 0.0001	evaluation reward: 1.61
episode: 688	score: 2.0	memory length: 100000 epsilon: 0.9454688200011838
steps: 200	lr: 0.0001	evaluation reward: 1.61
episode: 689	score: 1.0	memory length: 100000 epsilon: 0.9451342000011911
steps: 169	lr: 0.0001	evaluation reward: 1.61
episode: 690	score: 4.0	memory length: 100000 epsilon: 0.9445501000012038
steps: 295	lr: 0.0001	evaluation reward: 1.63
episode: 691	score: 3.0	memory length: 100000 epsilon: 0.9440986600012136

steps: 228	lr: 0.0001	evaluation reward: 1.64
episode: 692	score: 4.0	memory length: 100000 epsilon: 0.9435185200012262
steps: 293	lr: 0.0001	evaluation reward: 1.68
episode: 693	score: 1.0	memory length: 100000 epsilon: 0.9431839000012334
steps: 169	lr: 0.0001	evaluation reward: 1.68
episode: 694	score: 0.0	memory length: 100000 epsilon: 0.9429403600012387
steps: 123	lr: 0.0001	evaluation reward: 1.68
episode: 695	score: 1.0	memory length: 100000 epsilon: 0.942605740001246
steps: 169	lr: 0.0001	evaluation reward: 1.66
episode: 696	score: 7.0	memory length: 100000 epsilon: 0.9417998800012635
steps: 407	lr: 0.0001	evaluation reward: 1.72
episode: 697	score: 4.0	memory length: 100000 epsilon: 0.9412514200012754
steps: 277	lr: 0.0001	evaluation reward: 1.75
episode: 698	score: 1.0	memory length: 100000 epsilon: 0.9409524400012819
steps: 151	lr: 0.0001	evaluation reward: 1.76
episode: 699	score: 1.0	memory length: 100000 epsilon: 0.9406534600012884
steps: 151	lr: 0.0001	evaluation reward: 1.76
episode: 700	score: 2.0	memory length: 100000 epsilon: 0.9402970600012961
steps: 180	lr: 0.0001	evaluation reward: 1.78
episode: 701	score: 1.0	memory length: 100000 epsilon: 0.9399565000013035
steps: 172	lr: 0.0001	evaluation reward: 1.76
episode: 702	score: 1.0	memory length: 100000 epsilon: 0.93965752000131
steps: 151	lr: 0.0001	evaluation reward: 1.77
episode: 703	score: 1.0	memory length: 100000 epsilon: 0.9393248800013172
steps: 168	lr: 0.0001	evaluation reward: 1.76
episode: 704	score: 0.0	memory length: 100000 epsilon: 0.9390813400013225
steps: 123	lr: 0.0001	evaluation reward: 1.74
episode: 705	score: 2.0	memory length: 100000 epsilon: 0.9386477200013319
steps: 219	lr: 0.0001	evaluation reward: 1.76
episode: 706	score: 2.0	memory length: 100000 epsilon: 0.9382121200013414
steps: 220	lr: 0.0001	evaluation reward: 1.75
episode: 707	score: 2.0	memory length: 100000 epsilon: 0.9378200800013499
steps: 198	lr: 0.0001	evaluation reward: 1.76
episode: 708	score: 0.0	memory length: 100000 epsilon: 0.9375785200013551
steps: 122	lr: 0.0001	evaluation reward: 1.73
episode: 709	score: 2.0	memory length: 100000 epsilon: 0.9371488600013644
steps: 217	lr: 0.0001	evaluation reward: 1.73
episode: 710	score: 0.0	memory length: 100000 epsilon: 0.9369053200013697
steps: 123	lr: 0.0001	evaluation reward: 1.7
episode: 711	score: 4.0	memory length: 100000 epsilon: 0.9363608200013815
steps: 275	lr: 0.0001	evaluation reward: 1.74
episode: 712	score: 0.0	memory length: 100000 epsilon: 0.9361172800013868
steps: 123	lr: 0.0001	evaluation reward: 1.74
episode: 713	score: 0.0	memory length: 100000 epsilon: 0.9358757200013921
steps: 122	lr: 0.0001	evaluation reward: 1.73
episode: 714	score: 0.0	memory length: 100000 epsilon: 0.9356341600013973
steps: 122	lr: 0.0001	evaluation reward: 1.68
episode: 715	score: 0.0	memory length: 100000 epsilon: 0.9353906200014026

steps: 123	lr: 0.0001	evaluation reward: 1.66
episode: 716	score: 1.0	memory length: 100000 epsilon: 0.93505006000141
steps: 172	lr: 0.0001	evaluation reward: 1.66
episode: 717	score: 0.0	memory length: 100000 epsilon: 0.9348065200014153
steps: 123	lr: 0.0001	evaluation reward: 1.66
episode: 718	score: 3.0	memory length: 100000 epsilon: 0.934315480001426
steps: 248	lr: 0.0001	evaluation reward: 1.67
episode: 719	score: 1.0	memory length: 100000 epsilon: 0.9339769000014333
steps: 171	lr: 0.0001	evaluation reward: 1.64
episode: 720	score: 1.0	memory length: 100000 epsilon: 0.9336422800014406
steps: 169	lr: 0.0001	evaluation reward: 1.62
episode: 721	score: 3.0	memory length: 100000 epsilon: 0.933161140001451
steps: 243	lr: 0.0001	evaluation reward: 1.61
episode: 722	score: 0.0	memory length: 100000 epsilon: 0.9329176000014563
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 723	score: 3.0	memory length: 100000 epsilon: 0.932470120001466
steps: 226	lr: 0.0001	evaluation reward: 1.62
episode: 724	score: 2.0	memory length: 100000 epsilon: 0.9320780800014745
steps: 198	lr: 0.0001	evaluation reward: 1.61
episode: 725	score: 2.0	memory length: 100000 epsilon: 0.9317177200014823
steps: 182	lr: 0.0001	evaluation reward: 1.59
episode: 726	score: 0.0	memory length: 100000 epsilon: 0.9314741800014876
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 727	score: 1.0	memory length: 100000 epsilon: 0.9311752000014941
steps: 151	lr: 0.0001	evaluation reward: 1.54
episode: 728	score: 2.0	memory length: 100000 epsilon: 0.9307435600015035
steps: 218	lr: 0.0001	evaluation reward: 1.56
episode: 729	score: 3.0	memory length: 100000 epsilon: 0.9302564800015141
steps: 246	lr: 0.0001	evaluation reward: 1.58
episode: 730	score: 2.0	memory length: 100000 epsilon: 0.9298268200015234
steps: 217	lr: 0.0001	evaluation reward: 1.59
episode: 731	score: 1.0	memory length: 100000 epsilon: 0.9295278400015299
steps: 151	lr: 0.0001	evaluation reward: 1.57
episode: 732	score: 2.0	memory length: 100000 epsilon: 0.9291318400015385
steps: 200	lr: 0.0001	evaluation reward: 1.56
episode: 733	score: 0.0	memory length: 100000 epsilon: 0.9288883000015438
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 734	score: 1.0	memory length: 100000 epsilon: 0.928553680001551
steps: 169	lr: 0.0001	evaluation reward: 1.54
episode: 735	score: 0.0	memory length: 100000 epsilon: 0.9283101400015563
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 736	score: 3.0	memory length: 100000 epsilon: 0.927819100001567
steps: 248	lr: 0.0001	evaluation reward: 1.57
episode: 737	score: 3.0	memory length: 100000 epsilon: 0.9272904400015785
steps: 267	lr: 0.0001	evaluation reward: 1.59
episode: 738	score: 1.0	memory length: 100000 epsilon: 0.9269914600015849
steps: 151	lr: 0.0001	evaluation reward: 1.6
episode: 739	score: 2.0	memory length: 100000 epsilon: 0.9266014000015934

steps: 197	lr: 0.0001	evaluation reward: 1.6
episode: 740	score: 1.0	memory length: 100000 epsilon: 0.9262984600016
steps: 153	lr: 0.0001	evaluation reward: 1.58
episode: 741	score: 1.0	memory length: 100000 epsilon: 0.9259994800016065
steps: 151	lr: 0.0001	evaluation reward: 1.59
episode: 742	score: 1.0	memory length: 100000 epsilon: 0.9256609000016138
steps: 171	lr: 0.0001	evaluation reward: 1.59
episode: 743	score: 1.0	memory length: 100000 epsilon: 0.9253262800016211
steps: 169	lr: 0.0001	evaluation reward: 1.56
episode: 744	score: 1.0	memory length: 100000 epsilon: 0.9249916600016284
steps: 169	lr: 0.0001	evaluation reward: 1.57
episode: 745	score: 2.0	memory length: 100000 epsilon: 0.9246016000016368
steps: 197	lr: 0.0001	evaluation reward: 1.57
episode: 746	score: 0.0	memory length: 100000 epsilon: 0.9243600400016421
steps: 122	lr: 0.0001	evaluation reward: 1.57
episode: 747	score: 0.0	memory length: 100000 epsilon: 0.9241184800016473
steps: 122	lr: 0.0001	evaluation reward: 1.57
episode: 748	score: 3.0	memory length: 100000 epsilon: 0.9236650600016572
steps: 229	lr: 0.0001	evaluation reward: 1.59
episode: 749	score: 1.0	memory length: 100000 epsilon: 0.9233304400016644
steps: 169	lr: 0.0001	evaluation reward: 1.58
episode: 750	score: 0.0	memory length: 100000 epsilon: 0.9230869000016697
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 751	score: 3.0	memory length: 100000 epsilon: 0.9225998200016803
steps: 246	lr: 0.0001	evaluation reward: 1.57
episode: 752	score: 3.0	memory length: 100000 epsilon: 0.92215234000169
steps: 226	lr: 0.0001	evaluation reward: 1.6
episode: 753	score: 4.0	memory length: 100000 epsilon: 0.9215682400017027
steps: 295	lr: 0.0001	evaluation reward: 1.63
episode: 754	score: 0.0	memory length: 100000 epsilon: 0.9213266800017079
steps: 122	lr: 0.0001	evaluation reward: 1.63
episode: 755	score: 2.0	memory length: 100000 epsilon: 0.9208970200017172
steps: 217	lr: 0.0001	evaluation reward: 1.63
episode: 756	score: 3.0	memory length: 100000 epsilon: 0.9204059800017279
steps: 248	lr: 0.0001	evaluation reward: 1.64
episode: 757	score: 2.0	memory length: 100000 epsilon: 0.9200139400017364
steps: 198	lr: 0.0001	evaluation reward: 1.65
episode: 758	score: 5.0	memory length: 100000 epsilon: 0.9193546000017507
steps: 333	lr: 0.0001	evaluation reward: 1.66
episode: 759	score: 3.0	memory length: 100000 epsilon: 0.9188655400017613
steps: 247	lr: 0.0001	evaluation reward: 1.67
episode: 760	score: 1.0	memory length: 100000 epsilon: 0.9185665600017678
steps: 151	lr: 0.0001	evaluation reward: 1.68
episode: 761	score: 2.0	memory length: 100000 epsilon: 0.9181388800017771
steps: 216	lr: 0.0001	evaluation reward: 1.68
episode: 762	score: 3.0	memory length: 100000 epsilon: 0.9176498200017877
steps: 247	lr: 0.0001	evaluation reward: 1.7
episode: 763	score: 2.0	memory length: 100000 epsilon: 0.9172934200017955

steps: 180	lr: 0.0001	evaluation reward: 1.7
episode: 764	score: 2.0	memory length: 100000 epsilon: 0.9169370200018032
steps: 180	lr: 0.0001	evaluation reward: 1.7
episode: 765	score: 1.0	memory length: 100000 epsilon: 0.9165984400018106
steps: 171	lr: 0.0001	evaluation reward: 1.66
episode: 766	score: 0.0	memory length: 100000 epsilon: 0.9163549000018159
steps: 123	lr: 0.0001	evaluation reward: 1.64
episode: 767	score: 1.0	memory length: 100000 epsilon: 0.9160222600018231
steps: 168	lr: 0.0001	evaluation reward: 1.65
episode: 768	score: 2.0	memory length: 100000 epsilon: 0.9156302200018316
steps: 198	lr: 0.0001	evaluation reward: 1.66
episode: 769	score: 0.0	memory length: 100000 epsilon: 0.9153866800018369
steps: 123	lr: 0.0001	evaluation reward: 1.64
episode: 770	score: 2.0	memory length: 100000 epsilon: 0.9150263200018447
steps: 182	lr: 0.0001	evaluation reward: 1.66
episode: 771	score: 0.0	memory length: 100000 epsilon: 0.91478278000185
steps: 123	lr: 0.0001	evaluation reward: 1.65
episode: 772	score: 0.0	memory length: 100000 epsilon: 0.9145392400018553
steps: 123	lr: 0.0001	evaluation reward: 1.61
episode: 773	score: 2.0	memory length: 100000 epsilon: 0.9141095800018646
steps: 217	lr: 0.0001	evaluation reward: 1.61
episode: 774	score: 2.0	memory length: 100000 epsilon: 0.913677940001874
steps: 218	lr: 0.0001	evaluation reward: 1.62
episode: 775	score: 3.0	memory length: 100000 epsilon: 0.9131492800018854
steps: 267	lr: 0.0001	evaluation reward: 1.63
episode: 776	score: 4.0	memory length: 100000 epsilon: 0.9126008200018974
steps: 277	lr: 0.0001	evaluation reward: 1.65
episode: 777	score: 0.0	memory length: 100000 epsilon: 0.9123572800019026
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 778	score: 1.0	memory length: 100000 epsilon: 0.9120583000019091
steps: 151	lr: 0.0001	evaluation reward: 1.57
episode: 779	score: 7.0	memory length: 100000 epsilon: 0.9112742200019261
steps: 396	lr: 0.0001	evaluation reward: 1.63
episode: 780	score: 2.0	memory length: 100000 epsilon: 0.9108445600019355
steps: 217	lr: 0.0001	evaluation reward: 1.65
episode: 781	score: 2.0	memory length: 100000 epsilon: 0.910452520001944
steps: 198	lr: 0.0001	evaluation reward: 1.67
episode: 782	score: 3.0	memory length: 100000 epsilon: 0.9099654400019546
steps: 246	lr: 0.0001	evaluation reward: 1.69
episode: 783	score: 8.0	memory length: 100000 epsilon: 0.9093140200019687
steps: 329	lr: 0.0001	evaluation reward: 1.75
episode: 784	score: 3.0	memory length: 100000 epsilon: 0.9088249600019793
steps: 247	lr: 0.0001	evaluation reward: 1.78
episode: 785	score: 3.0	memory length: 100000 epsilon: 0.908377480001989
steps: 226	lr: 0.0001	evaluation reward: 1.78
episode: 786	score: 1.0	memory length: 100000 epsilon: 0.9080408800019963
steps: 170	lr: 0.0001	evaluation reward: 1.77
episode: 787	score: 3.0	memory length: 100000 epsilon: 0.9075835000020063

steps: 231	lr: 0.0001	evaluation reward: 1.79
episode: 788	score: 2.0	memory length: 100000 epsilon: 0.9071894800020148
steps: 199	lr: 0.0001	evaluation reward: 1.79
episode: 789	score: 1.0	memory length: 100000 epsilon: 0.9068905000020213
steps: 151	lr: 0.0001	evaluation reward: 1.79
episode: 790	score: 0.0	memory length: 100000 epsilon: 0.9066469600020266
steps: 123	lr: 0.0001	evaluation reward: 1.75
episode: 791	score: 0.0	memory length: 100000 epsilon: 0.9064034200020319
steps: 123	lr: 0.0001	evaluation reward: 1.72
episode: 792	score: 2.0	memory length: 100000 epsilon: 0.9060074200020405
steps: 200	lr: 0.0001	evaluation reward: 1.7
episode: 793	score: 4.0	memory length: 100000 epsilon: 0.9054569800020524
steps: 278	lr: 0.0001	evaluation reward: 1.73
episode: 794	score: 3.0	memory length: 100000 epsilon: 0.9049758400020629
steps: 243	lr: 0.0001	evaluation reward: 1.76
episode: 795	score: 3.0	memory length: 100000 epsilon: 0.9044491600020743
steps: 266	lr: 0.0001	evaluation reward: 1.78
episode: 796	score: 3.0	memory length: 100000 epsilon: 0.9039680200020848
steps: 243	lr: 0.0001	evaluation reward: 1.74
episode: 797	score: 1.0	memory length: 100000 epsilon: 0.9036274600020922
steps: 172	lr: 0.0001	evaluation reward: 1.71
episode: 798	score: 4.0	memory length: 100000 epsilon: 0.9030097000021056
steps: 312	lr: 0.0001	evaluation reward: 1.74
episode: 799	score: 0.0	memory length: 100000 epsilon: 0.9027661600021109
steps: 123	lr: 0.0001	evaluation reward: 1.73
episode: 800	score: 4.0	memory length: 100000 epsilon: 0.9022216600021227
steps: 275	lr: 0.0001	evaluation reward: 1.75
episode: 801	score: 4.0	memory length: 100000 epsilon: 0.9017088400021338
steps: 259	lr: 0.0001	evaluation reward: 1.78
episode: 802	score: 0.0	memory length: 100000 epsilon: 0.9014653000021391
steps: 123	lr: 0.0001	evaluation reward: 1.77
episode: 803	score: 3.0	memory length: 100000 epsilon: 0.9009287200021507
steps: 271	lr: 0.0001	evaluation reward: 1.79
episode: 804	score: 4.0	memory length: 100000 epsilon: 0.9003822400021626
steps: 276	lr: 0.0001	evaluation reward: 1.83
episode: 805	score: 2.0	memory length: 100000 epsilon: 0.8999921800021711
steps: 197	lr: 0.0001	evaluation reward: 1.83
episode: 806	score: 3.0	memory length: 100000 epsilon: 0.8995387600021809
steps: 229	lr: 0.0001	evaluation reward: 1.84
episode: 807	score: 2.0	memory length: 100000 epsilon: 0.8991823600021887
steps: 180	lr: 0.0001	evaluation reward: 1.84
episode: 808	score: 0.0	memory length: 100000 epsilon: 0.8989388200021939
steps: 123	lr: 0.0001	evaluation reward: 1.84
episode: 809	score: 0.0	memory length: 100000 epsilon: 0.8986952800021992
steps: 123	lr: 0.0001	evaluation reward: 1.82
episode: 810	score: 3.0	memory length: 100000 epsilon: 0.8982101800022098
steps: 245	lr: 0.0001	evaluation reward: 1.85
episode: 811	score: 3.0	memory length: 100000 epsilon: 0.8977270600022202

steps: 244	lr: 0.0001	evaluation reward: 1.84
episode: 812	score: 2.0	memory length: 100000 epsilon: 0.8973310600022288
steps: 200	lr: 0.0001	evaluation reward: 1.86
episode: 813	score: 3.0	memory length: 100000 epsilon: 0.8968024000022403
steps: 267	lr: 0.0001	evaluation reward: 1.89
episode: 814	score: 0.0	memory length: 100000 epsilon: 0.8965588600022456
steps: 123	lr: 0.0001	evaluation reward: 1.89
episode: 815	score: 2.0	memory length: 100000 epsilon: 0.8961292000022549
steps: 217	lr: 0.0001	evaluation reward: 1.91
episode: 816	score: 1.0	memory length: 100000 epsilon: 0.8957906200022623
steps: 171	lr: 0.0001	evaluation reward: 1.91
episode: 817	score: 0.0	memory length: 100000 epsilon: 0.8955470800022676
steps: 123	lr: 0.0001	evaluation reward: 1.91
episode: 818	score: 3.0	memory length: 100000 epsilon: 0.8950580200022782
steps: 247	lr: 0.0001	evaluation reward: 1.91
episode: 819	score: 3.0	memory length: 100000 epsilon: 0.8946105400022879
steps: 226	lr: 0.0001	evaluation reward: 1.93
episode: 820	score: 2.0	memory length: 100000 epsilon: 0.8941789000022973
steps: 218	lr: 0.0001	evaluation reward: 1.94
episode: 821	score: 2.0	memory length: 100000 epsilon: 0.8937512200023066
steps: 216	lr: 0.0001	evaluation reward: 1.93
episode: 822	score: 5.0	memory length: 100000 epsilon: 0.8930740600023213
steps: 342	lr: 0.0001	evaluation reward: 1.98
episode: 823	score: 1.0	memory length: 100000 epsilon: 0.8927750800023277
steps: 151	lr: 0.0001	evaluation reward: 1.96
episode: 824	score: 2.0	memory length: 100000 epsilon: 0.8923850200023362
steps: 197	lr: 0.0001	evaluation reward: 1.96
episode: 825	score: 3.0	memory length: 100000 epsilon: 0.891933580002346
steps: 228	lr: 0.0001	evaluation reward: 1.97
episode: 826	score: 2.0	memory length: 100000 epsilon: 0.8915019400023554
steps: 218	lr: 0.0001	evaluation reward: 1.99
episode: 827	score: 3.0	memory length: 100000 epsilon: 0.8909752600023668
steps: 266	lr: 0.0001	evaluation reward: 2.01
episode: 828	score: 1.0	memory length: 100000 epsilon: 0.8906406400023741
steps: 169	lr: 0.0001	evaluation reward: 2.0
episode: 829	score: 3.0	memory length: 100000 epsilon: 0.8901159400023855
steps: 265	lr: 0.0001	evaluation reward: 2.0
episode: 830	score: 3.0	memory length: 100000 epsilon: 0.8896605400023954
steps: 230	lr: 0.0001	evaluation reward: 2.01
episode: 831	score: 0.0	memory length: 100000 epsilon: 0.8894189800024006
steps: 122	lr: 0.0001	evaluation reward: 2.0
episode: 832	score: 1.0	memory length: 100000 epsilon: 0.8890843600024079
steps: 169	lr: 0.0001	evaluation reward: 1.99
episode: 833	score: 0.0	memory length: 100000 epsilon: 0.8888428000024131
steps: 122	lr: 0.0001	evaluation reward: 1.99
episode: 834	score: 4.0	memory length: 100000 epsilon: 0.8882983000024249
steps: 275	lr: 0.0001	evaluation reward: 2.02
episode: 835	score: 0.0	memory length: 100000 epsilon: 0.8880567400024302

steps: 122	lr: 0.0001	evaluation reward: 2.02
episode: 836	score: 2.0	memory length: 100000 epsilon: 0.8876587600024388
steps: 201	lr: 0.0001	evaluation reward: 2.01
episode: 837	score: 5.0	memory length: 100000 epsilon: 0.8869776400024536
steps: 344	lr: 0.0001	evaluation reward: 2.03
episode: 838	score: 2.0	memory length: 100000 epsilon: 0.8865875800024621
steps: 197	lr: 0.0001	evaluation reward: 2.04
episode: 839	score: 3.0	memory length: 100000 epsilon: 0.8860985200024727
steps: 247	lr: 0.0001	evaluation reward: 2.05
episode: 840	score: 4.0	memory length: 100000 epsilon: 0.8855500600024846
steps: 277	lr: 0.0001	evaluation reward: 2.08
episode: 841	score: 0.0	memory length: 100000 epsilon: 0.8853065200024899
steps: 123	lr: 0.0001	evaluation reward: 2.07
episode: 842	score: 4.0	memory length: 100000 epsilon: 0.8846848000025034
steps: 314	lr: 0.0001	evaluation reward: 2.1
episode: 843	score: 2.0	memory length: 100000 epsilon: 0.8843264200025112
steps: 181	lr: 0.0001	evaluation reward: 2.11
episode: 844	score: 1.0	memory length: 100000 epsilon: 0.8839937800025184
steps: 168	lr: 0.0001	evaluation reward: 2.11
episode: 845	score: 0.0	memory length: 100000 epsilon: 0.8837522200025236
steps: 122	lr: 0.0001	evaluation reward: 2.09
episode: 846	score: 3.0	memory length: 100000 epsilon: 0.8833007800025334
steps: 228	lr: 0.0001	evaluation reward: 2.12
episode: 847	score: 3.0	memory length: 100000 epsilon: 0.882811720002544
steps: 247	lr: 0.0001	evaluation reward: 2.15
episode: 848	score: 2.0	memory length: 100000 epsilon: 0.8824553200025518
steps: 180	lr: 0.0001	evaluation reward: 2.14
episode: 849	score: 1.0	memory length: 100000 epsilon: 0.8821147600025592
steps: 172	lr: 0.0001	evaluation reward: 2.14
episode: 850	score: 0.0	memory length: 100000 epsilon: 0.8818712200025645
steps: 123	lr: 0.0001	evaluation reward: 2.14
episode: 851	score: 1.0	memory length: 100000 epsilon: 0.881572240002571
steps: 151	lr: 0.0001	evaluation reward: 2.12
episode: 852	score: 5.0	memory length: 100000 epsilon: 0.880924780002585
steps: 327	lr: 0.0001	evaluation reward: 2.14
episode: 853	score: 4.0	memory length: 100000 epsilon: 0.8803387000025977
steps: 296	lr: 0.0001	evaluation reward: 2.14
episode: 854	score: 5.0	memory length: 100000 epsilon: 0.8796655000026123
steps: 340	lr: 0.0001	evaluation reward: 2.19
episode: 855	score: 11.0	memory length: 100000 epsilon: 0.8788794400026294
steps: 397	lr: 0.0001	evaluation reward: 2.28
episode: 856	score: 4.0	memory length: 100000 epsilon: 0.8782913800026422
steps: 297	lr: 0.0001	evaluation reward: 2.29
episode: 857	score: 4.0	memory length: 100000 epsilon: 0.877748860002654
steps: 274	lr: 0.0001	evaluation reward: 2.31
episode: 858	score: 4.0	memory length: 100000 epsilon: 0.8771588200026668
steps: 298	lr: 0.0001	evaluation reward: 2.3
episode: 859	score: 0.0	memory length: 100000 epsilon: 0.876917260002672

steps: 122	lr: 0.0001	evaluation reward: 2.27
episode: 860	score: 0.0	memory length: 100000 epsilon: 0.8766737200026773
steps: 123	lr: 0.0001	evaluation reward: 2.26
episode: 861	score: 2.0	memory length: 100000 epsilon: 0.8762816800026858
steps: 198	lr: 0.0001	evaluation reward: 2.26
episode: 862	score: 4.0	memory length: 100000 epsilon: 0.8757431200026975
steps: 272	lr: 0.0001	evaluation reward: 2.27
episode: 863	score: 1.0	memory length: 100000 epsilon: 0.8754025600027049
steps: 172	lr: 0.0001	evaluation reward: 2.26
episode: 864	score: 4.0	memory length: 100000 epsilon: 0.8748105400027177
steps: 299	lr: 0.0001	evaluation reward: 2.28
episode: 865	score: 0.0	memory length: 100000 epsilon: 0.874568980002723
steps: 122	lr: 0.0001	evaluation reward: 2.27
episode: 866	score: 1.0	memory length: 100000 epsilon: 0.8742323800027303
steps: 170	lr: 0.0001	evaluation reward: 2.28
episode: 867	score: 2.0	memory length: 100000 epsilon: 0.8738403400027388
steps: 198	lr: 0.0001	evaluation reward: 2.29
episode: 868	score: 1.0	memory length: 100000 epsilon: 0.873507700002746
steps: 168	lr: 0.0001	evaluation reward: 2.28
episode: 869	score: 2.0	memory length: 100000 epsilon: 0.8731176400027545
steps: 197	lr: 0.0001	evaluation reward: 2.3
episode: 870	score: 1.0	memory length: 100000 epsilon: 0.872818660002761
steps: 151	lr: 0.0001	evaluation reward: 2.29
episode: 871	score: 1.0	memory length: 100000 epsilon: 0.8724820600027683
steps: 170	lr: 0.0001	evaluation reward: 2.3
episode: 872	score: 2.0	memory length: 100000 epsilon: 0.8720920000027768
steps: 197	lr: 0.0001	evaluation reward: 2.32
episode: 873	score: 2.0	memory length: 100000 epsilon: 0.8716524400027863
steps: 222	lr: 0.0001	evaluation reward: 2.32
episode: 874	score: 5.0	memory length: 100000 epsilon: 0.8710762600027988
steps: 291	lr: 0.0001	evaluation reward: 2.35
episode: 875	score: 5.0	memory length: 100000 epsilon: 0.8704981000028114
steps: 292	lr: 0.0001	evaluation reward: 2.37
episode: 876	score: 0.0	memory length: 100000 epsilon: 0.8702545600028166
steps: 123	lr: 0.0001	evaluation reward: 2.33
episode: 877	score: 3.0	memory length: 100000 epsilon: 0.8698070800028264
steps: 226	lr: 0.0001	evaluation reward: 2.36
episode: 878	score: 3.0	memory length: 100000 epsilon: 0.869318020002837
steps: 247	lr: 0.0001	evaluation reward: 2.38
episode: 879	score: 3.0	memory length: 100000 epsilon: 0.8688289600028476
steps: 247	lr: 0.0001	evaluation reward: 2.34
episode: 880	score: 3.0	memory length: 100000 epsilon: 0.8683399000028582
steps: 247	lr: 0.0001	evaluation reward: 2.35
episode: 881	score: 2.0	memory length: 100000 epsilon: 0.8679062800028676
steps: 219	lr: 0.0001	evaluation reward: 2.35
episode: 882	score: 3.0	memory length: 100000 epsilon: 0.8673796000028791
steps: 266	lr: 0.0001	evaluation reward: 2.35
episode: 883	score: 2.0	memory length: 100000 epsilon: 0.8669816200028877

steps: 201	lr: 0.0001	evaluation reward: 2.29
episode: 884	score: 0.0	memory length: 100000 epsilon: 0.866738080002893
steps: 123	lr: 0.0001	evaluation reward: 2.26
episode: 885	score: 2.0	memory length: 100000 epsilon: 0.8663420800029016
steps: 200	lr: 0.0001	evaluation reward: 2.25
episode: 886	score: 4.0	memory length: 100000 epsilon: 0.8657282800029149
steps: 310	lr: 0.0001	evaluation reward: 2.28
episode: 887	score: 2.0	memory length: 100000 epsilon: 0.8653303000029235
steps: 201	lr: 0.0001	evaluation reward: 2.27
episode: 888	score: 2.0	memory length: 100000 epsilon: 0.864940240002932
steps: 197	lr: 0.0001	evaluation reward: 2.27
episode: 889	score: 1.0	memory length: 100000 epsilon: 0.8646412600029385
steps: 151	lr: 0.0001	evaluation reward: 2.27
episode: 890	score: 2.0	memory length: 100000 epsilon: 0.864249220002947
steps: 198	lr: 0.0001	evaluation reward: 2.29
episode: 891	score: 2.0	memory length: 100000 epsilon: 0.8638195600029563
steps: 217	lr: 0.0001	evaluation reward: 2.31
episode: 892	score: 2.0	memory length: 100000 epsilon: 0.8634295000029648
steps: 197	lr: 0.0001	evaluation reward: 2.31
episode: 893	score: 2.0	memory length: 100000 epsilon: 0.8630374600029733
steps: 198	lr: 0.0001	evaluation reward: 2.29
episode: 894	score: 3.0	memory length: 100000 epsilon: 0.8625167200029846
steps: 263	lr: 0.0001	evaluation reward: 2.29
episode: 895	score: 6.0	memory length: 100000 epsilon: 0.861898960002998
steps: 312	lr: 0.0001	evaluation reward: 2.32
episode: 896	score: 4.0	memory length: 100000 epsilon: 0.8612752600030116
steps: 315	lr: 0.0001	evaluation reward: 2.33
episode: 897	score: 0.0	memory length: 100000 epsilon: 0.8610317200030169
steps: 123	lr: 0.0001	evaluation reward: 2.32
episode: 898	score: 1.0	memory length: 100000 epsilon: 0.8607347200030233
steps: 150	lr: 0.0001	evaluation reward: 2.29
episode: 899	score: 4.0	memory length: 100000 epsilon: 0.8601882400030352
steps: 276	lr: 0.0001	evaluation reward: 2.33
episode: 900	score: 2.0	memory length: 100000 epsilon: 0.8597962000030437
steps: 198	lr: 0.0001	evaluation reward: 2.31
episode: 901	score: 1.0	memory length: 100000 epsilon: 0.8594972200030502
steps: 151	lr: 0.0001	evaluation reward: 2.28
episode: 902	score: 0.0	memory length: 100000 epsilon: 0.8592536800030555
steps: 123	lr: 0.0001	evaluation reward: 2.28
episode: 903	score: 2.0	memory length: 100000 epsilon: 0.8588260000030647
steps: 216	lr: 0.0001	evaluation reward: 2.27
episode: 904	score: 1.0	memory length: 100000 epsilon: 0.8585290000030712
steps: 150	lr: 0.0001	evaluation reward: 2.24
episode: 905	score: 6.0	memory length: 100000 epsilon: 0.8578538200030859
steps: 341	lr: 0.0001	evaluation reward: 2.28
episode: 906	score: 2.0	memory length: 100000 epsilon: 0.8574617800030944
steps: 198	lr: 0.0001	evaluation reward: 2.27
episode: 907	score: 4.0	memory length: 100000 epsilon: 0.8569133200031063

steps: 277	lr: 0.0001	evaluation reward: 2.29
episode: 908	score: 2.0	memory length: 100000 epsilon: 0.8565232600031147
steps: 197	lr: 0.0001	evaluation reward: 2.31
episode: 909	score: 6.0	memory length: 100000 epsilon: 0.85582234000313
steps: 354	lr: 0.0001	evaluation reward: 2.37
episode: 910	score: 2.0	memory length: 100000 epsilon: 0.8554303000031385
steps: 198	lr: 0.0001	evaluation reward: 2.36
episode: 911	score: 4.0	memory length: 100000 epsilon: 0.8549155000031496
steps: 260	lr: 0.0001	evaluation reward: 2.37
episode: 912	score: 2.0	memory length: 100000 epsilon: 0.8545254400031581
steps: 197	lr: 0.0001	evaluation reward: 2.37
episode: 913	score: 1.0	memory length: 100000 epsilon: 0.8542284400031646
steps: 150	lr: 0.0001	evaluation reward: 2.35
episode: 914	score: 7.0	memory length: 100000 epsilon: 0.8534107000031823
steps: 413	lr: 0.0001	evaluation reward: 2.42
episode: 915	score: 4.0	memory length: 100000 epsilon: 0.852824620003195
steps: 296	lr: 0.0001	evaluation reward: 2.44
episode: 916	score: 3.0	memory length: 100000 epsilon: 0.8523296200032058
steps: 250	lr: 0.0001	evaluation reward: 2.46
episode: 917	score: 1.0	memory length: 100000 epsilon: 0.8520306400032123
steps: 151	lr: 0.0001	evaluation reward: 2.47
episode: 918	score: 5.0	memory length: 100000 epsilon: 0.8513435800032272
steps: 347	lr: 0.0001	evaluation reward: 2.49
episode: 919	score: 0.0	memory length: 100000 epsilon: 0.8511000400032325
steps: 123	lr: 0.0001	evaluation reward: 2.46
episode: 920	score: 0.0	memory length: 100000 epsilon: 0.8508565000032378
steps: 123	lr: 0.0001	evaluation reward: 2.44
episode: 921	score: 4.0	memory length: 100000 epsilon: 0.8503021000032498
steps: 280	lr: 0.0001	evaluation reward: 2.46
episode: 922	score: 5.0	memory length: 100000 epsilon: 0.8495853400032654
steps: 362	lr: 0.0001	evaluation reward: 2.46
episode: 923	score: 4.0	memory length: 100000 epsilon: 0.8489933200032782
steps: 299	lr: 0.0001	evaluation reward: 2.49
episode: 924	score: 0.0	memory length: 100000 epsilon: 0.8487497800032835
steps: 123	lr: 0.0001	evaluation reward: 2.47
episode: 925	score: 0.0	memory length: 100000 epsilon: 0.8485082200032887
steps: 122	lr: 0.0001	evaluation reward: 2.44
episode: 926	score: 1.0	memory length: 100000 epsilon: 0.848173600003296
steps: 169	lr: 0.0001	evaluation reward: 2.43
episode: 927	score: 4.0	memory length: 100000 epsilon: 0.8475875200033087
steps: 296	lr: 0.0001	evaluation reward: 2.44
episode: 928	score: 3.0	memory length: 100000 epsilon: 0.8470964800033194
steps: 248	lr: 0.0001	evaluation reward: 2.46
episode: 929	score: 1.0	memory length: 100000 epsilon: 0.8467975000033259
steps: 151	lr: 0.0001	evaluation reward: 2.44
episode: 930	score: 5.0	memory length: 100000 epsilon: 0.8461955800033389
steps: 304	lr: 0.0001	evaluation reward: 2.46
episode: 931	score: 2.0	memory length: 100000 epsilon: 0.8457619600033484

steps: 219	lr: 0.0001	evaluation reward: 2.48
episode: 932	score: 5.0	memory length: 100000 epsilon: 0.8451164800033624
steps: 326	lr: 0.0001	evaluation reward: 2.52
episode: 933	score: 3.0	memory length: 100000 epsilon: 0.8446353400033728
steps: 243	lr: 0.0001	evaluation reward: 2.55
episode: 934	score: 2.0	memory length: 100000 epsilon: 0.8442730000033807
steps: 183	lr: 0.0001	evaluation reward: 2.53
episode: 935	score: 1.0	memory length: 100000 epsilon: 0.8439403600033879
steps: 168	lr: 0.0001	evaluation reward: 2.54
episode: 936	score: 5.0	memory length: 100000 epsilon: 0.8432533000034028
steps: 347	lr: 0.0001	evaluation reward: 2.57
episode: 937	score: 4.0	memory length: 100000 epsilon: 0.8427741400034132
steps: 242	lr: 0.0001	evaluation reward: 2.56
episode: 938	score: 5.0	memory length: 100000 epsilon: 0.8421326200034271
steps: 324	lr: 0.0001	evaluation reward: 2.59
episode: 939	score: 8.0	memory length: 100000 epsilon: 0.8412554800034462
steps: 443	lr: 0.0001	evaluation reward: 2.64
episode: 940	score: 3.0	memory length: 100000 epsilon: 0.8407941400034562
steps: 233	lr: 0.0001	evaluation reward: 2.63
episode: 941	score: 1.0	memory length: 100000 epsilon: 0.8404971400034627
steps: 150	lr: 0.0001	evaluation reward: 2.64
episode: 942	score: 2.0	memory length: 100000 epsilon: 0.8401387600034704
steps: 181	lr: 0.0001	evaluation reward: 2.62
episode: 943	score: 2.0	memory length: 100000 epsilon: 0.8397091000034798
steps: 217	lr: 0.0001	evaluation reward: 2.62
episode: 944	score: 6.0	memory length: 100000 epsilon: 0.8390200600034947
steps: 348	lr: 0.0001	evaluation reward: 2.67
episode: 945	score: 4.0	memory length: 100000 epsilon: 0.8384676400035067
steps: 279	lr: 0.0001	evaluation reward: 2.71
episode: 946	score: 1.0	memory length: 100000 epsilon: 0.8381706400035132
steps: 150	lr: 0.0001	evaluation reward: 2.69
episode: 947	score: 1.0	memory length: 100000 epsilon: 0.8378360200035204
steps: 169	lr: 0.0001	evaluation reward: 2.67
episode: 948	score: 6.0	memory length: 100000 epsilon: 0.8371469800035354
steps: 348	lr: 0.0001	evaluation reward: 2.71
episode: 949	score: 0.0	memory length: 100000 epsilon: 0.8369034400035407
steps: 123	lr: 0.0001	evaluation reward: 2.7
episode: 950	score: 6.0	memory length: 100000 epsilon: 0.8362322200035552
steps: 339	lr: 0.0001	evaluation reward: 2.76
episode: 951	score: 1.0	memory length: 100000 epsilon: 0.8358916600035626
steps: 172	lr: 0.0001	evaluation reward: 2.76
episode: 952	score: 3.0	memory length: 100000 epsilon: 0.8354382400035725
steps: 229	lr: 0.0001	evaluation reward: 2.74
episode: 953	score: 3.0	memory length: 100000 epsilon: 0.8349472000035831
steps: 248	lr: 0.0001	evaluation reward: 2.73
episode: 954	score: 1.0	memory length: 100000 epsilon: 0.8346482200035896
steps: 151	lr: 0.0001	evaluation reward: 2.69
episode: 955	score: 2.0	memory length: 100000 epsilon: 0.8342542000035982

steps: 199	lr: 0.0001	evaluation reward: 2.6
episode: 956	score: 3.0	memory length: 100000 epsilon: 0.8337631600036088
steps: 248	lr: 0.0001	evaluation reward: 2.59
episode: 957	score: 1.0	memory length: 100000 epsilon: 0.8334641800036153
steps: 151	lr: 0.0001	evaluation reward: 2.56
episode: 958	score: 4.0	memory length: 100000 epsilon: 0.8329177000036272
steps: 276	lr: 0.0001	evaluation reward: 2.56
episode: 959	score: 1.0	memory length: 100000 epsilon: 0.8326187200036337
steps: 151	lr: 0.0001	evaluation reward: 2.57
episode: 960	score: 4.0	memory length: 100000 epsilon: 0.8320227400036466
steps: 301	lr: 0.0001	evaluation reward: 2.61
episode: 961	score: 0.0	memory length: 100000 epsilon: 0.8317792000036519
steps: 123	lr: 0.0001	evaluation reward: 2.59
episode: 962	score: 5.0	memory length: 100000 epsilon: 0.8311416400036657
steps: 322	lr: 0.0001	evaluation reward: 2.6
episode: 963	score: 4.0	memory length: 100000 epsilon: 0.8305615000036783
steps: 293	lr: 0.0001	evaluation reward: 2.63
episode: 964	score: 5.0	memory length: 100000 epsilon: 0.8299516600036916
steps: 308	lr: 0.0001	evaluation reward: 2.64
episode: 965	score: 1.0	memory length: 100000 epsilon: 0.8296526800036981
steps: 151	lr: 0.0001	evaluation reward: 2.65
episode: 966	score: 2.0	memory length: 100000 epsilon: 0.8292586600037066
steps: 199	lr: 0.0001	evaluation reward: 2.66
episode: 967	score: 2.0	memory length: 100000 epsilon: 0.8288666200037151
steps: 198	lr: 0.0001	evaluation reward: 2.66
episode: 968	score: 3.0	memory length: 100000 epsilon: 0.8283815200037257
steps: 245	lr: 0.0001	evaluation reward: 2.68
episode: 969	score: 0.0	memory length: 100000 epsilon: 0.828137980003731
steps: 123	lr: 0.0001	evaluation reward: 2.66
episode: 970	score: 2.0	memory length: 100000 epsilon: 0.8277459400037395
steps: 198	lr: 0.0001	evaluation reward: 2.67
episode: 971	score: 3.0	memory length: 100000 epsilon: 0.8272984600037492
steps: 226	lr: 0.0001	evaluation reward: 2.69
episode: 972	score: 7.0	memory length: 100000 epsilon: 0.826569820003765
steps: 368	lr: 0.0001	evaluation reward: 2.74
episode: 973	score: 4.0	memory length: 100000 epsilon: 0.8260273000037768
steps: 274	lr: 0.0001	evaluation reward: 2.76
episode: 974	score: 0.0	memory length: 100000 epsilon: 0.825785740003782
steps: 122	lr: 0.0001	evaluation reward: 2.71
episode: 975	score: 3.0	memory length: 100000 epsilon: 0.8253679600037911
steps: 211	lr: 0.0001	evaluation reward: 2.69
episode: 976	score: 6.0	memory length: 100000 epsilon: 0.8246195200038073
steps: 378	lr: 0.0001	evaluation reward: 2.75
episode: 977	score: 2.0	memory length: 100000 epsilon: 0.8241859000038168
steps: 219	lr: 0.0001	evaluation reward: 2.74
episode: 978	score: 3.0	memory length: 100000 epsilon: 0.8237245600038268
steps: 233	lr: 0.0001	evaluation reward: 2.74
episode: 979	score: 4.0	memory length: 100000 epsilon: 0.8232196600038377

steps: 255	lr: 0.0001	evaluation reward: 2.75
episode: 980	score: 5.0	memory length: 100000 epsilon: 0.822609820003851
steps: 308	lr: 0.0001	evaluation reward: 2.77
episode: 981	score: 4.0	memory length: 100000 epsilon: 0.8220217600038637
steps: 297	lr: 0.0001	evaluation reward: 2.79
episode: 982	score: 2.0	memory length: 100000 epsilon: 0.8216297200038722
steps: 198	lr: 0.0001	evaluation reward: 2.78
episode: 983	score: 2.0	memory length: 100000 epsilon: 0.8212337200038808
steps: 200	lr: 0.0001	evaluation reward: 2.78
episode: 984	score: 1.0	memory length: 100000 epsilon: 0.8208971200038881
steps: 170	lr: 0.0001	evaluation reward: 2.79
episode: 985	score: 1.0	memory length: 100000 epsilon: 0.8205625000038954
steps: 169	lr: 0.0001	evaluation reward: 2.78
episode: 986	score: 5.0	memory length: 100000 epsilon: 0.8199605800039085
steps: 304	lr: 0.0001	evaluation reward: 2.79
episode: 987	score: 1.0	memory length: 100000 epsilon: 0.819661600003915
steps: 151	lr: 0.0001	evaluation reward: 2.78
episode: 988	score: 0.0	memory length: 100000 epsilon: 0.8194200400039202
steps: 122	lr: 0.0001	evaluation reward: 2.76
episode: 989	score: 0.0	memory length: 100000 epsilon: 0.8191765000039255
steps: 123	lr: 0.0001	evaluation reward: 2.75
episode: 990	score: 3.0	memory length: 100000 epsilon: 0.8186894200039361
steps: 246	lr: 0.0001	evaluation reward: 2.76
episode: 991	score: 4.0	memory length: 100000 epsilon: 0.8181449200039479
steps: 275	lr: 0.0001	evaluation reward: 2.78
episode: 992	score: 1.0	memory length: 100000 epsilon: 0.8178103000039552
steps: 169	lr: 0.0001	evaluation reward: 2.77
episode: 993	score: 4.0	memory length: 100000 epsilon: 0.8172262000039678
steps: 295	lr: 0.0001	evaluation reward: 2.79
episode: 994	score: 2.0	memory length: 100000 epsilon: 0.8168302000039764
steps: 200	lr: 0.0001	evaluation reward: 2.78
episode: 995	score: 1.0	memory length: 100000 epsilon: 0.8164936000039837
steps: 170	lr: 0.0001	evaluation reward: 2.73
episode: 996	score: 4.0	memory length: 100000 epsilon: 0.8159134600039963
steps: 293	lr: 0.0001	evaluation reward: 2.73
episode: 997	score: 4.0	memory length: 100000 epsilon: 0.8153372800040088
steps: 291	lr: 0.0001	evaluation reward: 2.77
episode: 998	score: 1.0	memory length: 100000 epsilon: 0.8150383000040153
steps: 151	lr: 0.0001	evaluation reward: 2.77
episode: 999	score: 3.0	memory length: 100000 epsilon: 0.8145551800040258
steps: 244	lr: 0.0001	evaluation reward: 2.76
episode: 1000	score: 1.0	memory length: 100000 epsilon: 0.8142562000040323
steps: 151	lr: 0.0001	evaluation reward: 2.75
episode: 1001	score: 3.0	memory length: 100000 epsilon: 0.8138027800040422
steps: 229	lr: 0.0001	evaluation reward: 2.77
episode: 1002	score: 5.0	memory length: 100000 epsilon: 0.8131929400040554
steps: 308	lr: 0.0001	evaluation reward: 2.82
episode: 1003	score: 2.0	memory length: 100000 epsilon: 0.8127613000040648

steps: 218	lr: 0.0001	evaluation reward: 2.82	
episode: 1004	score: 2.0	memory length: 100000	epsilon: 0.8124009400040726
steps: 182	lr: 0.0001	evaluation reward: 2.83	
episode: 1005	score: 1.0	memory length: 100000	epsilon: 0.8121019600040791
steps: 151	lr: 0.0001	evaluation reward: 2.78	
episode: 1006	score: 3.0	memory length: 100000	epsilon: 0.8116505200040889
steps: 228	lr: 0.0001	evaluation reward: 2.79	
episode: 1007	score: 0.0	memory length: 100000	epsilon: 0.8114089600040941
steps: 122	lr: 0.0001	evaluation reward: 2.75	
episode: 1008	score: 4.0	memory length: 100000	epsilon: 0.8109040600041051
steps: 255	lr: 0.0001	evaluation reward: 2.77	
episode: 1009	score: 1.0	memory length: 100000	epsilon: 0.8105654800041124
steps: 171	lr: 0.0001	evaluation reward: 2.72	
episode: 1010	score: 6.0	memory length: 100000	epsilon: 0.8098883200041271
steps: 342	lr: 0.0001	evaluation reward: 2.76	
episode: 1011	score: 2.0	memory length: 100000	epsilon: 0.8094962800041356
steps: 198	lr: 0.0001	evaluation reward: 2.74	
episode: 1012	score: 1.0	memory length: 100000	epsilon: 0.8091973000041421
steps: 151	lr: 0.0001	evaluation reward: 2.73	
episode: 1013	score: 4.0	memory length: 100000	epsilon: 0.8087300200041523
steps: 236	lr: 0.0001	evaluation reward: 2.76	
episode: 1014	score: 4.0	memory length: 100000	epsilon: 0.8082211600041633
steps: 257	lr: 0.0001	evaluation reward: 2.73	
episode: 1015	score: 3.0	memory length: 100000	epsilon: 0.8077360600041739
steps: 245	lr: 0.0001	evaluation reward: 2.72	
episode: 1016	score: 1.0	memory length: 100000	epsilon: 0.8074370800041804
steps: 151	lr: 0.0001	evaluation reward: 2.7	
episode: 1017	score: 0.0	memory length: 100000	epsilon: 0.8071935400041856
steps: 123	lr: 0.0001	evaluation reward: 2.69	
episode: 1018	score: 4.0	memory length: 100000	epsilon: 0.8066074600041984
steps: 296	lr: 0.0001	evaluation reward: 2.68	
episode: 1019	score: 4.0	memory length: 100000	epsilon: 0.8059738600042121
steps: 320	lr: 0.0001	evaluation reward: 2.72	
episode: 1020	score: 6.0	memory length: 100000	epsilon: 0.8051838400042293
steps: 399	lr: 0.0001	evaluation reward: 2.78	
episode: 1021	score: 4.0	memory length: 100000	epsilon: 0.8046353800042412
steps: 277	lr: 0.0001	evaluation reward: 2.78	
episode: 1022	score: 3.0	memory length: 100000	epsilon: 0.804183940004251
steps: 228	lr: 0.0001	evaluation reward: 2.76	
episode: 1023	score: 0.0	memory length: 100000	epsilon: 0.8039404000042563
steps: 123	lr: 0.0001	evaluation reward: 2.72	
episode: 1024	score: 3.0	memory length: 100000	epsilon: 0.8034513400042669
steps: 247	lr: 0.0001	evaluation reward: 2.75	
episode: 1025	score: 3.0	memory length: 100000	epsilon: 0.8030256400042761
steps: 215	lr: 0.0001	evaluation reward: 2.78	
episode: 1026	score: 8.0	memory length: 100000	epsilon: 0.8022336400042933
steps: 400	lr: 0.0001	evaluation reward: 2.85	
episode: 1027	score: 5.0	memory length: 100000	epsilon: 0.8015604400043079

steps: 340	lr: 4e-05	evaluation reward: 2.86	
episode: 1028	score: 5.0	memory length: 100000	epsilon: 0.8008852600043226
steps: 341	lr: 4e-05	evaluation reward: 2.88	
episode: 1029	score: 3.0	memory length: 100000	epsilon: 0.8004338200043324
steps: 228	lr: 4e-05	evaluation reward: 2.9	
episode: 1030	score: 4.0	memory length: 100000	epsilon: 0.7998556600043449
steps: 292	lr: 4e-05	evaluation reward: 2.89	
episode: 1031	score: 2.0	memory length: 100000	epsilon: 0.7994972800043527
steps: 181	lr: 4e-05	evaluation reward: 2.89	
episode: 1032	score: 8.0	memory length: 100000	epsilon: 0.7986676600043707
steps: 419	lr: 4e-05	evaluation reward: 2.92	
episode: 1033	score: 1.0	memory length: 100000	epsilon: 0.7983290800043781
steps: 171	lr: 4e-05	evaluation reward: 2.9	
episode: 1034	score: 1.0	memory length: 100000	epsilon: 0.7980301000043846
steps: 151	lr: 4e-05	evaluation reward: 2.89	
episode: 1035	score: 3.0	memory length: 100000	epsilon: 0.7976182600043935
steps: 208	lr: 4e-05	evaluation reward: 2.91	
episode: 1036	score: 2.0	memory length: 100000	epsilon: 0.7971866200044029
steps: 218	lr: 4e-05	evaluation reward: 2.88	
episode: 1037	score: 1.0	memory length: 100000	epsilon: 0.7968539800044101
steps: 168	lr: 4e-05	evaluation reward: 2.85	
episode: 1038	score: 2.0	memory length: 100000	epsilon: 0.7964619400044186
steps: 198	lr: 4e-05	evaluation reward: 2.82	
episode: 1039	score: 2.0	memory length: 100000	epsilon: 0.7961035600044264
steps: 181	lr: 4e-05	evaluation reward: 2.76	
episode: 1040	score: 3.0	memory length: 100000	epsilon: 0.7956046000044372
steps: 252	lr: 4e-05	evaluation reward: 2.76	
episode: 1041	score: 2.0	memory length: 100000	epsilon: 0.7951729600044466
steps: 218	lr: 4e-05	evaluation reward: 2.77	
episode: 1042	score: 3.0	memory length: 100000	epsilon: 0.7947571600044556
steps: 210	lr: 4e-05	evaluation reward: 2.78	
episode: 1043	score: 3.0	memory length: 100000	epsilon: 0.7942601800044664
steps: 251	lr: 4e-05	evaluation reward: 2.79	
episode: 1044	score: 5.0	memory length: 100000	epsilon: 0.7936661800044793
steps: 300	lr: 4e-05	evaluation reward: 2.78	
episode: 1045	score: 4.0	memory length: 100000	epsilon: 0.7930860400044919
steps: 293	lr: 4e-05	evaluation reward: 2.78	
episode: 1046	score: 3.0	memory length: 100000	epsilon: 0.7925890600045027
steps: 251	lr: 4e-05	evaluation reward: 2.8	
episode: 1047	score: 3.0	memory length: 100000	epsilon: 0.7921000000045133
steps: 247	lr: 4e-05	evaluation reward: 2.82	
episode: 1048	score: 7.0	memory length: 100000	epsilon: 0.7912585000045316
steps: 425	lr: 4e-05	evaluation reward: 2.83	
episode: 1049	score: 2.0	memory length: 100000	epsilon: 0.7908644800045401
steps: 199	lr: 4e-05	evaluation reward: 2.85	
episode: 1050	score: 4.0	memory length: 100000	epsilon: 0.790361560004551
steps: 254	lr: 4e-05	evaluation reward: 2.83	
episode: 1051	score: 1.0	memory length: 100000	epsilon: 0.7900229800045584

steps: 171	lr: 4e-05	evaluation reward: 2.83	
episode: 1052	score: 6.0	memory length: 100000	epsilon: 0.7892864200045744
steps: 372	lr: 4e-05	evaluation reward: 2.86	
episode: 1053	score: 5.0	memory length: 100000	epsilon: 0.7885993600045893
steps: 347	lr: 4e-05	evaluation reward: 2.88	
episode: 1054	score: 6.0	memory length: 100000	epsilon: 0.7878944800046046
steps: 356	lr: 4e-05	evaluation reward: 2.93	
episode: 1055	score: 0.0	memory length: 100000	epsilon: 0.7876509400046099
steps: 123	lr: 4e-05	evaluation reward: 2.91	
episode: 1056	score: 1.0	memory length: 100000	epsilon: 0.7873123600046172
steps: 171	lr: 4e-05	evaluation reward: 2.89	
episode: 1057	score: 4.0	memory length: 100000	epsilon: 0.78672430000463
steps: 297	lr: 4e-05	evaluation reward: 2.92	
episode: 1058	score: 4.0	memory length: 100000	epsilon: 0.7861758400046419
steps: 277	lr: 4e-05	evaluation reward: 2.92	
episode: 1059	score: 3.0	memory length: 100000	epsilon: 0.7856848000046526
steps: 248	lr: 4e-05	evaluation reward: 2.94	
episode: 1060	score: 5.0	memory length: 100000	epsilon: 0.7850908000046655
steps: 300	lr: 4e-05	evaluation reward: 2.95	
episode: 1061	score: 5.0	memory length: 100000	epsilon: 0.7844393800046796
steps: 329	lr: 4e-05	evaluation reward: 3.0	
episode: 1062	score: 4.0	memory length: 100000	epsilon: 0.7838968600046914
steps: 274	lr: 4e-05	evaluation reward: 2.99	
episode: 1063	score: 1.0	memory length: 100000	epsilon: 0.7835978800046979
steps: 151	lr: 4e-05	evaluation reward: 2.96	
episode: 1064	score: 4.0	memory length: 100000	epsilon: 0.7829702200047115
steps: 317	lr: 4e-05	evaluation reward: 2.95	
episode: 1065	score: 0.0	memory length: 100000	epsilon: 0.7827266800047168
steps: 123	lr: 4e-05	evaluation reward: 2.94	
episode: 1066	score: 2.0	memory length: 100000	epsilon: 0.7823346400047253
steps: 198	lr: 4e-05	evaluation reward: 2.94	
episode: 1067	score: 2.0	memory length: 100000	epsilon: 0.7819030000047347
steps: 218	lr: 4e-05	evaluation reward: 2.94	
episode: 1068	score: 2.0	memory length: 100000	epsilon: 0.7815466000047424
steps: 180	lr: 4e-05	evaluation reward: 2.93	
episode: 1069	score: 3.0	memory length: 100000	epsilon: 0.7810634800047529
steps: 244	lr: 4e-05	evaluation reward: 2.96	
episode: 1070	score: 4.0	memory length: 100000	epsilon: 0.7805823400047633
steps: 243	lr: 4e-05	evaluation reward: 2.98	
episode: 1071	score: 1.0	memory length: 100000	epsilon: 0.7802853400047698
steps: 150	lr: 4e-05	evaluation reward: 2.96	
episode: 1072	score: 4.0	memory length: 100000	epsilon: 0.7797784600047808
steps: 256	lr: 4e-05	evaluation reward: 2.93	
episode: 1073	score: 3.0	memory length: 100000	epsilon: 0.7792894000047914
steps: 247	lr: 4e-05	evaluation reward: 2.92	
episode: 1074	score: 4.0	memory length: 100000	epsilon: 0.7787825200048024
steps: 256	lr: 4e-05	evaluation reward: 2.96	
episode: 1075	score: 1.0	memory length: 100000	epsilon: 0.7784459200048097

steps: 170	lr: 4e-05	evaluation reward: 2.94	
episode: 1076	score: 8.0	memory length: 100000	epsilon: 0.7776321400048274
steps: 411	lr: 4e-05	evaluation reward: 2.96	
episode: 1077	score: 1.0	memory length: 100000	epsilon: 0.7773351400048338
steps: 150	lr: 4e-05	evaluation reward: 2.95	
episode: 1078	score: 3.0	memory length: 100000	epsilon: 0.7768896400048435
steps: 225	lr: 4e-05	evaluation reward: 2.95	
episode: 1079	score: 1.0	memory length: 100000	epsilon: 0.7765530400048508
steps: 170	lr: 4e-05	evaluation reward: 2.92	
episode: 1080	score: 3.0	memory length: 100000	epsilon: 0.7761035800048606
steps: 227	lr: 4e-05	evaluation reward: 2.9	
episode: 1081	score: 1.0	memory length: 100000	epsilon: 0.775806580004867
steps: 150	lr: 4e-05	evaluation reward: 2.87	
episode: 1082	score: 5.0	memory length: 100000	epsilon: 0.77521060000488
steps: 301	lr: 4e-05	evaluation reward: 2.9	
episode: 1083	score: 2.0	memory length: 100000	epsilon: 0.7747809400048893
steps: 217	lr: 4e-05	evaluation reward: 2.9	
episode: 1084	score: 0.0	memory length: 100000	epsilon: 0.7745393800048945
steps: 122	lr: 4e-05	evaluation reward: 2.89	
episode: 1085	score: 3.0	memory length: 100000	epsilon: 0.7740186400049058
steps: 263	lr: 4e-05	evaluation reward: 2.91	
episode: 1086	score: 1.0	memory length: 100000	epsilon: 0.7736840200049131
steps: 169	lr: 4e-05	evaluation reward: 2.87	
episode: 1087	score: 2.0	memory length: 100000	epsilon: 0.7732919800049216
steps: 198	lr: 4e-05	evaluation reward: 2.88	
episode: 1088	score: 4.0	memory length: 100000	epsilon: 0.7727375800049336
steps: 280	lr: 4e-05	evaluation reward: 2.92	
episode: 1089	score: 5.0	memory length: 100000	epsilon: 0.7721970400049454
steps: 273	lr: 4e-05	evaluation reward: 2.97	
episode: 1090	score: 2.0	memory length: 100000	epsilon: 0.771801040004954
steps: 200	lr: 4e-05	evaluation reward: 2.96	
episode: 1091	score: 5.0	memory length: 100000	epsilon: 0.7711219000049687
steps: 343	lr: 4e-05	evaluation reward: 2.97	
episode: 1092	score: 10.0	memory length: 100000	epsilon: 0.7703398000049857
steps: 395	lr: 4e-05	evaluation reward: 3.06	
episode: 1093	score: 4.0	memory length: 100000	epsilon: 0.7697933200049976
steps: 276	lr: 4e-05	evaluation reward: 3.06	
episode: 1094	score: 7.0	memory length: 100000	epsilon: 0.7692706000050089
steps: 264	lr: 4e-05	evaluation reward: 3.11	
episode: 1095	score: 2.0	memory length: 100000	epsilon: 0.7688785600050174
steps: 198	lr: 4e-05	evaluation reward: 3.12	
episode: 1096	score: 2.0	memory length: 100000	epsilon: 0.7685201800050252
steps: 181	lr: 4e-05	evaluation reward: 3.1	
episode: 1097	score: 3.0	memory length: 100000	epsilon: 0.7681004200050343
steps: 212	lr: 4e-05	evaluation reward: 3.09	
episode: 1098	score: 3.0	memory length: 100000	epsilon: 0.7676430400050442
steps: 231	lr: 4e-05	evaluation reward: 3.11	
episode: 1099	score: 5.0	memory length: 100000	epsilon: 0.7669936000050583

steps: 328	lr: 4e-05	evaluation reward: 3.13	
episode: 1100	score: 6.0	memory length: 100000	epsilon: 0.7663481200050724
steps: 326	lr: 4e-05	evaluation reward: 3.18	
episode: 1101	score: 1.0	memory length: 100000	epsilon: 0.7660511200050788
steps: 150	lr: 4e-05	evaluation reward: 3.16	
episode: 1102	score: 4.0	memory length: 100000	epsilon: 0.7655066200050906
steps: 275	lr: 4e-05	evaluation reward: 3.15	
episode: 1103	score: 7.0	memory length: 100000	epsilon: 0.7647839200051063
steps: 365	lr: 4e-05	evaluation reward: 3.2	
episode: 1104	score: 2.0	memory length: 100000	epsilon: 0.7643879200051149
steps: 200	lr: 4e-05	evaluation reward: 3.2	
episode: 1105	score: 1.0	memory length: 100000	epsilon: 0.7640533000051222
steps: 169	lr: 4e-05	evaluation reward: 3.2	
episode: 1106	score: 3.0	memory length: 100000	epsilon: 0.7636038400051319
steps: 227	lr: 4e-05	evaluation reward: 3.2	
episode: 1107	score: 3.0	memory length: 100000	epsilon: 0.7631207200051424
steps: 244	lr: 4e-05	evaluation reward: 3.23	
episode: 1108	score: 5.0	memory length: 100000	epsilon: 0.7625128600051556
steps: 307	lr: 4e-05	evaluation reward: 3.24	
episode: 1109	score: 7.0	memory length: 100000	epsilon: 0.7618258000051705
steps: 347	lr: 4e-05	evaluation reward: 3.3	
episode: 1110	score: 3.0	memory length: 100000	epsilon: 0.7613723800051804
steps: 229	lr: 4e-05	evaluation reward: 3.27	
episode: 1111	score: 4.0	memory length: 100000	epsilon: 0.760792240005193
steps: 293	lr: 4e-05	evaluation reward: 3.29	
episode: 1112	score: 3.0	memory length: 100000	epsilon: 0.7603467400052026
steps: 225	lr: 4e-05	evaluation reward: 3.31	
episode: 1113	score: 6.0	memory length: 100000	epsilon: 0.7596438400052179
steps: 355	lr: 4e-05	evaluation reward: 3.33	
episode: 1114	score: 2.0	memory length: 100000	epsilon: 0.7592478400052265
steps: 200	lr: 4e-05	evaluation reward: 3.31	
episode: 1115	score: 5.0	memory length: 100000	epsilon: 0.7586914600052386
steps: 281	lr: 4e-05	evaluation reward: 3.33	
episode: 1116	score: 3.0	memory length: 100000	epsilon: 0.758210320005249
steps: 243	lr: 4e-05	evaluation reward: 3.35	
episode: 1117	score: 2.0	memory length: 100000	epsilon: 0.7578182800052575
steps: 198	lr: 4e-05	evaluation reward: 3.37	
episode: 1118	score: 5.0	memory length: 100000	epsilon: 0.7571233000052726
steps: 351	lr: 4e-05	evaluation reward: 3.38	
episode: 1119	score: 4.0	memory length: 100000	epsilon: 0.7564916800052863
steps: 319	lr: 4e-05	evaluation reward: 3.38	
episode: 1120	score: 3.0	memory length: 100000	epsilon: 0.7560343000052963
steps: 231	lr: 4e-05	evaluation reward: 3.35	
episode: 1121	score: 3.0	memory length: 100000	epsilon: 0.755539300005307
steps: 250	lr: 4e-05	evaluation reward: 3.34	
episode: 1122	score: 3.0	memory length: 100000	epsilon: 0.7550898400053168
steps: 227	lr: 4e-05	evaluation reward: 3.34	
episode: 1123	score: 5.0	memory length: 100000	epsilon: 0.7544423800053308

steps: 327	lr: 4e-05	evaluation reward: 3.39	
episode: 1124	score: 3.0	memory length: 100000	epsilon: 0.7539968800053405
steps: 225	lr: 4e-05	evaluation reward: 3.39	
episode: 1125	score: 3.0	memory length: 100000	epsilon: 0.753511780005351
steps: 245	lr: 4e-05	evaluation reward: 3.39	
episode: 1126	score: 3.0	memory length: 100000	epsilon: 0.7530643000053607
steps: 226	lr: 4e-05	evaluation reward: 3.34	
episode: 1127	score: 3.0	memory length: 100000	epsilon: 0.7526108800053706
steps: 229	lr: 4e-05	evaluation reward: 3.32	
episode: 1128	score: 2.0	memory length: 100000	epsilon: 0.7522188400053791
steps: 198	lr: 4e-05	evaluation reward: 3.29	
episode: 1129	score: 3.0	memory length: 100000	epsilon: 0.7517654200053889
steps: 229	lr: 4e-05	evaluation reward: 3.29	
episode: 1130	score: 1.0	memory length: 100000	epsilon: 0.7514684200053954
steps: 150	lr: 4e-05	evaluation reward: 3.26	
episode: 1131	score: 2.0	memory length: 100000	epsilon: 0.7510744000054039
steps: 199	lr: 4e-05	evaluation reward: 3.26	
episode: 1132	score: 3.0	memory length: 100000	epsilon: 0.7506229600054137
steps: 228	lr: 4e-05	evaluation reward: 3.21	
episode: 1133	score: 2.0	memory length: 100000	epsilon: 0.7502309200054222
steps: 198	lr: 4e-05	evaluation reward: 3.22	
episode: 1134	score: 4.0	memory length: 100000	epsilon: 0.7497121600054335
steps: 262	lr: 4e-05	evaluation reward: 3.25	
episode: 1135	score: 3.0	memory length: 100000	epsilon: 0.7492607200054433
steps: 228	lr: 4e-05	evaluation reward: 3.25	
episode: 1136	score: 2.0	memory length: 100000	epsilon: 0.7488290800054527
steps: 218	lr: 4e-05	evaluation reward: 3.25	
episode: 1137	score: 6.0	memory length: 100000	epsilon: 0.7480489600054696
steps: 394	lr: 4e-05	evaluation reward: 3.3	
episode: 1138	score: 4.0	memory length: 100000	epsilon: 0.7475024800054815
steps: 276	lr: 4e-05	evaluation reward: 3.32	
episode: 1139	score: 5.0	memory length: 100000	epsilon: 0.7469005600054945
steps: 304	lr: 4e-05	evaluation reward: 3.35	
episode: 1140	score: 4.0	memory length: 100000	epsilon: 0.7463164600055072
steps: 295	lr: 4e-05	evaluation reward: 3.36	
episode: 1141	score: 5.0	memory length: 100000	epsilon: 0.745637320005522
steps: 343	lr: 4e-05	evaluation reward: 3.39	
episode: 1142	score: 4.0	memory length: 100000	epsilon: 0.745126480005533
steps: 258	lr: 4e-05	evaluation reward: 3.4	
episode: 1143	score: 2.0	memory length: 100000	epsilon: 0.7446968200055424
steps: 217	lr: 4e-05	evaluation reward: 3.39	
episode: 1144	score: 7.0	memory length: 100000	epsilon: 0.7439246200055591
steps: 390	lr: 4e-05	evaluation reward: 3.41	
episode: 1145	score: 8.0	memory length: 100000	epsilon: 0.7431346000055763
steps: 399	lr: 4e-05	evaluation reward: 3.45	
episode: 1146	score: 4.0	memory length: 100000	epsilon: 0.7425901000055881
steps: 275	lr: 4e-05	evaluation reward: 3.46	
episode: 1147	score: 3.0	memory length: 100000	epsilon: 0.7421406400055979

steps: 227	lr: 4e-05	evaluation reward: 3.46	
episode: 1148	score: 5.0	memory length: 100000	epsilon: 0.7415248600056112
steps: 311	lr: 4e-05	evaluation reward: 3.44	
episode: 1149	score: 2.0	memory length: 100000	epsilon: 0.7411328200056198
steps: 198	lr: 4e-05	evaluation reward: 3.44	
episode: 1150	score: 2.0	memory length: 100000	epsilon: 0.7407031600056291
steps: 217	lr: 4e-05	evaluation reward: 3.42	
episode: 1151	score: 5.0	memory length: 100000	epsilon: 0.7400576800056431
steps: 326	lr: 4e-05	evaluation reward: 3.46	
episode: 1152	score: 5.0	memory length: 100000	epsilon: 0.739370620005658
steps: 347	lr: 4e-05	evaluation reward: 3.45	
episode: 1153	score: 4.0	memory length: 100000	epsilon: 0.738863740005669
steps: 256	lr: 4e-05	evaluation reward: 3.44	
episode: 1154	score: 4.0	memory length: 100000	epsilon: 0.7383172600056809
steps: 276	lr: 4e-05	evaluation reward: 3.42	
episode: 1155	score: 3.0	memory length: 100000	epsilon: 0.7378658200056907
steps: 228	lr: 4e-05	evaluation reward: 3.45	
episode: 1156	score: 3.0	memory length: 100000	epsilon: 0.7374084400057006
steps: 231	lr: 4e-05	evaluation reward: 3.47	
episode: 1157	score: 8.0	memory length: 100000	epsilon: 0.7365768400057187
steps: 420	lr: 4e-05	evaluation reward: 3.51	
episode: 1158	score: 1.0	memory length: 100000	epsilon: 0.7362442000057259
steps: 168	lr: 4e-05	evaluation reward: 3.48	
episode: 1159	score: 4.0	memory length: 100000	epsilon: 0.7356601000057386
steps: 295	lr: 4e-05	evaluation reward: 3.49	
episode: 1160	score: 0.0	memory length: 100000	epsilon: 0.7354185400057438
steps: 122	lr: 4e-05	evaluation reward: 3.44	
episode: 1161	score: 3.0	memory length: 100000	epsilon: 0.7349275000057545
steps: 248	lr: 4e-05	evaluation reward: 3.42	
episode: 1162	score: 7.0	memory length: 100000	epsilon: 0.7340978800057725
steps: 419	lr: 4e-05	evaluation reward: 3.45	
episode: 1163	score: 4.0	memory length: 100000	epsilon: 0.7335177400057851
steps: 293	lr: 4e-05	evaluation reward: 3.48	
episode: 1164	score: 3.0	memory length: 100000	epsilon: 0.7330286800057957
steps: 247	lr: 4e-05	evaluation reward: 3.47	
episode: 1165	score: 6.0	memory length: 100000	epsilon: 0.7322941000058116
steps: 371	lr: 4e-05	evaluation reward: 3.53	
episode: 1166	score: 3.0	memory length: 100000	epsilon: 0.7318446400058214
steps: 227	lr: 4e-05	evaluation reward: 3.54	
episode: 1167	score: 1.0	memory length: 100000	epsilon: 0.7315456600058279
steps: 151	lr: 4e-05	evaluation reward: 3.53	
episode: 1168	score: 4.0	memory length: 100000	epsilon: 0.7310011600058397
steps: 275	lr: 4e-05	evaluation reward: 3.55	
episode: 1169	score: 4.0	memory length: 100000	epsilon: 0.7304111200058525
steps: 298	lr: 4e-05	evaluation reward: 3.56	
episode: 1170	score: 0.0	memory length: 100000	epsilon: 0.7301675800058578
steps: 123	lr: 4e-05	evaluation reward: 3.52	
episode: 1171	score: 5.0	memory length: 100000	epsilon: 0.7295300200058716

steps: 322	lr: 4e-05	evaluation reward: 3.56	
episode: 1172	score: 1.0	memory length: 100000	epsilon: 0.729191440005879
steps: 171	lr: 4e-05	evaluation reward: 3.53	
episode: 1173	score: 3.0	memory length: 100000	epsilon: 0.7287697000058881
steps: 213	lr: 4e-05	evaluation reward: 3.53	
episode: 1174	score: 0.0	memory length: 100000	epsilon: 0.7285261600058934
steps: 123	lr: 4e-05	evaluation reward: 3.49	
episode: 1175	score: 5.0	memory length: 100000	epsilon: 0.727809400005909
steps: 362	lr: 4e-05	evaluation reward: 3.53	
episode: 1176	score: 3.0	memory length: 100000	epsilon: 0.7273223200059196
steps: 246	lr: 4e-05	evaluation reward: 3.48	
episode: 1177	score: 2.0	memory length: 100000	epsilon: 0.7269659200059273
steps: 180	lr: 4e-05	evaluation reward: 3.49	
episode: 1178	score: 4.0	memory length: 100000	epsilon: 0.72638182000594
steps: 295	lr: 4e-05	evaluation reward: 3.5	
episode: 1179	score: 8.0	memory length: 100000	epsilon: 0.7255264600059586
steps: 432	lr: 4e-05	evaluation reward: 3.57	
episode: 1180	score: 5.0	memory length: 100000	epsilon: 0.7249601800059708
steps: 286	lr: 4e-05	evaluation reward: 3.59	
episode: 1181	score: 5.0	memory length: 100000	epsilon: 0.724354300005984
steps: 306	lr: 4e-05	evaluation reward: 3.63	
episode: 1182	score: 5.0	memory length: 100000	epsilon: 0.7237028800059981
steps: 329	lr: 4e-05	evaluation reward: 3.63	
episode: 1183	score: 4.0	memory length: 100000	epsilon: 0.7231940200060092
steps: 257	lr: 4e-05	evaluation reward: 3.65	
episode: 1184	score: 6.0	memory length: 100000	epsilon: 0.7224119200060262
steps: 395	lr: 4e-05	evaluation reward: 3.71	
episode: 1185	score: 5.0	memory length: 100000	epsilon: 0.7217209000060412
steps: 349	lr: 4e-05	evaluation reward: 3.73	
episode: 1186	score: 5.0	memory length: 100000	epsilon: 0.7210734400060552
steps: 327	lr: 4e-05	evaluation reward: 3.77	
episode: 1187	score: 3.0	memory length: 100000	epsilon: 0.7206517000060644
steps: 213	lr: 4e-05	evaluation reward: 3.78	
episode: 1188	score: 3.0	memory length: 100000	epsilon: 0.7201210600060759
steps: 268	lr: 4e-05	evaluation reward: 3.77	
episode: 1189	score: 6.0	memory length: 100000	epsilon: 0.719380540006092
steps: 374	lr: 4e-05	evaluation reward: 3.78	
episode: 1190	score: 2.0	memory length: 100000	epsilon: 0.7189845400061006
steps: 200	lr: 4e-05	evaluation reward: 3.78	
episode: 1191	score: 4.0	memory length: 100000	epsilon: 0.7183945000061134
steps: 298	lr: 4e-05	evaluation reward: 3.77	
episode: 1192	score: 2.0	memory length: 100000	epsilon: 0.7180024600061219
steps: 198	lr: 4e-05	evaluation reward: 3.69	
episode: 1193	score: 5.0	memory length: 100000	epsilon: 0.7174361800061342
steps: 286	lr: 4e-05	evaluation reward: 3.7	
episode: 1194	score: 2.0	memory length: 100000	epsilon: 0.7170421600061427
steps: 199	lr: 4e-05	evaluation reward: 3.65	
episode: 1195	score: 6.0	memory length: 100000	epsilon: 0.7163749000061572

steps: 337	lr: 4e-05	evaluation reward: 3.69	
episode: 1196	score: 0.0	memory length: 100000	epsilon: 0.7161313600061625
steps: 123	lr: 4e-05	evaluation reward: 3.67	
episode: 1197	score: 6.0	memory length: 100000	epsilon: 0.7154740000061768
steps: 332	lr: 4e-05	evaluation reward: 3.7	
episode: 1198	score: 3.0	memory length: 100000	epsilon: 0.7149493000061882
steps: 265	lr: 4e-05	evaluation reward: 3.7	
episode: 1199	score: 3.0	memory length: 100000	epsilon: 0.714495880006198
steps: 229	lr: 4e-05	evaluation reward: 3.68	
episode: 1200	score: 1.0	memory length: 100000	epsilon: 0.7141969000062045
steps: 151	lr: 4e-05	evaluation reward: 3.63	
episode: 1201	score: 4.0	memory length: 100000	epsilon: 0.7136504200062164
steps: 276	lr: 4e-05	evaluation reward: 3.66	
episode: 1202	score: 2.0	memory length: 100000	epsilon: 0.7132583800062249
steps: 198	lr: 4e-05	evaluation reward: 3.64	
episode: 1203	score: 3.0	memory length: 100000	epsilon: 0.7127752600062354
steps: 244	lr: 4e-05	evaluation reward: 3.6	
episode: 1204	score: 2.0	memory length: 100000	epsilon: 0.712377280006244
steps: 201	lr: 4e-05	evaluation reward: 3.6	
episode: 1205	score: 3.0	memory length: 100000	epsilon: 0.711963460006253
steps: 209	lr: 4e-05	evaluation reward: 3.62	
episode: 1206	score: 1.0	memory length: 100000	epsilon: 0.7116288400062603
steps: 169	lr: 4e-05	evaluation reward: 3.6	
episode: 1207	score: 1.0	memory length: 100000	epsilon: 0.7113298600062667
steps: 151	lr: 4e-05	evaluation reward: 3.58	
episode: 1208	score: 5.0	memory length: 100000	epsilon: 0.7106863600062807
steps: 325	lr: 4e-05	evaluation reward: 3.58	
episode: 1209	score: 3.0	memory length: 100000	epsilon: 0.7101973000062913
steps: 247	lr: 4e-05	evaluation reward: 3.54	
episode: 1210	score: 4.0	memory length: 100000	epsilon: 0.7097221000063016
steps: 240	lr: 4e-05	evaluation reward: 3.55	
episode: 1211	score: 1.0	memory length: 100000	epsilon: 0.7093874800063089
steps: 169	lr: 4e-05	evaluation reward: 3.52	
episode: 1212	score: 5.0	memory length: 100000	epsilon: 0.7087083400063237
steps: 343	lr: 4e-05	evaluation reward: 3.54	
episode: 1213	score: 7.0	memory length: 100000	epsilon: 0.707864860006342
steps: 426	lr: 4e-05	evaluation reward: 3.55	
episode: 1214	score: 4.0	memory length: 100000	epsilon: 0.707310460006354
steps: 280	lr: 4e-05	evaluation reward: 3.57	
episode: 1215	score: 5.0	memory length: 100000	epsilon: 0.7066946800063674
steps: 311	lr: 4e-05	evaluation reward: 3.57	
episode: 1216	score: 3.0	memory length: 100000	epsilon: 0.706205620006378
steps: 247	lr: 4e-05	evaluation reward: 3.57	
episode: 1217	score: 5.0	memory length: 100000	epsilon: 0.7056353800063904
steps: 288	lr: 4e-05	evaluation reward: 3.6	
episode: 1218	score: 7.0	memory length: 100000	epsilon: 0.704825560006408
steps: 409	lr: 4e-05	evaluation reward: 3.62	
episode: 1219	score: 8.0	memory length: 100000	epsilon: 0.7039306000064274

steps: 452	lr: 4e-05	evaluation reward: 3.66	
episode: 1220	score: 2.0	memory length: 100000	epsilon: 0.7034989600064367
steps: 218	lr: 4e-05	evaluation reward: 3.65	
episode: 1221	score: 2.0	memory length: 100000	epsilon: 0.7031386000064446
steps: 182	lr: 4e-05	evaluation reward: 3.64	
episode: 1222	score: 3.0	memory length: 100000	epsilon: 0.7026911200064543
steps: 226	lr: 4e-05	evaluation reward: 3.64	
episode: 1223	score: 5.0	memory length: 100000	epsilon: 0.7020654400064679
steps: 316	lr: 4e-05	evaluation reward: 3.64	
episode: 1224	score: 3.0	memory length: 100000	epsilon: 0.7015387600064793
steps: 266	lr: 4e-05	evaluation reward: 3.64	
episode: 1225	score: 8.0	memory length: 100000	epsilon: 0.7006853800064978
steps: 431	lr: 4e-05	evaluation reward: 3.69	
episode: 1226	score: 3.0	memory length: 100000	epsilon: 0.7001923600065085
steps: 249	lr: 4e-05	evaluation reward: 3.69	
episode: 1227	score: 0.0	memory length: 100000	epsilon: 0.6999488200065138
steps: 123	lr: 4e-05	evaluation reward: 3.66	
episode: 1228	score: 3.0	memory length: 100000	epsilon: 0.6994993600065236
steps: 227	lr: 4e-05	evaluation reward: 3.67	
episode: 1229	score: 4.0	memory length: 100000	epsilon: 0.6989865400065347
steps: 259	lr: 4e-05	evaluation reward: 3.68	
episode: 1230	score: 1.0	memory length: 100000	epsilon: 0.6986875600065412
steps: 151	lr: 4e-05	evaluation reward: 3.68	
episode: 1231	score: 5.0	memory length: 100000	epsilon: 0.6980401000065553
steps: 327	lr: 4e-05	evaluation reward: 3.71	
episode: 1232	score: 1.0	memory length: 100000	epsilon: 0.6977411200065617
steps: 151	lr: 4e-05	evaluation reward: 3.69	
episode: 1233	score: 7.0	memory length: 100000	epsilon: 0.6969709000065785
steps: 389	lr: 4e-05	evaluation reward: 3.74	
episode: 1234	score: 5.0	memory length: 100000	epsilon: 0.6963689800065915
steps: 304	lr: 4e-05	evaluation reward: 3.75	
episode: 1235	score: 8.0	memory length: 100000	epsilon: 0.6954839200066107
steps: 447	lr: 4e-05	evaluation reward: 3.8	
episode: 1236	score: 5.0	memory length: 100000	epsilon: 0.6948681400066241
steps: 311	lr: 4e-05	evaluation reward: 3.83	
episode: 1237	score: 3.0	memory length: 100000	epsilon: 0.6944206600066338
steps: 226	lr: 4e-05	evaluation reward: 3.8	
episode: 1238	score: 6.0	memory length: 100000	epsilon: 0.693765280006648
steps: 331	lr: 4e-05	evaluation reward: 3.82	
episode: 1239	score: 10.0	memory length: 100000	epsilon:
0.6927515200066701	steps: 512	lr: 4e-05	evaluation reward: 3.87
episode: 1240	score: 3.0	memory length: 100000	epsilon: 0.6922862200066802
steps: 235	lr: 4e-05	evaluation reward: 3.86	
episode: 1241	score: 4.0	memory length: 100000	epsilon: 0.6917456800066919
steps: 273	lr: 4e-05	evaluation reward: 3.85	
episode: 1242	score: 9.0	memory length: 100000	epsilon: 0.6908705200067109
steps: 442	lr: 4e-05	evaluation reward: 3.9	
episode: 1243	score: 5.0	memory length: 100000	epsilon: 0.6902567200067242

steps: 310	lr: 4e-05	evaluation reward: 3.93	
episode: 1244	score: 3.0	memory length: 100000	epsilon: 0.6898349800067334
steps: 213	lr: 4e-05	evaluation reward: 3.89	
episode: 1245	score: 4.0	memory length: 100000	epsilon: 0.6892845400067453
steps: 278	lr: 4e-05	evaluation reward: 3.85	
episode: 1246	score: 5.0	memory length: 100000	epsilon: 0.68860936000676
steps: 341	lr: 4e-05	evaluation reward: 3.86	
episode: 1247	score: 4.0	memory length: 100000	epsilon: 0.6880609000067719
steps: 277	lr: 4e-05	evaluation reward: 3.87	
episode: 1248	score: 4.0	memory length: 100000	epsilon: 0.687548080006783
steps: 259	lr: 4e-05	evaluation reward: 3.86	
episode: 1249	score: 7.0	memory length: 100000	epsilon: 0.6870570400067937
steps: 248	lr: 4e-05	evaluation reward: 3.91	
episode: 1250	score: 1.0	memory length: 100000	epsilon: 0.686718460006801
steps: 171	lr: 4e-05	evaluation reward: 3.9	
episode: 1251	score: 8.0	memory length: 100000	epsilon: 0.6859482400068178
steps: 389	lr: 4e-05	evaluation reward: 3.93	
episode: 1252	score: 3.0	memory length: 100000	epsilon: 0.6854968000068276
steps: 228	lr: 4e-05	evaluation reward: 3.91	
episode: 1253	score: 3.0	memory length: 100000	epsilon: 0.6850037800068383
steps: 249	lr: 4e-05	evaluation reward: 3.9	
episode: 1254	score: 8.0	memory length: 100000	epsilon: 0.6841662400068564
steps: 423	lr: 4e-05	evaluation reward: 3.94	
episode: 1255	score: 2.0	memory length: 100000	epsilon: 0.683770240006865
steps: 200	lr: 4e-05	evaluation reward: 3.93	
episode: 1256	score: 2.0	memory length: 100000	epsilon: 0.6833782000068735
steps: 198	lr: 4e-05	evaluation reward: 3.92	
episode: 1257	score: 6.0	memory length: 100000	epsilon: 0.6827287600068876
steps: 328	lr: 4e-05	evaluation reward: 3.9	
episode: 1258	score: 5.0	memory length: 100000	epsilon: 0.6820793200069017
steps: 328	lr: 4e-05	evaluation reward: 3.94	
episode: 1259	score: 2.0	memory length: 100000	epsilon: 0.6816853000069103
steps: 199	lr: 4e-05	evaluation reward: 3.92	
episode: 1260	score: 8.0	memory length: 100000	epsilon: 0.68077648000693
steps: 459	lr: 4e-05	evaluation reward: 4.0	
episode: 1261	score: 7.0	memory length: 100000	epsilon: 0.6800082400069467
steps: 388	lr: 4e-05	evaluation reward: 4.04	
episode: 1262	score: 3.0	memory length: 100000	epsilon: 0.6795548200069566
steps: 229	lr: 4e-05	evaluation reward: 4.0	
episode: 1263	score: 6.0	memory length: 100000	epsilon: 0.6788855800069711
steps: 338	lr: 4e-05	evaluation reward: 4.02	
episode: 1264	score: 3.0	memory length: 100000	epsilon: 0.6783490000069827
steps: 271	lr: 4e-05	evaluation reward: 4.02	
episode: 1265	score: 1.0	memory length: 100000	epsilon: 0.6780460600069893
steps: 153	lr: 4e-05	evaluation reward: 3.97	
episode: 1266	score: 6.0	memory length: 100000	epsilon: 0.6773075200070053
steps: 373	lr: 4e-05	evaluation reward: 4.0	
episode: 1267	score: 3.0	memory length: 100000	epsilon: 0.6768501400070153

steps: 231	lr: 4e-05	evaluation reward: 4.02	
episode: 1268	score: 8.0	memory length: 100000	epsilon: 0.6759136000070356
steps: 473	lr: 4e-05	evaluation reward: 4.06	
episode: 1269	score: 5.0	memory length: 100000	epsilon: 0.6753017800070489
steps: 309	lr: 4e-05	evaluation reward: 4.07	
episode: 1270	score: 3.0	memory length: 100000	epsilon: 0.6748087600070596
steps: 249	lr: 4e-05	evaluation reward: 4.1	
episode: 1271	score: 3.0	memory length: 100000	epsilon: 0.6742741600070712
steps: 270	lr: 4e-05	evaluation reward: 4.08	
episode: 1272	score: 8.0	memory length: 100000	epsilon: 0.6736880800070839
steps: 296	lr: 4e-05	evaluation reward: 4.15	
episode: 1273	score: 2.0	memory length: 100000	epsilon: 0.6732980200070924
steps: 197	lr: 4e-05	evaluation reward: 4.14	
episode: 1274	score: 4.0	memory length: 100000	epsilon: 0.6727515400071042
steps: 276	lr: 4e-05	evaluation reward: 4.18	
episode: 1275	score: 5.0	memory length: 100000	epsilon: 0.6720664600071191
steps: 346	lr: 4e-05	evaluation reward: 4.18	
episode: 1276	score: 2.0	memory length: 100000	epsilon: 0.6716744200071276
steps: 198	lr: 4e-05	evaluation reward: 4.17	
episode: 1277	score: 2.0	memory length: 100000	epsilon: 0.6713140600071354
steps: 182	lr: 4e-05	evaluation reward: 4.17	
episode: 1278	score: 1.0	memory length: 100000	epsilon: 0.671013100007142
steps: 152	lr: 4e-05	evaluation reward: 4.14	
episode: 1279	score: 6.0	memory length: 100000	epsilon: 0.6702409000071587
steps: 390	lr: 4e-05	evaluation reward: 4.12	
episode: 1280	score: 10.0	memory length: 100000	epsilon: 0.669351880007178
steps: 449	lr: 4e-05	evaluation reward: 4.17	
episode: 1281	score: 9.0	memory length: 100000	epsilon: 0.6684094000071985
steps: 476	lr: 4e-05	evaluation reward: 4.21	
episode: 1282	score: 1.0	memory length: 100000	epsilon: 0.6680747800072058
steps: 169	lr: 4e-05	evaluation reward: 4.17	
episode: 1283	score: 3.0	memory length: 100000	epsilon: 0.6675916600072163
steps: 244	lr: 4e-05	evaluation reward: 4.16	
episode: 1284	score: 9.0	memory length: 100000	epsilon: 0.6666551200072366
steps: 473	lr: 4e-05	evaluation reward: 4.19	
episode: 1285	score: 5.0	memory length: 100000	epsilon: 0.6660472600072498
steps: 307	lr: 4e-05	evaluation reward: 4.19	
episode: 1286	score: 3.0	memory length: 100000	epsilon: 0.6655146400072613
steps: 269	lr: 4e-05	evaluation reward: 4.17	
episode: 1287	score: 2.0	memory length: 100000	epsilon: 0.6650810200072708
steps: 219	lr: 4e-05	evaluation reward: 4.16	
episode: 1288	score: 3.0	memory length: 100000	epsilon: 0.6646632400072798
steps: 211	lr: 4e-05	evaluation reward: 4.16	
episode: 1289	score: 5.0	memory length: 100000	epsilon: 0.6640791400072925
steps: 295	lr: 4e-05	evaluation reward: 4.15	
episode: 1290	score: 6.0	memory length: 100000	epsilon: 0.663411880007307
steps: 337	lr: 4e-05	evaluation reward: 4.19	
episode: 1291	score: 5.0	memory length: 100000	epsilon: 0.6627980800073203

```

steps: 310    lr: 4e-05    evaluation reward: 4.2
episode: 1292  score: 4.0    memory length: 100000  epsilon: 0.6622872400073314
steps: 258    lr: 4e-05    evaluation reward: 4.22
episode: 1293  score: 11.0   memory length: 100000  epsilon:
0.6612437800073541    steps: 527    lr: 4e-05    evaluation reward: 4.28
episode: 1294  score: 4.0    memory length: 100000  epsilon: 0.660695320007366
steps: 277    lr: 4e-05    evaluation reward: 4.3
episode: 1295  score: 6.0    memory length: 100000  epsilon: 0.6599706400073817
steps: 366    lr: 4e-05    evaluation reward: 4.3
episode: 1296  score: 7.0    memory length: 100000  epsilon: 0.6591469600073996
steps: 416    lr: 4e-05    evaluation reward: 4.37
episode: 1297  score: 10.0   memory length: 100000  epsilon:
0.6580322200074238    steps: 563    lr: 4e-05    evaluation reward: 4.41
episode: 1298  score: 5.0    memory length: 100000  epsilon: 0.657424360007437
steps: 307    lr: 4e-05    evaluation reward: 4.43
episode: 1299  score: 5.0    memory length: 100000  epsilon: 0.6568184800074501
steps: 306    lr: 4e-05    evaluation reward: 4.45
episode: 1300  score: 5.0    memory length: 100000  epsilon: 0.6561631000074644
steps: 331    lr: 4e-05    evaluation reward: 4.49
episode: 1301  score: 5.0    memory length: 100000  epsilon: 0.6555136600074785
steps: 328    lr: 4e-05    evaluation reward: 4.5
episode: 1302  score: 4.0    memory length: 100000  epsilon: 0.6549691600074903
steps: 275    lr: 4e-05    evaluation reward: 4.52
episode: 1303  score: 4.0    memory length: 100000  epsilon: 0.6544226800075021
steps: 276    lr: 4e-05    evaluation reward: 4.53
episode: 1304  score: 9.0    memory length: 100000  epsilon: 0.6535475200075211
steps: 442    lr: 4e-05    evaluation reward: 4.6
episode: 1305  score: 5.0    memory length: 100000  epsilon: 0.6529357000075344
steps: 309    lr: 4e-05    evaluation reward: 4.62
episode: 1306  score: 3.0    memory length: 100000  epsilon: 0.6524882200075441
steps: 226    lr: 4e-05    evaluation reward: 4.64
episode: 1307  score: 8.0    memory length: 100000  epsilon: 0.6516289000075628
steps: 434    lr: 4e-05    evaluation reward: 4.71
episode: 1308  score: 4.0    memory length: 100000  epsilon: 0.651114100007574
steps: 260    lr: 4e-05    evaluation reward: 4.7
episode: 1309  score: 7.0    memory length: 100000  epsilon: 0.6503518000075905
steps: 385    lr: 4e-05    evaluation reward: 4.74
episode: 1310  score: 3.0    memory length: 100000  epsilon: 0.6498944200076004
steps: 231    lr: 4e-05    evaluation reward: 4.73
episode: 1311  score: 6.0    memory length: 100000  epsilon: 0.6491875600076158
steps: 357    lr: 4e-05    evaluation reward: 4.78
episode: 1312  score: 9.0    memory length: 100000  epsilon: 0.6483223000076346
steps: 437    lr: 4e-05    evaluation reward: 4.82
episode: 1313  score: 5.0    memory length: 100000  epsilon: 0.6477104800076479
steps: 309    lr: 4e-05    evaluation reward: 4.8
episode: 1314  score: 2.0    memory length: 100000  epsilon: 0.6473164600076564
steps: 199    lr: 4e-05    evaluation reward: 4.78
episode: 1315  score: 8.0    memory length: 100000  epsilon: 0.6464413000076754

```



```

steps: 442    lr: 4e-05    evaluation reward: 4.81
episode: 1316  score: 6.0    memory length: 100000  epsilon: 0.6456651400076923
steps: 392    lr: 4e-05    evaluation reward: 4.84
episode: 1317  score: 5.0    memory length: 100000  epsilon: 0.6450592600077054
steps: 306    lr: 4e-05    evaluation reward: 4.84
episode: 1318  score: 2.0    memory length: 100000  epsilon: 0.6446593000077141
steps: 202    lr: 4e-05    evaluation reward: 4.79
episode: 1319  score: 8.0    memory length: 100000  epsilon: 0.6438277000077322
steps: 420    lr: 4e-05    evaluation reward: 4.79
episode: 1320  score: 5.0    memory length: 100000  epsilon: 0.6432198400077453
steps: 307    lr: 4e-05    evaluation reward: 4.82
episode: 1321  score: 6.0    memory length: 100000  epsilon: 0.6425169400077606
steps: 355    lr: 4e-05    evaluation reward: 4.86
episode: 1322  score: 6.0    memory length: 100000  epsilon: 0.6418199800077757
steps: 352    lr: 4e-05    evaluation reward: 4.89
episode: 1323  score: 11.0   memory length: 100000  epsilon:
0.6407171200077997  steps: 557    lr: 4e-05    evaluation reward: 4.95
episode: 1324  score: 2.0    memory length: 100000  epsilon: 0.6403270600078081
steps: 197    lr: 4e-05    evaluation reward: 4.94
episode: 1325  score: 3.0    memory length: 100000  epsilon: 0.6398399800078187
steps: 246    lr: 4e-05    evaluation reward: 4.89
episode: 1326  score: 3.0    memory length: 100000  epsilon: 0.6393925000078284
steps: 226    lr: 4e-05    evaluation reward: 4.89
episode: 1327  score: 5.0    memory length: 100000  epsilon: 0.6387173200078431
steps: 341    lr: 4e-05    evaluation reward: 4.94
episode: 1328  score: 2.0    memory length: 100000  epsilon: 0.6383589400078509
steps: 181    lr: 4e-05    evaluation reward: 4.93
episode: 1329  score: 4.0    memory length: 100000  epsilon: 0.6378144400078627
steps: 275    lr: 4e-05    evaluation reward: 4.93
episode: 1330  score: 2.0    memory length: 100000  epsilon: 0.6374540800078705
steps: 182    lr: 4e-05    evaluation reward: 4.94
episode: 1331  score: 5.0    memory length: 100000  epsilon: 0.6368680000078832
steps: 296    lr: 4e-05    evaluation reward: 4.94
episode: 1332  score: 6.0    memory length: 100000  epsilon: 0.6360878800079002
steps: 394    lr: 4e-05    evaluation reward: 4.99
episode: 1333  score: 5.0    memory length: 100000  epsilon: 0.6355216000079125
steps: 286    lr: 4e-05    evaluation reward: 4.97
episode: 1334  score: 4.0    memory length: 100000  epsilon: 0.6350087800079236
steps: 259    lr: 4e-05    evaluation reward: 4.96
episode: 1335  score: 1.0    memory length: 100000  epsilon: 0.63471178000793
steps: 150    lr: 4e-05    evaluation reward: 4.89
episode: 1336  score: 10.0   memory length: 100000  epsilon:
0.6336980200079521  steps: 512    lr: 4e-05    evaluation reward: 4.94
episode: 1337  score: 8.0    memory length: 100000  epsilon: 0.6328030600079715
steps: 452    lr: 4e-05    evaluation reward: 4.99
episode: 1338  score: 9.0    memory length: 100000  epsilon: 0.6318942400079912
steps: 459    lr: 4e-05    evaluation reward: 5.02
episode: 1339  score: 7.0    memory length: 100000  epsilon: 0.6311022400080084

```

```

steps: 400    lr: 4e-05    evaluation reward: 4.99
episode: 1340  score: 7.0    memory length: 100000  epsilon: 0.6303062800080257
steps: 402    lr: 4e-05    evaluation reward: 5.03
episode: 1341  score: 11.0   memory length: 100000  epsilon:
0.6292113400080495    steps: 553    lr: 4e-05    evaluation reward: 5.1
episode: 1342  score: 6.0    memory length: 100000  epsilon: 0.6285797200080632
steps: 319    lr: 4e-05    evaluation reward: 5.07
episode: 1343  score: 5.0    memory length: 100000  epsilon: 0.6279362200080771
steps: 325    lr: 4e-05    evaluation reward: 5.07
episode: 1344  score: 5.0    memory length: 100000  epsilon: 0.6272887600080912
steps: 327    lr: 4e-05    evaluation reward: 5.09
episode: 1345  score: 3.0    memory length: 100000  epsilon: 0.6268412800081009
steps: 226    lr: 4e-05    evaluation reward: 5.08
episode: 1346  score: 6.0    memory length: 100000  epsilon: 0.626146300008116
steps: 351    lr: 4e-05    evaluation reward: 5.09
episode: 1347  score: 7.0    memory length: 100000  epsilon: 0.6253028200081343
steps: 426    lr: 4e-05    evaluation reward: 5.12
episode: 1348  score: 4.0    memory length: 100000  epsilon: 0.6247147600081471
steps: 297    lr: 4e-05    evaluation reward: 5.12
episode: 1349  score: 3.0    memory length: 100000  epsilon: 0.624259360008157
steps: 230    lr: 4e-05    evaluation reward: 5.08
episode: 1350  score: 5.0    memory length: 100000  epsilon: 0.623613880008171
steps: 326    lr: 4e-05    evaluation reward: 5.12
episode: 1351  score: 5.0    memory length: 100000  epsilon: 0.6230317600081836
steps: 294    lr: 4e-05    evaluation reward: 5.09
episode: 1352  score: 4.0    memory length: 100000  epsilon: 0.6225189400081947
steps: 259    lr: 4e-05    evaluation reward: 5.1
episode: 1353  score: 7.0    memory length: 100000  epsilon: 0.6217566400082113
steps: 385    lr: 4e-05    evaluation reward: 5.14
episode: 1354  score: 5.0    memory length: 100000  epsilon: 0.6211151200082252
steps: 324    lr: 4e-05    evaluation reward: 5.11
episode: 1355  score: 3.0    memory length: 100000  epsilon: 0.6206676400082349
steps: 226    lr: 4e-05    evaluation reward: 5.12
episode: 1356  score: 4.0    memory length: 100000  epsilon: 0.6201073000082471
steps: 283    lr: 4e-05    evaluation reward: 5.14
episode: 1357  score: 7.0    memory length: 100000  epsilon: 0.6193054000082645
steps: 405    lr: 4e-05    evaluation reward: 5.15
episode: 1358  score: 7.0    memory length: 100000  epsilon: 0.6186044800082797
steps: 354    lr: 4e-05    evaluation reward: 5.17
episode: 1359  score: 3.0    memory length: 100000  epsilon: 0.6181570000082894
steps: 226    lr: 4e-05    evaluation reward: 5.18
episode: 1360  score: 8.0    memory length: 100000  epsilon: 0.6174303400083052
steps: 367    lr: 4e-05    evaluation reward: 5.18
episode: 1361  score: 8.0    memory length: 100000  epsilon: 0.6165591400083241
steps: 440    lr: 4e-05    evaluation reward: 5.19
episode: 1362  score: 7.0    memory length: 100000  epsilon: 0.6157235800083423
steps: 422    lr: 4e-05    evaluation reward: 5.23
episode: 1363  score: 5.0    memory length: 100000  epsilon: 0.6150068200083578

```

```

steps: 362    lr: 4e-05    evaluation reward: 5.22
episode: 1364  score: 0.0    memory length: 100000  epsilon: 0.6147632800083631
steps: 123    lr: 4e-05    evaluation reward: 5.19
episode: 1365  score: 11.0   memory length: 100000  epsilon:
0.6139336600083811  steps: 419    lr: 4e-05    evaluation reward: 5.29
episode: 1366  score: 4.0    memory length: 100000  epsilon: 0.6133891600083929
steps: 275    lr: 4e-05    evaluation reward: 5.27
episode: 1367  score: 2.0    memory length: 100000  epsilon: 0.6129931600084015
steps: 200    lr: 4e-05    evaluation reward: 5.26
episode: 1368  score: 5.0    memory length: 100000  epsilon: 0.612418960008414
steps: 290    lr: 4e-05    evaluation reward: 5.23
episode: 1369  score: 7.0    memory length: 100000  epsilon: 0.6116467600084308
steps: 390    lr: 4e-05    evaluation reward: 5.25
episode: 1370  score: 9.0    memory length: 100000  epsilon: 0.610809220008449
steps: 423    lr: 4e-05    evaluation reward: 5.31
episode: 1371  score: 5.0    memory length: 100000  epsilon: 0.6101974000084622
steps: 309    lr: 4e-05    evaluation reward: 5.33
episode: 1372  score: 6.0    memory length: 100000  epsilon: 0.6094885600084776
steps: 358    lr: 4e-05    evaluation reward: 5.31
episode: 1373  score: 9.0    memory length: 100000  epsilon: 0.6085361800084983
steps: 481    lr: 4e-05    evaluation reward: 5.38
episode: 1374  score: 7.0    memory length: 100000  epsilon: 0.6077283400085158
steps: 408    lr: 4e-05    evaluation reward: 5.41
episode: 1375  score: 3.0    memory length: 100000  epsilon: 0.6073026400085251
steps: 215    lr: 4e-05    evaluation reward: 5.39
episode: 1376  score: 5.0    memory length: 100000  epsilon: 0.6066195400085399
steps: 345    lr: 4e-05    evaluation reward: 5.42
episode: 1377  score: 7.0    memory length: 100000  epsilon: 0.6058671400085562
steps: 380    lr: 4e-05    evaluation reward: 5.47
episode: 1378  score: 8.0    memory length: 100000  epsilon: 0.604956340008576
steps: 460    lr: 4e-05    evaluation reward: 5.54
episode: 1379  score: 9.0    memory length: 100000  epsilon: 0.6040653400085954
steps: 450    lr: 4e-05    evaluation reward: 5.57
episode: 1380  score: 8.0    memory length: 100000  epsilon: 0.6032495800086131
steps: 412    lr: 1.6000000000000003e-05    evaluation reward: 5.55
episode: 1381  score: 3.0    memory length: 100000  epsilon: 0.6027585400086237
steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 5.49
episode: 1382  score: 6.0    memory length: 100000  epsilon: 0.6020952400086381
steps: 335    lr: 1.6000000000000003e-05    evaluation reward: 5.54
episode: 1383  score: 7.0    memory length: 100000  epsilon: 0.6012874000086557
steps: 408    lr: 1.6000000000000003e-05    evaluation reward: 5.58
episode: 1384  score: 2.0    memory length: 100000  epsilon: 0.6008953600086642
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 5.51
episode: 1385  score: 1.0    memory length: 100000  epsilon: 0.6005607400086714
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 5.47
episode: 1386  score: 3.0    memory length: 100000  epsilon: 0.600028120008683
steps: 269    lr: 1.6000000000000003e-05    evaluation reward: 5.47
episode: 1387  score: 3.0    memory length: 100000  epsilon: 0.5995311400086938

```

```

steps: 251    lr: 1.6000000000000003e-05    evaluation reward: 5.48
episode: 1388  score: 7.0    memory length: 100000    epsilon: 0.5987569600087106
steps: 391    lr: 1.6000000000000003e-05    evaluation reward: 5.52
episode: 1389  score: 8.0    memory length: 100000    epsilon: 0.5979629800087278
steps: 401    lr: 1.6000000000000003e-05    evaluation reward: 5.55
episode: 1390  score: 4.0    memory length: 100000    epsilon: 0.5974858000087382
steps: 241    lr: 1.6000000000000003e-05    evaluation reward: 5.53
episode: 1391  score: 6.0    memory length: 100000    epsilon: 0.5968225000087526
steps: 335    lr: 1.6000000000000003e-05    evaluation reward: 5.54
episode: 1392  score: 7.0    memory length: 100000    epsilon: 0.59602258000877
steps: 404    lr: 1.6000000000000003e-05    evaluation reward: 5.57
episode: 1393  score: 6.0    memory length: 100000    epsilon: 0.5953790800087839
steps: 325    lr: 1.6000000000000003e-05    evaluation reward: 5.52
episode: 1394  score: 8.0    memory length: 100000    epsilon: 0.5947969600087966
steps: 294    lr: 1.6000000000000003e-05    evaluation reward: 5.56
episode: 1395  score: 7.0    memory length: 100000    epsilon: 0.5939732800088144
steps: 416    lr: 1.6000000000000003e-05    evaluation reward: 5.57
episode: 1396  score: 1.0    memory length: 100000    epsilon: 0.5936743000088209
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 5.51
episode: 1397  score: 7.0    memory length: 100000    epsilon: 0.5929199200088373
steps: 381    lr: 1.6000000000000003e-05    evaluation reward: 5.48
episode: 1398  score: 8.0    memory length: 100000    epsilon: 0.5920269400088567
steps: 451    lr: 1.6000000000000003e-05    evaluation reward: 5.51
episode: 1399  score: 10.0   memory length: 100000    epsilon:
0.5910329800088783    steps: 502    lr: 1.6000000000000003e-05    evaluation
reward: 5.56
episode: 1400  score: 3.0    memory length: 100000    epsilon: 0.5905874800088879
steps: 225    lr: 1.6000000000000003e-05    evaluation reward: 5.54
episode: 1401  score: 12.0   memory length: 100000    epsilon:
0.5895103600089113    steps: 544    lr: 1.6000000000000003e-05    evaluation
reward: 5.61
episode: 1402  score: 5.0    memory length: 100000    epsilon: 0.5888985400089246
steps: 309    lr: 1.6000000000000003e-05    evaluation reward: 5.62
episode: 1403  score: 6.0    memory length: 100000    epsilon: 0.5882293000089391
steps: 338    lr: 1.6000000000000003e-05    evaluation reward: 5.64
episode: 1404  score: 4.0    memory length: 100000    epsilon: 0.5875996600089528
steps: 318    lr: 1.6000000000000003e-05    evaluation reward: 5.59
episode: 1405  score: 7.0    memory length: 100000    epsilon: 0.5868472600089691
steps: 380    lr: 1.6000000000000003e-05    evaluation reward: 5.61
episode: 1406  score: 8.0    memory length: 100000    epsilon: 0.5860354600089868
steps: 410    lr: 1.6000000000000003e-05    evaluation reward: 5.66
episode: 1407  score: 5.0    memory length: 100000    epsilon: 0.5853959200090006
steps: 323    lr: 1.6000000000000003e-05    evaluation reward: 5.63
episode: 1408  score: 5.0    memory length: 100000    epsilon: 0.584782120009014
steps: 310    lr: 1.6000000000000003e-05    evaluation reward: 5.64
episode: 1409  score: 1.0    memory length: 100000    epsilon: 0.5844851200090204
steps: 150    lr: 1.6000000000000003e-05    evaluation reward: 5.58
episode: 1410  score: 9.0    memory length: 100000    epsilon: 0.5835921400090398

```

```

steps: 451    lr: 1.6000000000000003e-05    evaluation reward: 5.64
episode: 1411  score: 2.0    memory length: 100000    epsilon: 0.5832337600090476
steps: 181    lr: 1.6000000000000003e-05    evaluation reward: 5.6
episode: 1412  score: 6.0    memory length: 100000    epsilon: 0.5825308600090628
steps: 355    lr: 1.6000000000000003e-05    evaluation reward: 5.57
episode: 1413  score: 2.0    memory length: 100000    epsilon: 0.5820992200090722
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 5.54
episode: 1414  score: 7.0    memory length: 100000    epsilon: 0.5814121600090871
steps: 347    lr: 1.6000000000000003e-05    evaluation reward: 5.59
episode: 1415  score: 1.0    memory length: 100000    epsilon: 0.5811112000090937
steps: 152    lr: 1.6000000000000003e-05    evaluation reward: 5.52
episode: 1416  score: 8.0    memory length: 100000    epsilon: 0.5803192000091109
steps: 400    lr: 1.6000000000000003e-05    evaluation reward: 5.54
episode: 1417  score: 7.0    memory length: 100000    epsilon: 0.579527200009128
steps: 400    lr: 1.6000000000000003e-05    evaluation reward: 5.56
episode: 1418  score: 9.0    memory length: 100000    epsilon: 0.5785906600091484
steps: 473    lr: 1.6000000000000003e-05    evaluation reward: 5.63
episode: 1419  score: 7.0    memory length: 100000    epsilon: 0.5778818200091638
steps: 358    lr: 1.6000000000000003e-05    evaluation reward: 5.62
episode: 1420  score: 6.0    memory length: 100000    epsilon: 0.5771373400091799
steps: 376    lr: 1.6000000000000003e-05    evaluation reward: 5.63
episode: 1421  score: 5.0    memory length: 100000    epsilon: 0.5765195800091933
steps: 312    lr: 1.6000000000000003e-05    evaluation reward: 5.62
episode: 1422  score: 6.0    memory length: 100000    epsilon: 0.5758107400092087
steps: 358    lr: 1.6000000000000003e-05    evaluation reward: 5.62
episode: 1423  score: 11.0   memory length: 100000    epsilon:
0.5747554000092316    steps: 533    lr: 1.6000000000000003e-05    evaluation
reward: 5.62
episode: 1424  score: 6.0    memory length: 100000    epsilon: 0.574046560009247
steps: 358    lr: 1.6000000000000003e-05    evaluation reward: 5.66
episode: 1425  score: 9.0    memory length: 100000    epsilon: 0.573127840009267
steps: 464    lr: 1.6000000000000003e-05    evaluation reward: 5.72
episode: 1426  score: 5.0    memory length: 100000    epsilon: 0.5725576000092794
steps: 288    lr: 1.6000000000000003e-05    evaluation reward: 5.74
episode: 1427  score: 5.0    memory length: 100000    epsilon: 0.5719438000092927
steps: 310    lr: 1.6000000000000003e-05    evaluation reward: 5.74
episode: 1428  score: 2.0    memory length: 100000    epsilon: 0.5715478000093013
steps: 200    lr: 1.6000000000000003e-05    evaluation reward: 5.74
episode: 1429  score: 5.0    memory length: 100000    epsilon: 0.5709835000093135
steps: 285    lr: 1.6000000000000003e-05    evaluation reward: 5.75
episode: 1430  score: 15.0   memory length: 100000    epsilon:
0.5698628200093379    steps: 566    lr: 1.6000000000000003e-05    evaluation
reward: 5.88
episode: 1431  score: 11.0   memory length: 100000    epsilon:
0.5687243200093626    steps: 575    lr: 1.6000000000000003e-05    evaluation
reward: 5.94
episode: 1432  score: 4.0    memory length: 100000    epsilon: 0.5681362600093753
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 5.92

```

episode: 1433 score: 11.0 memory length: 100000 epsilon:
 0.5671304200093972 steps: 508 lr: 1.6000000000000003e-05 evaluation
 reward: 5.98
 episode: 1434 score: 3.0 memory length: 100000 epsilon: 0.5666750200094071
 steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 5.97
 episode: 1435 score: 6.0 memory length: 100000 epsilon: 0.5659305400094232
 steps: 376 lr: 1.6000000000000003e-05 evaluation reward: 6.02
 episode: 1436 score: 4.0 memory length: 100000 epsilon: 0.5654117800094345
 steps: 262 lr: 1.6000000000000003e-05 evaluation reward: 5.96
 episode: 1437 score: 4.0 memory length: 100000 epsilon: 0.564928660009445
 steps: 244 lr: 1.6000000000000003e-05 evaluation reward: 5.92
 episode: 1438 score: 5.0 memory length: 100000 epsilon: 0.5643544600094574
 steps: 290 lr: 1.6000000000000003e-05 evaluation reward: 5.88
 episode: 1439 score: 13.0 memory length: 100000 epsilon:
 0.5630516200094857 steps: 658 lr: 1.6000000000000003e-05 evaluation
 reward: 5.94
 episode: 1440 score: 10.0 memory length: 100000 epsilon:
 0.5619428200095098 steps: 560 lr: 1.6000000000000003e-05 evaluation
 reward: 5.97
 episode: 1441 score: 10.0 memory length: 100000 epsilon:
 0.5612003200095259 steps: 375 lr: 1.6000000000000003e-05 evaluation
 reward: 5.96
 episode: 1442 score: 10.0 memory length: 100000 epsilon: 0.560184580009548
 steps: 513 lr: 1.6000000000000003e-05 evaluation reward: 6.0
 episode: 1443 score: 6.0 memory length: 100000 epsilon: 0.559446040009564
 steps: 373 lr: 1.6000000000000003e-05 evaluation reward: 6.01
 episode: 1444 score: 10.0 memory length: 100000 epsilon:
 0.5584699000095852 steps: 493 lr: 1.6000000000000003e-05 evaluation
 reward: 6.06
 episode: 1445 score: 10.0 memory length: 100000 epsilon:
 0.5573967400096085 steps: 542 lr: 1.6000000000000003e-05 evaluation
 reward: 6.13
 episode: 1446 score: 5.0 memory length: 100000 epsilon: 0.5567789800096219
 steps: 312 lr: 1.6000000000000003e-05 evaluation reward: 6.12
 episode: 1447 score: 8.0 memory length: 100000 epsilon: 0.5559355000096402
 steps: 426 lr: 1.6000000000000003e-05 evaluation reward: 6.13
 episode: 1448 score: 5.0 memory length: 100000 epsilon: 0.5553296200096534
 steps: 306 lr: 1.6000000000000003e-05 evaluation reward: 6.14
 episode: 1449 score: 10.0 memory length: 100000 epsilon:
 0.5544128800096733 steps: 463 lr: 1.6000000000000003e-05 evaluation
 reward: 6.21
 episode: 1450 score: 6.0 memory length: 100000 epsilon: 0.5536783000096892
 steps: 371 lr: 1.6000000000000003e-05 evaluation reward: 6.22
 episode: 1451 score: 3.0 memory length: 100000 epsilon: 0.5532526000096984
 steps: 215 lr: 1.6000000000000003e-05 evaluation reward: 6.2
 episode: 1452 score: 7.0 memory length: 100000 epsilon: 0.5524091200097168
 steps: 426 lr: 1.6000000000000003e-05 evaluation reward: 6.23
 episode: 1453 score: 3.0 memory length: 100000 epsilon: 0.5519200600097274

```

steps: 247    lr: 1.6000000000000003e-05    evaluation reward: 6.19
episode: 1454  score: 5.0    memory length: 100000    epsilon: 0.5512805200097413
steps: 323    lr: 1.6000000000000003e-05    evaluation reward: 6.19
episode: 1455  score: 6.0    memory length: 100000    epsilon: 0.5505776200097565
steps: 355    lr: 1.6000000000000003e-05    evaluation reward: 6.22
episode: 1456  score: 6.0    memory length: 100000    epsilon: 0.5497975000097735
steps: 394    lr: 1.6000000000000003e-05    evaluation reward: 6.24
episode: 1457  score: 5.0    memory length: 100000    epsilon: 0.5492094400097862
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 6.22
episode: 1458  score: 13.0   memory length: 100000    epsilon:
0.5481897400098084    steps: 515    lr: 1.6000000000000003e-05    evaluation
reward: 6.28
episode: 1459  score: 3.0    memory length: 100000    epsilon: 0.5477699800098175
steps: 212    lr: 1.6000000000000003e-05    evaluation reward: 6.28
episode: 1460  score: 8.0    memory length: 100000    epsilon: 0.5468334400098378
steps: 473    lr: 1.6000000000000003e-05    evaluation reward: 6.28
episode: 1461  score: 4.0    memory length: 100000    epsilon: 0.5462790400098498
steps: 280    lr: 1.6000000000000003e-05    evaluation reward: 6.24
episode: 1462  score: 4.0    memory length: 100000    epsilon: 0.5457345400098617
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 6.21
episode: 1463  score: 6.0    memory length: 100000    epsilon: 0.5449880800098779
steps: 377    lr: 1.6000000000000003e-05    evaluation reward: 6.22
episode: 1464  score: 6.0    memory length: 100000    epsilon: 0.5441960800098951
steps: 400    lr: 1.6000000000000003e-05    evaluation reward: 6.28
episode: 1465  score: 8.0    memory length: 100000    epsilon: 0.5433605200099132
steps: 422    lr: 1.6000000000000003e-05    evaluation reward: 6.25
episode: 1466  score: 6.0    memory length: 100000    epsilon: 0.5426180200099293
steps: 375    lr: 1.6000000000000003e-05    evaluation reward: 6.27
episode: 1467  score: 7.0    memory length: 100000    epsilon: 0.5418616600099457
steps: 382    lr: 1.6000000000000003e-05    evaluation reward: 6.32
episode: 1468  score: 4.0    memory length: 100000    epsilon: 0.5413171600099576
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 6.31
episode: 1469  score: 5.0    memory length: 100000    epsilon: 0.5406776200099714
steps: 323    lr: 1.6000000000000003e-05    evaluation reward: 6.29
episode: 1470  score: 6.0    memory length: 100000    epsilon: 0.5399747200099867
steps: 355    lr: 1.6000000000000003e-05    evaluation reward: 6.26
episode: 1471  score: 6.0    memory length: 100000    epsilon: 0.5393411200100005
steps: 320    lr: 1.6000000000000003e-05    evaluation reward: 6.27
episode: 1472  score: 7.0    memory length: 100000    epsilon: 0.5385709000100172
steps: 389    lr: 1.6000000000000003e-05    evaluation reward: 6.28
episode: 1473  score: 9.0    memory length: 100000    epsilon: 0.5377195000100357
steps: 430    lr: 1.6000000000000003e-05    evaluation reward: 6.28
episode: 1474  score: 7.0    memory length: 100000    epsilon: 0.5369552200100522
steps: 386    lr: 1.6000000000000003e-05    evaluation reward: 6.28
episode: 1475  score: 3.0    memory length: 100000    epsilon: 0.5365394200100613
steps: 210    lr: 1.6000000000000003e-05    evaluation reward: 6.28
episode: 1476  score: 11.0   memory length: 100000    epsilon: 0.535585060010082
steps: 482    lr: 1.6000000000000003e-05    evaluation reward: 6.34

```

episode: 1477 score: 7.0 memory length: 100000 epsilon: 0.5348524600100979
steps: 370 lr: 1.6000000000000003e-05 evaluation reward: 6.34
episode: 1478 score: 12.0 memory length: 100000 epsilon:
0.5339060200101184 steps: 478 lr: 1.6000000000000003e-05 evaluation
reward: 6.38
episode: 1479 score: 4.0 memory length: 100000 epsilon: 0.5333120200101313
steps: 300 lr: 1.6000000000000003e-05 evaluation reward: 6.33
episode: 1480 score: 10.0 memory length: 100000 epsilon:
0.5322487600101544 steps: 537 lr: 1.6000000000000003e-05 evaluation
reward: 6.35
episode: 1481 score: 3.0 memory length: 100000 epsilon: 0.531763660010165
steps: 245 lr: 1.6000000000000003e-05 evaluation reward: 6.35
episode: 1482 score: 3.0 memory length: 100000 epsilon: 0.5313082600101748
steps: 230 lr: 1.6000000000000003e-05 evaluation reward: 6.32
episode: 1483 score: 6.0 memory length: 100000 epsilon: 0.5305004200101924
steps: 408 lr: 1.6000000000000003e-05 evaluation reward: 6.31
episode: 1484 score: 7.0 memory length: 100000 epsilon: 0.5297024800102097
steps: 403 lr: 1.6000000000000003e-05 evaluation reward: 6.36
episode: 1485 score: 3.0 memory length: 100000 epsilon: 0.5292074800102204
steps: 250 lr: 1.6000000000000003e-05 evaluation reward: 6.38
episode: 1486 score: 11.0 memory length: 100000 epsilon:
0.5281343200102437 steps: 542 lr: 1.6000000000000003e-05 evaluation
reward: 6.46
episode: 1487 score: 11.0 memory length: 100000 epsilon:
0.5270888800102664 steps: 528 lr: 1.6000000000000003e-05 evaluation
reward: 6.54
episode: 1488 score: 2.0 memory length: 100000 epsilon: 0.526696840010275
steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 6.49
episode: 1489 score: 6.0 memory length: 100000 epsilon: 0.525958300010291
steps: 373 lr: 1.6000000000000003e-05 evaluation reward: 6.47
episode: 1490 score: 5.0 memory length: 100000 epsilon: 0.525312820010305
steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 6.48
episode: 1491 score: 11.0 memory length: 100000 epsilon:
0.5242951000103271 steps: 514 lr: 1.6000000000000003e-05 evaluation
reward: 6.53
episode: 1492 score: 5.0 memory length: 100000 epsilon: 0.5236912000103402
steps: 305 lr: 1.6000000000000003e-05 evaluation reward: 6.51
episode: 1493 score: 4.0 memory length: 100000 epsilon: 0.5231407600103521
steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 6.49
episode: 1494 score: 7.0 memory length: 100000 epsilon: 0.5223586600103691
steps: 395 lr: 1.6000000000000003e-05 evaluation reward: 6.48
episode: 1495 score: 7.0 memory length: 100000 epsilon: 0.5215112200103875
steps: 428 lr: 1.6000000000000003e-05 evaluation reward: 6.48
episode: 1496 score: 6.0 memory length: 100000 epsilon: 0.5207905000104032
steps: 364 lr: 1.6000000000000003e-05 evaluation reward: 6.53
episode: 1497 score: 6.0 memory length: 100000 epsilon: 0.5201371000104174
steps: 330 lr: 1.6000000000000003e-05 evaluation reward: 6.52
episode: 1498 score: 12.0 memory length: 100000 epsilon:

0.5188580200104451 steps: 646 lr: 1.6000000000000003e-05 evaluation
 reward: 6.56
 episode: 1499 score: 8.0 memory length: 100000 epsilon: 0.5180482000104627
 steps: 409 lr: 1.6000000000000003e-05 evaluation reward: 6.54
 episode: 1500 score: 5.0 memory length: 100000 epsilon: 0.5173750000104773
 steps: 340 lr: 1.6000000000000003e-05 evaluation reward: 6.56
 episode: 1501 score: 2.0 memory length: 100000 epsilon: 0.5169809800104859
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 6.46
 episode: 1502 score: 11.0 memory length: 100000 epsilon:
 0.5158900000105096 steps: 551 lr: 1.6000000000000003e-05 evaluation
 reward: 6.52
 episode: 1503 score: 5.0 memory length: 100000 epsilon: 0.5152445200105236
 steps: 326 lr: 1.6000000000000003e-05 evaluation reward: 6.51
 episode: 1504 score: 15.0 memory length: 100000 epsilon:
 0.5141099800105482 steps: 573 lr: 1.6000000000000003e-05 evaluation
 reward: 6.62
 episode: 1505 score: 8.0 memory length: 100000 epsilon: 0.5131378000105693
 steps: 491 lr: 1.6000000000000003e-05 evaluation reward: 6.63
 episode: 1506 score: 2.0 memory length: 100000 epsilon: 0.5127774400105771
 steps: 182 lr: 1.6000000000000003e-05 evaluation reward: 6.57
 episode: 1507 score: 8.0 memory length: 100000 epsilon: 0.5119636600105948
 steps: 411 lr: 1.6000000000000003e-05 evaluation reward: 6.6
 episode: 1508 score: 8.0 memory length: 100000 epsilon: 0.5111399800106127
 steps: 416 lr: 1.6000000000000003e-05 evaluation reward: 6.63
 episode: 1509 score: 6.0 memory length: 100000 epsilon: 0.5104588600106275
 steps: 344 lr: 1.6000000000000003e-05 evaluation reward: 6.68
 episode: 1510 score: 9.0 memory length: 100000 epsilon: 0.5095183600106479
 steps: 475 lr: 1.6000000000000003e-05 evaluation reward: 6.68
 episode: 1511 score: 9.0 memory length: 100000 epsilon: 0.5086115200106676
 steps: 458 lr: 1.6000000000000003e-05 evaluation reward: 6.75
 episode: 1512 score: 6.0 memory length: 100000 epsilon: 0.5079125800106827
 steps: 353 lr: 1.6000000000000003e-05 evaluation reward: 6.75
 episode: 1513 score: 9.0 memory length: 100000 epsilon: 0.5069859400107029
 steps: 468 lr: 1.6000000000000003e-05 evaluation reward: 6.82
 episode: 1514 score: 7.0 memory length: 100000 epsilon: 0.5062513600107188
 steps: 371 lr: 1.6000000000000003e-05 evaluation reward: 6.82
 episode: 1515 score: 8.0 memory length: 100000 epsilon: 0.5053544200107383
 steps: 453 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1516 score: 7.0 memory length: 100000 epsilon: 0.5045426200107559
 steps: 410 lr: 1.6000000000000003e-05 evaluation reward: 6.88
 episode: 1517 score: 10.0 memory length: 100000 epsilon:
 0.5035209400107781 steps: 516 lr: 1.6000000000000003e-05 evaluation
 reward: 6.91
 episode: 1518 score: 6.0 memory length: 100000 epsilon: 0.5028081400107935
 steps: 360 lr: 1.6000000000000003e-05 evaluation reward: 6.88
 episode: 1519 score: 8.0 memory length: 100000 epsilon: 0.5019805000108115
 steps: 418 lr: 1.6000000000000003e-05 evaluation reward: 6.89
 episode: 1520 score: 2.0 memory length: 100000 epsilon: 0.50158846001082

```

steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 6.85
episode: 1521  score: 8.0  memory length: 100000  epsilon: 0.5007232000108388
steps: 437    lr: 1.6000000000000003e-05    evaluation reward: 6.88
episode: 1522  score: 8.0  memory length: 100000  epsilon: 0.4998737800108537
steps: 429    lr: 1.6000000000000003e-05    evaluation reward: 6.9
episode: 1523  score: 3.0  memory length: 100000  epsilon:
0.49945798001085107  steps: 210  lr: 1.6000000000000003e-05    evaluation
reward: 6.82
episode: 1524  score: 7.0  memory length: 100000  epsilon: 0.4986897400108462
steps: 388    lr: 1.6000000000000003e-05    evaluation reward: 6.83
episode: 1525  score: 8.0  memory length: 100000  epsilon:
0.49785616001084093  steps: 421  lr: 1.6000000000000003e-05    evaluation
reward: 6.82
episode: 1526  score: 8.0  memory length: 100000  epsilon: 0.4970463400108358
steps: 409    lr: 1.6000000000000003e-05    evaluation reward: 6.85
episode: 1527  score: 7.0  memory length: 100000  epsilon:
0.49624642001083075  steps: 404  lr: 1.6000000000000003e-05    evaluation
reward: 6.87
episode: 1528  score: 6.0  memory length: 100000  epsilon:
0.49556926001082646  steps: 342  lr: 1.6000000000000003e-05    evaluation
reward: 6.91
episode: 1529  score: 4.0  memory length: 100000  epsilon: 0.4950564400108232
steps: 259    lr: 1.6000000000000003e-05    evaluation reward: 6.9
episode: 1530  score: 6.0  memory length: 100000  epsilon: 0.4943773000108189
steps: 343    lr: 1.6000000000000003e-05    evaluation reward: 6.81
episode: 1531  score: 2.0  memory length: 100000  epsilon:
0.49398328001081643  steps: 199  lr: 1.6000000000000003e-05    evaluation
reward: 6.72
episode: 1532  score: 1.0  memory length: 100000  epsilon:
0.49368628001081455  steps: 150  lr: 1.6000000000000003e-05    evaluation
reward: 6.69
episode: 1533  score: 11.0  memory length: 100000  epsilon:
0.4927576600108087  steps: 469  lr: 1.6000000000000003e-05    evaluation
reward: 6.69
episode: 1534  score: 3.0  memory length: 100000  epsilon:
0.49226662001080557  steps: 248  lr: 1.6000000000000003e-05    evaluation
reward: 6.69
episode: 1535  score: 10.0  memory length: 100000  epsilon: 0.491225140010799
steps: 526    lr: 1.6000000000000003e-05    evaluation reward: 6.73
episode: 1536  score: 7.0  memory length: 100000  epsilon: 0.4905024400107944
steps: 365    lr: 1.6000000000000003e-05    evaluation reward: 6.76
episode: 1537  score: 16.0  memory length: 100000  epsilon:
0.4889501200107846  steps: 784  lr: 1.6000000000000003e-05    evaluation
reward: 6.88
episode: 1538  score: 4.0  memory length: 100000  epsilon: 0.4883640400107809
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 6.87
episode: 1539  score: 4.0  memory length: 100000  epsilon: 0.4878433000107776
steps: 263    lr: 1.6000000000000003e-05    evaluation reward: 6.78

```

episode: 1540 score: 14.0 memory length: 100000 epsilon:
 0.4865464000107694 steps: 655 lr: 1.6000000000000003e-05 evaluation
 reward: 6.82
 episode: 1541 score: 9.0 memory length: 100000 epsilon: 0.4857247000107642
 steps: 415 lr: 1.6000000000000003e-05 evaluation reward: 6.81
 episode: 1542 score: 2.0 memory length: 100000 epsilon: 0.4853326600107617
 steps: 198 lr: 1.6000000000000003e-05 evaluation reward: 6.73
 episode: 1543 score: 9.0 memory length: 100000 epsilon: 0.4843525600107555
 steps: 495 lr: 1.6000000000000003e-05 evaluation reward: 6.76
 episode: 1544 score: 11.0 memory length: 100000 epsilon:
 0.4832140600107483 steps: 575 lr: 1.6000000000000003e-05 evaluation
 reward: 6.77
 episode: 1545 score: 5.0 memory length: 100000 epsilon: 0.4825982800107444
 steps: 311 lr: 1.6000000000000003e-05 evaluation reward: 6.72
 episode: 1546 score: 6.0 memory length: 100000 epsilon:
 0.48186568001073976 steps: 370 lr: 1.6000000000000003e-05 evaluation
 reward: 6.73
 episode: 1547 score: 10.0 memory length: 100000 epsilon:
 0.48094498001073394 steps: 465 lr: 1.6000000000000003e-05 evaluation
 reward: 6.75
 episode: 1548 score: 10.0 memory length: 100000 epsilon: 0.480010420010728
 steps: 472 lr: 1.6000000000000003e-05 evaluation reward: 6.8
 episode: 1549 score: 4.0 memory length: 100000 epsilon:
 0.47949166001072474 steps: 262 lr: 1.6000000000000003e-05 evaluation
 reward: 6.74
 episode: 1550 score: 8.0 memory length: 100000 epsilon:
 0.47860660001071914 steps: 447 lr: 1.6000000000000003e-05 evaluation
 reward: 6.76
 episode: 1551 score: 12.0 memory length: 100000 epsilon: 0.477634420010713
 steps: 491 lr: 1.6000000000000003e-05 evaluation reward: 6.85
 episode: 1552 score: 10.0 memory length: 100000 epsilon:
 0.4767196600107072 steps: 462 lr: 1.6000000000000003e-05 evaluation
 reward: 6.88
 episode: 1553 score: 7.0 memory length: 100000 epsilon:
 0.47591974001070214 steps: 404 lr: 1.6000000000000003e-05 evaluation
 reward: 6.92
 episode: 1554 score: 9.0 memory length: 100000 epsilon:
 0.47495944001069607 steps: 485 lr: 1.6000000000000003e-05 evaluation
 reward: 6.96
 episode: 1555 score: 7.0 memory length: 100000 epsilon: 0.474159520010691
 steps: 404 lr: 1.6000000000000003e-05 evaluation reward: 6.97
 episode: 1556 score: 6.0 memory length: 100000 epsilon: 0.4734328600106864
 steps: 367 lr: 1.6000000000000003e-05 evaluation reward: 6.97
 episode: 1557 score: 4.0 memory length: 100000 epsilon:
 0.47288638001068295 steps: 276 lr: 1.6000000000000003e-05 evaluation
 reward: 6.96
 episode: 1558 score: 5.0 memory length: 100000 epsilon:
 0.47228446001067914 steps: 304 lr: 1.6000000000000003e-05 evaluation

```

reward: 6.88
episode: 1559  score: 4.0  memory length: 100000  epsilon: 0.4717399600106757
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 6.89
episode: 1560  score: 5.0  memory length: 100000  epsilon:
0.47113408001067186  steps: 306    lr: 1.6000000000000003e-05    evaluation
reward: 6.86
episode: 1561  score: 6.0  memory length: 100000  epsilon:
0.47049652001066783  steps: 322    lr: 1.6000000000000003e-05    evaluation
reward: 6.88
episode: 1562  score: 4.0  memory length: 100000  epsilon:
0.46991242001066413  steps: 295    lr: 1.6000000000000003e-05    evaluation
reward: 6.88
episode: 1563  score: 7.0  memory length: 100000  epsilon: 0.4690828000106589
steps: 419    lr: 1.6000000000000003e-05    evaluation reward: 6.89
episode: 1564  score: 6.0  memory length: 100000  epsilon: 0.4683442600106542
steps: 373    lr: 1.6000000000000003e-05    evaluation reward: 6.89
episode: 1565  score: 11.0  memory length: 100000  epsilon:
0.4675027600106489  steps: 425    lr: 1.6000000000000003e-05    evaluation
reward: 6.92
episode: 1566  score: 5.0  memory length: 100000  epsilon:
0.46681768001064455  steps: 346    lr: 1.6000000000000003e-05    evaluation
reward: 6.91
episode: 1567  score: 4.0  memory length: 100000  epsilon:
0.46626130001064103  steps: 281    lr: 1.6000000000000003e-05    evaluation
reward: 6.88
episode: 1568  score: 5.0  memory length: 100000  epsilon: 0.4656098800106369
steps: 329    lr: 1.6000000000000003e-05    evaluation reward: 6.89
episode: 1569  score: 7.0  memory length: 100000  epsilon: 0.4647723400106316
steps: 423    lr: 1.6000000000000003e-05    evaluation reward: 6.91
episode: 1570  score: 12.0  memory length: 100000  epsilon:
0.4639288600106263  steps: 426    lr: 1.6000000000000003e-05    evaluation
reward: 6.97
episode: 1571  score: 4.0  memory length: 100000  epsilon:
0.46338436001062283  steps: 275    lr: 1.6000000000000003e-05    evaluation
reward: 6.95
episode: 1572  score: 9.0  memory length: 100000  epsilon: 0.4624181200106167
steps: 488    lr: 1.6000000000000003e-05    evaluation reward: 6.97
episode: 1573  score: 10.0  memory length: 100000  epsilon:
0.46139842001061027  steps: 515    lr: 1.6000000000000003e-05    evaluation
reward: 6.98
episode: 1574  score: 4.0  memory length: 100000  epsilon: 0.4608519400106068
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 6.95
episode: 1575  score: 6.0  memory length: 100000  epsilon: 0.4601054800106021
steps: 377    lr: 1.6000000000000003e-05    evaluation reward: 6.98
episode: 1576  score: 6.0  memory length: 100000  epsilon: 0.4594322800105978
steps: 340    lr: 1.6000000000000003e-05    evaluation reward: 6.93
episode: 1577  score: 5.0  memory length: 100000  epsilon: 0.458826400010594
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 6.91

```

episode: 1578 score: 10.0 memory length: 100000 epsilon:
 0.45796510001058854 steps: 435 lr: 1.6000000000000003e-05 evaluation
 reward: 6.89
 episode: 1579 score: 6.0 memory length: 100000 epsilon: 0.4572958600105843
 steps: 338 lr: 1.6000000000000003e-05 evaluation reward: 6.91
 episode: 1580 score: 7.0 memory length: 100000 epsilon:
 0.45655732001057964 steps: 373 lr: 1.6000000000000003e-05 evaluation
 reward: 6.88
 episode: 1581 score: 14.0 memory length: 100000 epsilon:
 0.45521884001057117 steps: 676 lr: 1.6000000000000003e-05 evaluation
 reward: 6.99
 episode: 1582 score: 5.0 memory length: 100000 epsilon:
 0.45457138001056707 steps: 327 lr: 1.6000000000000003e-05 evaluation
 reward: 7.01
 episode: 1583 score: 11.0 memory length: 100000 epsilon:
 0.4534408000105599 steps: 571 lr: 1.6000000000000003e-05 evaluation
 reward: 7.06
 episode: 1584 score: 12.0 memory length: 100000 epsilon:
 0.4526329600105548 steps: 408 lr: 1.6000000000000003e-05 evaluation
 reward: 7.11
 episode: 1585 score: 7.0 memory length: 100000 epsilon: 0.4518409600105498
 steps: 400 lr: 1.6000000000000003e-05 evaluation reward: 7.15
 episode: 1586 score: 7.0 memory length: 100000 epsilon: 0.451078660010545
 steps: 385 lr: 1.6000000000000003e-05 evaluation reward: 7.11
 episode: 1587 score: 9.0 memory length: 100000 epsilon: 0.4503064600105401
 steps: 390 lr: 1.6000000000000003e-05 evaluation reward: 7.09
 episode: 1588 score: 6.0 memory length: 100000 epsilon: 0.4495996000105356
 steps: 357 lr: 1.6000000000000003e-05 evaluation reward: 7.13
 episode: 1589 score: 7.0 memory length: 100000 epsilon:
 0.44876602001053034 steps: 421 lr: 1.6000000000000003e-05 evaluation
 reward: 7.14
 episode: 1590 score: 11.0 memory length: 100000 epsilon:
 0.44769088001052354 steps: 543 lr: 1.6000000000000003e-05 evaluation
 reward: 7.2
 episode: 1591 score: 12.0 memory length: 100000 epsilon:
 0.44680978001051797 steps: 445 lr: 1.6000000000000003e-05 evaluation
 reward: 7.21
 episode: 1592 score: 9.0 memory length: 100000 epsilon:
 0.44577820001051144 steps: 521 lr: 1.6000000000000003e-05 evaluation
 reward: 7.25
 episode: 1593 score: 6.0 memory length: 100000 epsilon: 0.4449545200105062
 steps: 416 lr: 1.6000000000000003e-05 evaluation reward: 7.27
 episode: 1594 score: 5.0 memory length: 100000 epsilon:
 0.44434270001050236 steps: 309 lr: 1.6000000000000003e-05 evaluation
 reward: 7.25
 episode: 1595 score: 8.0 memory length: 100000 epsilon: 0.443495260010497
 steps: 428 lr: 1.6000000000000003e-05 evaluation reward: 7.26
 episode: 1596 score: 6.0 memory length: 100000 epsilon: 0.4427983000104926

```

steps: 352    lr: 1.6000000000000003e-05    evaluation reward: 7.26
episode: 1597  score: 10.0  memory length: 100000  epsilon:
0.44172118001048577  steps: 544    lr: 1.6000000000000003e-05    evaluation
reward: 7.3
episode: 1598  score: 8.0  memory length: 100000  epsilon:
0.44094700001048087  steps: 391    lr: 1.6000000000000003e-05    evaluation
reward: 7.26
episode: 1599  score: 3.0  memory length: 100000  epsilon: 0.4404619000104778
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 7.21
episode: 1600  score: 5.0  memory length: 100000  epsilon:
0.43978672001047353  steps: 341    lr: 1.6000000000000003e-05    evaluation
reward: 7.21
episode: 1601  score: 10.0  memory length: 100000  epsilon:
0.4388363200104675  steps: 480    lr: 1.6000000000000003e-05    evaluation
reward: 7.29
episode: 1602  score: 6.0  memory length: 100000  epsilon: 0.4380938200104628
steps: 375    lr: 1.6000000000000003e-05    evaluation reward: 7.24
episode: 1603  score: 8.0  memory length: 100000  epsilon:
0.43722658001045733  steps: 438    lr: 1.6000000000000003e-05    evaluation
reward: 7.27
episode: 1604  score: 6.0  memory length: 100000  epsilon:
0.43651972001045286  steps: 357    lr: 1.6000000000000003e-05    evaluation
reward: 7.18
episode: 1605  score: 7.0  memory length: 100000  epsilon: 0.4356386200104473
steps: 445    lr: 1.6000000000000003e-05    evaluation reward: 7.17
episode: 1606  score: 9.0  memory length: 100000  epsilon:
0.43498522001044315  steps: 330    lr: 1.6000000000000003e-05    evaluation
reward: 7.24
episode: 1607  score: 9.0  memory length: 100000  epsilon: 0.434011060010437
steps: 492    lr: 1.6000000000000003e-05    evaluation reward: 7.25
episode: 1608  score: 5.0  memory length: 100000  epsilon:
0.43340320001043314  steps: 307    lr: 1.6000000000000003e-05    evaluation
reward: 7.22
episode: 1609  score: 9.0  memory length: 100000  epsilon: 0.432433000010427
steps: 490    lr: 1.6000000000000003e-05    evaluation reward: 7.25
episode: 1610  score: 7.0  memory length: 100000  epsilon:
0.43164694001042203  steps: 397    lr: 1.6000000000000003e-05    evaluation
reward: 7.23
episode: 1611  score: 6.0  memory length: 100000  epsilon:
0.43094008001041756  steps: 357    lr: 1.6000000000000003e-05    evaluation
reward: 7.2
episode: 1612  score: 6.0  memory length: 100000  epsilon:
0.43030648001041355  steps: 320    lr: 1.6000000000000003e-05    evaluation
reward: 7.2
episode: 1613  score: 5.0  memory length: 100000  epsilon:
0.42974020001040997  steps: 286    lr: 1.6000000000000003e-05    evaluation
reward: 7.16
episode: 1614  score: 7.0  memory length: 100000  epsilon: 0.4289204800104048

```

```

steps: 414    lr: 1.6000000000000003e-05    evaluation reward: 7.16
episode: 1615  score: 7.0  memory length: 100000  epsilon: 0.4281779800104001
steps: 375    lr: 1.6000000000000003e-05    evaluation reward: 7.15
episode: 1616  score: 4.0  memory length: 100000  epsilon:
0.42758992001039636  steps: 297  lr: 1.6000000000000003e-05    evaluation
reward: 7.12
episode: 1617  score: 6.0  memory length: 100000  epsilon: 0.4268058400103914
steps: 396    lr: 1.6000000000000003e-05    evaluation reward: 7.08
episode: 1618  score: 11.0  memory length: 100000  epsilon:
0.4258059400103851  steps: 505  lr: 1.6000000000000003e-05    evaluation
reward: 7.13
episode: 1619  score: 3.0  memory length: 100000  epsilon: 0.4253842000103824
steps: 213    lr: 1.6000000000000003e-05    evaluation reward: 7.08
episode: 1620  score: 10.0  memory length: 100000  epsilon:
0.42434470001037583  steps: 525  lr: 1.6000000000000003e-05    evaluation
reward: 7.16
episode: 1621  score: 8.0  memory length: 100000  epsilon: 0.4234556800103702
steps: 449    lr: 1.6000000000000003e-05    evaluation reward: 7.16
episode: 1622  score: 13.0  memory length: 100000  epsilon: 0.422315200010363
steps: 576    lr: 1.6000000000000003e-05    evaluation reward: 7.21
episode: 1623  score: 5.0  memory length: 100000  epsilon: 0.4217647600103595
steps: 278    lr: 1.6000000000000003e-05    evaluation reward: 7.23
episode: 1624  score: 4.0  memory length: 100000  epsilon:
0.42117076001035575  steps: 300  lr: 1.6000000000000003e-05    evaluation
reward: 7.2
episode: 1625  score: 4.0  memory length: 100000  epsilon:
0.42061438001035223  steps: 281  lr: 1.6000000000000003e-05    evaluation
reward: 7.16
episode: 1626  score: 4.0  memory length: 100000  epsilon:
0.42010948001034903  steps: 255  lr: 1.6000000000000003e-05    evaluation
reward: 7.12
episode: 1627  score: 10.0  memory length: 100000  epsilon:
0.41918086001034316  steps: 469  lr: 1.6000000000000003e-05    evaluation
reward: 7.15
episode: 1628  score: 5.0  memory length: 100000  epsilon:
0.41849974001033885  steps: 344  lr: 1.6000000000000003e-05    evaluation
reward: 7.14
episode: 1629  score: 6.0  memory length: 100000  epsilon:
0.41782258001033457  steps: 342  lr: 1.6000000000000003e-05    evaluation
reward: 7.16
episode: 1630  score: 9.0  memory length: 100000  epsilon: 0.4168919800103287
steps: 470    lr: 1.6000000000000003e-05    evaluation reward: 7.19
episode: 1631  score: 11.0  memory length: 100000  epsilon:
0.4159118800103225  steps: 495  lr: 1.6000000000000003e-05    evaluation
reward: 7.28
episode: 1632  score: 6.0  memory length: 100000  epsilon: 0.415205020010318
steps: 357    lr: 1.6000000000000003e-05    evaluation reward: 7.33
episode: 1633  score: 4.0  memory length: 100000  epsilon: 0.4146486400103145

```

```

steps: 281    lr: 1.6000000000000003e-05    evaluation reward: 7.26
episode: 1634  score: 7.0  memory length: 100000  epsilon: 0.4138091200103092
steps: 424    lr: 1.6000000000000003e-05    evaluation reward: 7.3
episode: 1635  score: 9.0  memory length: 100000  epsilon: 0.4129478200103037
steps: 435    lr: 1.6000000000000003e-05    evaluation reward: 7.29
episode: 1636  score: 11.0  memory length: 100000  epsilon:
0.41203702001029796  steps: 460  lr: 1.6000000000000003e-05    evaluation
reward: 7.33
episode: 1637  score: 3.0  memory length: 100000  epsilon: 0.4115539000102949
steps: 244    lr: 1.6000000000000003e-05    evaluation reward: 7.2
episode: 1638  score: 9.0  memory length: 100000  epsilon:
0.41080150001029014  steps: 380  lr: 1.6000000000000003e-05    evaluation
reward: 7.25
episode: 1639  score: 6.0  memory length: 100000  epsilon: 0.4100847400102856
steps: 362    lr: 1.6000000000000003e-05    evaluation reward: 7.27
episode: 1640  score: 7.0  memory length: 100000  epsilon: 0.4092967000102806
steps: 398    lr: 1.6000000000000003e-05    evaluation reward: 7.2
episode: 1641  score: 8.0  memory length: 100000  epsilon: 0.4085165800102757
steps: 394    lr: 1.6000000000000003e-05    evaluation reward: 7.19
episode: 1642  score: 14.0  memory length: 100000  epsilon:
0.4074394600102689  steps: 544  lr: 1.6000000000000003e-05    evaluation
reward: 7.31
episode: 1643  score: 6.0  memory length: 100000  epsilon: 0.4067207200102643
steps: 363    lr: 1.6000000000000003e-05    evaluation reward: 7.28
episode: 1644  score: 3.0  memory length: 100000  epsilon: 0.4062237400102612
steps: 251    lr: 1.6000000000000003e-05    evaluation reward: 7.2
episode: 1645  score: 10.0  memory length: 100000  epsilon:
0.4052139400102548  steps: 510  lr: 6.4000000000000001e-06    evaluation
reward: 7.25
episode: 1646  score: 5.0  memory length: 100000  epsilon:
0.40457242001025073  steps: 324  lr: 6.4000000000000001e-06    evaluation
reward: 7.24
episode: 1647  score: 21.0  memory length: 100000  epsilon:
0.40332700001024285  steps: 629  lr: 6.4000000000000001e-06    evaluation
reward: 7.35
episode: 1648  score: 9.0  memory length: 100000  epsilon: 0.4023884800102369
steps: 474    lr: 6.4000000000000001e-06    evaluation reward: 7.34
episode: 1649  score: 4.0  memory length: 100000  epsilon: 0.4018360600102334
steps: 279    lr: 6.4000000000000001e-06    evaluation reward: 7.34
episode: 1650  score: 8.0  memory length: 100000  epsilon: 0.4009925800102281
steps: 426    lr: 6.4000000000000001e-06    evaluation reward: 7.34
episode: 1651  score: 14.0  memory length: 100000  epsilon:
0.3999709000102216  steps: 516  lr: 6.4000000000000001e-06    evaluation
reward: 7.36
episode: 1652  score: 14.0  memory length: 100000  epsilon:
0.39882052001021434  steps: 581  lr: 6.4000000000000001e-06    evaluation
reward: 7.4
episode: 1653  score: 7.0  memory length: 100000  epsilon: 0.3980681200102096

```



```

steps: 380    lr: 6.4000000000000001e-06    evaluation reward: 7.4
episode: 1654    score: 12.0    memory length: 100000    epsilon:
0.3968821000102021    steps: 599    lr: 6.4000000000000001e-06    evaluation
reward: 7.43
episode: 1655    score: 5.0    memory length: 100000    epsilon: 0.3962663200101982
steps: 311    lr: 6.4000000000000001e-06    evaluation reward: 7.41
episode: 1656    score: 8.0    memory length: 100000    epsilon: 0.3954307600101929
steps: 422    lr: 6.4000000000000001e-06    evaluation reward: 7.43
episode: 1657    score: 7.0    memory length: 100000    epsilon:
0.39471598001018837    steps: 361    lr: 6.4000000000000001e-06    evaluation
reward: 7.46
episode: 1658    score: 2.0    memory length: 100000    epsilon:
0.39428236001018563    steps: 219    lr: 6.4000000000000001e-06    evaluation
reward: 7.43
episode: 1659    score: 7.0    memory length: 100000    epsilon: 0.3934705600101805
steps: 410    lr: 6.4000000000000001e-06    evaluation reward: 7.46
episode: 1660    score: 7.0    memory length: 100000    epsilon: 0.3926330200101752
steps: 423    lr: 6.4000000000000001e-06    evaluation reward: 7.48
episode: 1661    score: 9.0    memory length: 100000    epsilon: 0.3917638000101697
steps: 439    lr: 6.4000000000000001e-06    evaluation reward: 7.51
episode: 1662    score: 4.0    memory length: 100000    epsilon:
0.39121930001016625    steps: 275    lr: 6.4000000000000001e-06    evaluation
reward: 7.51
episode: 1663    score: 4.0    memory length: 100000    epsilon: 0.390702520010163
steps: 261    lr: 6.4000000000000001e-06    evaluation reward: 7.48
episode: 1664    score: 3.0    memory length: 100000    epsilon:
0.39025504001016015    steps: 226    lr: 6.4000000000000001e-06    evaluation
reward: 7.45
episode: 1665    score: 7.0    memory length: 100000    epsilon: 0.3894590800101551
steps: 402    lr: 6.4000000000000001e-06    evaluation reward: 7.41
episode: 1666    score: 11.0    memory length: 100000    epsilon:
0.3884294800101486    steps: 520    lr: 6.4000000000000001e-06    evaluation
reward: 7.47
episode: 1667    score: 4.0    memory length: 100000    epsilon:
0.38788498001014515    steps: 275    lr: 6.4000000000000001e-06    evaluation
reward: 7.47
episode: 1668    score: 9.0    memory length: 100000    epsilon: 0.3869425000101392
steps: 476    lr: 6.4000000000000001e-06    evaluation reward: 7.51
episode: 1669    score: 16.0    memory length: 100000    epsilon:
0.38579806001013195    steps: 578    lr: 6.4000000000000001e-06    evaluation
reward: 7.6
episode: 1670    score: 6.0    memory length: 100000    epsilon:
0.38512090001012766    steps: 342    lr: 6.4000000000000001e-06    evaluation
reward: 7.54
episode: 1671    score: 10.0    memory length: 100000    epsilon:
0.38409328001012116    steps: 519    lr: 6.4000000000000001e-06    evaluation
reward: 7.6
episode: 1672    score: 21.0    memory length: 100000    epsilon:

```

0.38274688001011264 steps: 680 lr: 6.4000000000000001e-06 evaluation
 reward: 7.72
 episode: 1673 score: 9.0 memory length: 100000 epsilon:
 0.38185984001010703 steps: 448 lr: 6.4000000000000001e-06 evaluation
 reward: 7.71
 episode: 1674 score: 9.0 memory length: 100000 epsilon:
 0.38094310001010123 steps: 463 lr: 6.4000000000000001e-06 evaluation
 reward: 7.76
 episode: 1675 score: 14.0 memory length: 100000 epsilon:
 0.37969570001009334 steps: 630 lr: 6.4000000000000001e-06 evaluation
 reward: 7.84
 episode: 1676 score: 8.0 memory length: 100000 epsilon:
 0.37886014001008805 steps: 422 lr: 6.4000000000000001e-06 evaluation
 reward: 7.86
 episode: 1677 score: 7.0 memory length: 100000 epsilon:
 0.37813348001008346 steps: 367 lr: 6.4000000000000001e-06 evaluation
 reward: 7.88
 episode: 1678 score: 10.0 memory length: 100000 epsilon:
 0.3772088200100776 steps: 467 lr: 6.4000000000000001e-06 evaluation
 reward: 7.88
 episode: 1679 score: 12.0 memory length: 100000 epsilon:
 0.3760247800100701 steps: 598 lr: 6.4000000000000001e-06 evaluation
 reward: 7.94
 episode: 1680 score: 8.0 memory length: 100000 epsilon:
 0.37513180001006446 steps: 451 lr: 6.4000000000000001e-06 evaluation
 reward: 7.95
 episode: 1681 score: 2.0 memory length: 100000 epsilon:
 0.37473184001006193 steps: 202 lr: 6.4000000000000001e-06 evaluation
 reward: 7.83
 episode: 1682 score: 9.0 memory length: 100000 epsilon:
 0.37381906001005616 steps: 461 lr: 6.4000000000000001e-06 evaluation
 reward: 7.87
 episode: 1683 score: 5.0 memory length: 100000 epsilon:
 0.37316764001005204 steps: 329 lr: 6.4000000000000001e-06 evaluation
 reward: 7.81
 episode: 1684 score: 8.0 memory length: 100000 epsilon: 0.3722311000100461
 steps: 473 lr: 6.4000000000000001e-06 evaluation reward: 7.77
 episode: 1685 score: 1.0 memory length: 100000 epsilon: 0.3719281600100442
 steps: 153 lr: 6.4000000000000001e-06 evaluation reward: 7.71
 episode: 1686 score: 7.0 memory length: 100000 epsilon:
 0.37119556001003956 steps: 370 lr: 6.4000000000000001e-06 evaluation
 reward: 7.71
 episode: 1687 score: 3.0 memory length: 100000 epsilon:
 0.37070452001003645 steps: 248 lr: 6.4000000000000001e-06 evaluation
 reward: 7.65
 episode: 1688 score: 13.0 memory length: 100000 epsilon:
 0.36947296001002866 steps: 622 lr: 6.4000000000000001e-06 evaluation
 reward: 7.72

episode: 1689 score: 7.0 memory length: 100000 epsilon:
 0.36867898001002364 steps: 401 lr: 6.4000000000000001e-06 evaluation
 reward: 7.72
 episode: 1690 score: 12.0 memory length: 100000 epsilon:
 0.36758998001001675 steps: 550 lr: 6.4000000000000001e-06 evaluation
 reward: 7.73
 episode: 1691 score: 9.0 memory length: 100000 epsilon: 0.3667108600100112
 steps: 444 lr: 6.4000000000000001e-06 evaluation reward: 7.7
 episode: 1692 score: 2.0 memory length: 100000 epsilon: 0.3663188200100087
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 7.63
 episode: 1693 score: 8.0 memory length: 100000 epsilon: 0.3654476200100032
 steps: 440 lr: 6.4000000000000001e-06 evaluation reward: 7.65
 episode: 1694 score: 17.0 memory length: 100000 epsilon:
 0.3641388400099949 steps: 661 lr: 6.4000000000000001e-06 evaluation
 reward: 7.77
 episode: 1695 score: 10.0 memory length: 100000 epsilon:
 0.3630894400099883 steps: 530 lr: 6.4000000000000001e-06 evaluation
 reward: 7.79
 episode: 1696 score: 7.0 memory length: 100000 epsilon:
 0.36229348000998324 steps: 402 lr: 6.4000000000000001e-06 evaluation
 reward: 7.8
 episode: 1697 score: 8.0 memory length: 100000 epsilon: 0.3614222800099777
 steps: 440 lr: 6.4000000000000001e-06 evaluation reward: 7.78
 episode: 1698 score: 13.0 memory length: 100000 epsilon:
 0.36017488000996983 steps: 630 lr: 6.4000000000000001e-06 evaluation
 reward: 7.83
 episode: 1699 score: 9.0 memory length: 100000 epsilon:
 0.35933932000996455 steps: 422 lr: 6.4000000000000001e-06 evaluation
 reward: 7.89
 episode: 1700 score: 12.0 memory length: 100000 epsilon: 0.358143400009957
 steps: 604 lr: 6.4000000000000001e-06 evaluation reward: 7.96
 episode: 1701 score: 8.0 memory length: 100000 epsilon: 0.3572167600099511
 steps: 468 lr: 6.4000000000000001e-06 evaluation reward: 7.94
 episode: 1702 score: 11.0 memory length: 100000 epsilon:
 0.35625844000994505 steps: 484 lr: 6.4000000000000001e-06 evaluation
 reward: 7.99
 episode: 1703 score: 4.0 memory length: 100000 epsilon:
 0.35570800000994157 steps: 278 lr: 6.4000000000000001e-06 evaluation
 reward: 7.95
 episode: 1704 score: 4.0 memory length: 100000 epsilon: 0.3551892400099383
 steps: 262 lr: 6.4000000000000001e-06 evaluation reward: 7.93
 episode: 1705 score: 9.0 memory length: 100000 epsilon: 0.3542546800099324
 steps: 472 lr: 6.4000000000000001e-06 evaluation reward: 7.95
 episode: 1706 score: 7.0 memory length: 100000 epsilon: 0.3535141600099277
 steps: 374 lr: 6.4000000000000001e-06 evaluation reward: 7.93
 episode: 1707 score: 8.0 memory length: 100000 epsilon:
 0.35265088000992223 steps: 436 lr: 6.4000000000000001e-06 evaluation
 reward: 7.92

episode: 1708 score: 13.0 memory length: 100000 epsilon:
 0.3515203000099151 steps: 571 lr: 6.4000000000000001e-06 evaluation
 reward: 8.0
 episode: 1709 score: 5.0 memory length: 100000 epsilon:
 0.35094610000991144 steps: 290 lr: 6.4000000000000001e-06 evaluation
 reward: 7.96
 episode: 1710 score: 7.0 memory length: 100000 epsilon:
 0.35015608000990645 steps: 399 lr: 6.4000000000000001e-06 evaluation
 reward: 7.96
 episode: 1711 score: 5.0 memory length: 100000 epsilon:
 0.34947694000990215 steps: 343 lr: 6.4000000000000001e-06 evaluation
 reward: 7.95
 episode: 1712 score: 12.0 memory length: 100000 epsilon:
 0.3483602200098951 steps: 564 lr: 6.4000000000000001e-06 evaluation
 reward: 8.01
 episode: 1713 score: 8.0 memory length: 100000 epsilon:
 0.34753060000988983 steps: 419 lr: 6.4000000000000001e-06 evaluation
 reward: 8.04
 episode: 1714 score: 4.0 memory length: 100000 epsilon: 0.3469742200098863
 steps: 281 lr: 6.4000000000000001e-06 evaluation reward: 8.01
 episode: 1715 score: 11.0 memory length: 100000 epsilon:
 0.3458614600098793 steps: 562 lr: 6.4000000000000001e-06 evaluation
 reward: 8.05
 episode: 1716 score: 6.0 memory length: 100000 epsilon: 0.3451585600098748
 steps: 355 lr: 6.4000000000000001e-06 evaluation reward: 8.07
 episode: 1717 score: 7.0 memory length: 100000 epsilon: 0.3444893200098706
 steps: 338 lr: 6.4000000000000001e-06 evaluation reward: 8.08
 episode: 1718 score: 8.0 memory length: 100000 epsilon: 0.3436537600098653
 steps: 422 lr: 6.4000000000000001e-06 evaluation reward: 8.05
 episode: 1719 score: 20.0 memory length: 100000 epsilon:
 0.34211332000985556 steps: 778 lr: 6.4000000000000001e-06 evaluation
 reward: 8.22
 episode: 1720 score: 8.0 memory length: 100000 epsilon: 0.34123816000985
 steps: 442 lr: 6.4000000000000001e-06 evaluation reward: 8.2
 episode: 1721 score: 5.0 memory length: 100000 epsilon:
 0.34059664000984596 steps: 324 lr: 6.4000000000000001e-06 evaluation
 reward: 8.17
 episode: 1722 score: 4.0 memory length: 100000 epsilon:
 0.34011748000984293 steps: 242 lr: 6.4000000000000001e-06 evaluation
 reward: 8.08
 episode: 1723 score: 10.0 memory length: 100000 epsilon:
 0.33915520000983684 steps: 486 lr: 6.4000000000000001e-06 evaluation
 reward: 8.13
 episode: 1724 score: 5.0 memory length: 100000 epsilon:
 0.33858496000983324 steps: 288 lr: 6.4000000000000001e-06 evaluation
 reward: 8.14
 episode: 1725 score: 8.0 memory length: 100000 epsilon:
 0.33768604000982755 steps: 454 lr: 6.4000000000000001e-06 evaluation

```

reward: 8.18
episode: 1726  score: 5.0  memory length: 100000  epsilon: 0.3370346200098234
steps: 329  lr: 6.4000000000000001e-06  evaluation reward: 8.19
episode: 1727  score: 11.0  memory length: 100000  epsilon: 0.336022840009817
steps: 511  lr: 6.4000000000000001e-06  evaluation reward: 8.2
episode: 1728  score: 8.0  memory length: 100000  epsilon: 0.3352110400098119
steps: 410  lr: 6.4000000000000001e-06  evaluation reward: 8.23
episode: 1729  score: 4.0  memory length: 100000  epsilon: 0.3346606000098084
steps: 278  lr: 6.4000000000000001e-06  evaluation reward: 8.21
episode: 1730  score: 7.0  memory length: 100000  epsilon:
0.33389236000980355  steps: 388  lr: 6.4000000000000001e-06  evaluation
reward: 8.19
episode: 1731  score: 8.0  memory length: 100000  epsilon: 0.3330607600097983
steps: 420  lr: 6.4000000000000001e-06  evaluation reward: 8.16
episode: 1732  score: 5.0  memory length: 100000  epsilon:
0.33243904000979435  steps: 314  lr: 6.4000000000000001e-06  evaluation
reward: 8.15
episode: 1733  score: 9.0  memory length: 100000  epsilon: 0.3314807200097883
steps: 484  lr: 6.4000000000000001e-06  evaluation reward: 8.2
episode: 1734  score: 15.0  memory length: 100000  epsilon:
0.33032044000978095  steps: 586  lr: 6.4000000000000001e-06  evaluation
reward: 8.28
episode: 1735  score: 13.0  memory length: 100000  epsilon:
0.32900968000977265  steps: 662  lr: 6.4000000000000001e-06  evaluation
reward: 8.32
episode: 1736  score: 8.0  memory length: 100000  epsilon:
0.32820184000976754  steps: 408  lr: 6.4000000000000001e-06  evaluation
reward: 8.29
episode: 1737  score: 11.0  memory length: 100000  epsilon:
0.32715640000976093  steps: 528  lr: 6.4000000000000001e-06  evaluation
reward: 8.37
episode: 1738  score: 7.0  memory length: 100000  epsilon: 0.3263941000097561
steps: 385  lr: 6.4000000000000001e-06  evaluation reward: 8.35
episode: 1739  score: 6.0  memory length: 100000  epsilon:
0.32567338000975155  steps: 364  lr: 6.4000000000000001e-06  evaluation
reward: 8.35
episode: 1740  score: 11.0  memory length: 100000  epsilon: 0.324635860009745
steps: 524  lr: 6.4000000000000001e-06  evaluation reward: 8.39
episode: 1741  score: 5.0  memory length: 100000  epsilon:
0.32406166000974135  steps: 290  lr: 6.4000000000000001e-06  evaluation
reward: 8.36
episode: 1742  score: 10.0  memory length: 100000  epsilon:
0.3230419600097349  steps: 515  lr: 6.4000000000000001e-06  evaluation
reward: 8.32
episode: 1743  score: 4.0  memory length: 100000  epsilon: 0.3225232000097316
steps: 262  lr: 6.4000000000000001e-06  evaluation reward: 8.3
episode: 1744  score: 7.0  memory length: 100000  epsilon: 0.321790600009727
steps: 370  lr: 6.4000000000000001e-06  evaluation reward: 8.34

```

episode: 1745 score: 7.0 memory length: 100000 epsilon:
 0.32106196000972237 steps: 368 lr: 6.4000000000000001e-06 evaluation
 reward: 8.31
 episode: 1746 score: 11.0 memory length: 100000 epsilon:
 0.3200680000097161 steps: 502 lr: 6.4000000000000001e-06 evaluation
 reward: 8.37
 episode: 1747 score: 18.0 memory length: 100000 epsilon: 0.31910374000971
 steps: 487 lr: 6.4000000000000001e-06 evaluation reward: 8.34
 episode: 1748 score: 12.0 memory length: 100000 epsilon:
 0.31794544000970265 steps: 585 lr: 6.4000000000000001e-06 evaluation
 reward: 8.37
 episode: 1749 score: 8.0 memory length: 100000 epsilon: 0.3170960200096973
 steps: 429 lr: 6.4000000000000001e-06 evaluation reward: 8.41
 episode: 1750 score: 8.0 memory length: 100000 epsilon:
 0.31622086000969174 steps: 442 lr: 6.4000000000000001e-06 evaluation
 reward: 8.41
 episode: 1751 score: 8.0 memory length: 100000 epsilon: 0.3153595600096863
 steps: 435 lr: 6.4000000000000001e-06 evaluation reward: 8.35
 episode: 1752 score: 10.0 memory length: 100000 epsilon:
 0.3144428200096805 steps: 463 lr: 6.4000000000000001e-06 evaluation
 reward: 8.31
 episode: 1753 score: 8.0 memory length: 100000 epsilon: 0.3136349800096754
 steps: 408 lr: 6.4000000000000001e-06 evaluation reward: 8.32
 episode: 1754 score: 4.0 memory length: 100000 epsilon: 0.3131300800096722
 steps: 255 lr: 6.4000000000000001e-06 evaluation reward: 8.24
 episode: 1755 score: 8.0 memory length: 100000 epsilon:
 0.31235194000966726 steps: 393 lr: 6.4000000000000001e-06 evaluation
 reward: 8.27
 episode: 1756 score: 3.0 memory length: 100000 epsilon: 0.3118965400096644
 steps: 230 lr: 6.4000000000000001e-06 evaluation reward: 8.22
 episode: 1757 score: 7.0 memory length: 100000 epsilon:
 0.31111444000965943 steps: 395 lr: 6.4000000000000001e-06 evaluation
 reward: 8.22
 episode: 1758 score: 10.0 memory length: 100000 epsilon:
 0.31009276000965297 steps: 516 lr: 6.4000000000000001e-06 evaluation
 reward: 8.3
 episode: 1759 score: 5.0 memory length: 100000 epsilon: 0.3094472800096489
 steps: 326 lr: 6.4000000000000001e-06 evaluation reward: 8.28
 episode: 1760 score: 4.0 memory length: 100000 epsilon: 0.3088631800096452
 steps: 295 lr: 6.4000000000000001e-06 evaluation reward: 8.25
 episode: 1761 score: 14.0 memory length: 100000 epsilon:
 0.30760588000963723 steps: 635 lr: 6.4000000000000001e-06 evaluation
 reward: 8.3
 episode: 1762 score: 6.0 memory length: 100000 epsilon: 0.3069524800096331
 steps: 330 lr: 6.4000000000000001e-06 evaluation reward: 8.32
 episode: 1763 score: 14.0 memory length: 100000 epsilon:
 0.3059090200096265 steps: 527 lr: 6.4000000000000001e-06 evaluation
 reward: 8.42

episode: 1764 score: 8.0 memory length: 100000 epsilon: 0.3050200000096209
 steps: 449 lr: 6.4000000000000001e-06 evaluation reward: 8.47
 episode: 1765 score: 6.0 memory length: 100000 epsilon:
 0.30432106000961645 steps: 353 lr: 6.4000000000000001e-06 evaluation
 reward: 8.46
 episode: 1766 score: 7.0 memory length: 100000 epsilon: 0.3035528200096116
 steps: 388 lr: 6.4000000000000001e-06 evaluation reward: 8.42
 episode: 1767 score: 5.0 memory length: 100000 epsilon: 0.3029033800096075
 steps: 328 lr: 6.4000000000000001e-06 evaluation reward: 8.43
 episode: 1768 score: 9.0 memory length: 100000 epsilon: 0.3019708000096016
 steps: 471 lr: 6.4000000000000001e-06 evaluation reward: 8.43
 episode: 1769 score: 14.0 memory length: 100000 epsilon:
 0.30110950000959613 steps: 435 lr: 6.4000000000000001e-06 evaluation
 reward: 8.41
 episode: 1770 score: 12.0 memory length: 100000 epsilon:
 0.30021058000959044 steps: 454 lr: 6.4000000000000001e-06 evaluation
 reward: 8.47
 episode: 1771 score: 3.0 memory length: 100000 epsilon: 0.2997967600095878
 steps: 209 lr: 6.4000000000000001e-06 evaluation reward: 8.4
 episode: 1772 score: 9.0 memory length: 100000 epsilon:
 0.29879092000958146 steps: 508 lr: 6.4000000000000001e-06 evaluation
 reward: 8.28
 episode: 1773 score: 7.0 memory length: 100000 epsilon: 0.2980246600095766
 steps: 387 lr: 6.4000000000000001e-06 evaluation reward: 8.26
 episode: 1774 score: 9.0 memory length: 100000 epsilon: 0.297298000009572
 steps: 367 lr: 6.4000000000000001e-06 evaluation reward: 8.26
 episode: 1775 score: 6.0 memory length: 100000 epsilon:
 0.29665846000956797 steps: 323 lr: 6.4000000000000001e-06 evaluation
 reward: 8.18
 episode: 1776 score: 6.0 memory length: 100000 epsilon: 0.2959654600095636
 steps: 350 lr: 6.4000000000000001e-06 evaluation reward: 8.16
 episode: 1777 score: 11.0 memory length: 100000 epsilon:
 0.2948467600095565 steps: 565 lr: 6.4000000000000001e-06 evaluation
 reward: 8.2
 episode: 1778 score: 10.0 memory length: 100000 epsilon:
 0.2937518200095496 steps: 553 lr: 6.4000000000000001e-06 evaluation
 reward: 8.2
 episode: 1779 score: 8.0 memory length: 100000 epsilon:
 0.29294398000954447 steps: 408 lr: 6.4000000000000001e-06 evaluation
 reward: 8.16
 episode: 1780 score: 6.0 memory length: 100000 epsilon:
 0.29224702000954006 steps: 352 lr: 6.4000000000000001e-06 evaluation
 reward: 8.14
 episode: 1781 score: 7.0 memory length: 100000 epsilon: 0.2915263000095355
 steps: 364 lr: 6.4000000000000001e-06 evaluation reward: 8.19
 episode: 1782 score: 0.0 memory length: 100000 epsilon:
 0.29128276000953396 steps: 123 lr: 6.4000000000000001e-06 evaluation
 reward: 8.1

episode: 1783 score: 8.0 memory length: 100000 epsilon:
 0.29040958000952843 steps: 441 lr: 6.4000000000000001e-06 evaluation
 reward: 8.13
 episode: 1784 score: 7.0 memory length: 100000 epsilon:
 0.28962154000952345 steps: 398 lr: 6.4000000000000001e-06 evaluation
 reward: 8.12
 episode: 1785 score: 12.0 memory length: 100000 epsilon:
 0.28845334000951606 steps: 590 lr: 6.4000000000000001e-06 evaluation
 reward: 8.23
 episode: 1786 score: 4.0 memory length: 100000 epsilon:
 0.28793062000951275 steps: 264 lr: 6.4000000000000001e-06 evaluation
 reward: 8.2
 episode: 1787 score: 11.0 memory length: 100000 epsilon:
 0.2867347000095052 steps: 604 lr: 6.4000000000000001e-06 evaluation
 reward: 8.28
 episode: 1788 score: 9.0 memory length: 100000 epsilon: 0.2858041000094993
 steps: 470 lr: 6.4000000000000001e-06 evaluation reward: 8.24
 episode: 1789 score: 12.0 memory length: 100000 epsilon: 0.284649760009492
 steps: 583 lr: 6.4000000000000001e-06 evaluation reward: 8.29
 episode: 1790 score: 6.0 memory length: 100000 epsilon: 0.2839389400094875
 steps: 359 lr: 6.4000000000000001e-06 evaluation reward: 8.23
 episode: 1791 score: 7.0 memory length: 100000 epsilon:
 0.28313902000948243 steps: 404 lr: 6.4000000000000001e-06 evaluation
 reward: 8.21
 episode: 1792 score: 8.0 memory length: 100000 epsilon: 0.2823787000094776
 steps: 384 lr: 6.4000000000000001e-06 evaluation reward: 8.27
 episode: 1793 score: 9.0 memory length: 100000 epsilon: 0.2815411600094723
 steps: 423 lr: 6.4000000000000001e-06 evaluation reward: 8.28
 episode: 1794 score: 10.0 memory length: 100000 epsilon:
 0.28064620000946666 steps: 452 lr: 6.4000000000000001e-06 evaluation
 reward: 8.21
 episode: 1795 score: 5.0 memory length: 100000 epsilon:
 0.27999874000946257 steps: 327 lr: 6.4000000000000001e-06 evaluation
 reward: 8.16
 episode: 1796 score: 10.0 memory length: 100000 epsilon:
 0.27890776000945566 steps: 551 lr: 6.4000000000000001e-06 evaluation
 reward: 8.19
 episode: 1797 score: 11.0 memory length: 100000 epsilon:
 0.27784846000944896 steps: 535 lr: 6.4000000000000001e-06 evaluation
 reward: 8.22
 episode: 1798 score: 7.0 memory length: 100000 epsilon:
 0.27707428000944406 steps: 391 lr: 6.4000000000000001e-06 evaluation
 reward: 8.16
 episode: 1799 score: 13.0 memory length: 100000 epsilon:
 0.27605854000943764 steps: 513 lr: 6.4000000000000001e-06 evaluation
 reward: 8.2
 episode: 1800 score: 11.0 memory length: 100000 epsilon:
 0.2750230000094311 steps: 523 lr: 6.4000000000000001e-06 evaluation


```

reward: 8.19
episode: 1801  score: 13.0  memory length: 100000  epsilon:
0.2738686600094238  steps: 583  lr: 6.4000000000000001e-06  evaluation
reward: 8.24
episode: 1802  score: 8.0  memory length: 100000  epsilon:
0.27300934000941834  steps: 434  lr: 6.4000000000000001e-06  evaluation
reward: 8.21
episode: 1803  score: 10.0  memory length: 100000  epsilon:
0.27215992000941297  steps: 429  lr: 6.4000000000000001e-06  evaluation
reward: 8.27
episode: 1804  score: 12.0  memory length: 100000  epsilon:
0.27114814000940657  steps: 511  lr: 6.4000000000000001e-06  evaluation
reward: 8.35
episode: 1805  score: 7.0  memory length: 100000  epsilon: 0.2703799000094017
steps: 388  lr: 6.4000000000000001e-06  evaluation reward: 8.33
episode: 1806  score: 10.0  memory length: 100000  epsilon:
0.26931070000939494  steps: 540  lr: 6.4000000000000001e-06  evaluation
reward: 8.36
episode: 1807  score: 8.0  memory length: 100000  epsilon:
0.26840782000938923  steps: 456  lr: 6.4000000000000001e-06  evaluation
reward: 8.36
episode: 1808  score: 12.0  memory length: 100000  epsilon:
0.2672990200093822  steps: 560  lr: 6.4000000000000001e-06  evaluation
reward: 8.35
episode: 1809  score: 6.0  memory length: 100000  epsilon: 0.266631760009378
steps: 337  lr: 6.4000000000000001e-06  evaluation reward: 8.36
episode: 1810  score: 9.0  memory length: 100000  epsilon: 0.2657288800093723
steps: 456  lr: 6.4000000000000001e-06  evaluation reward: 8.38
episode: 1811  score: 6.0  memory length: 100000  epsilon: 0.2650675600093681
steps: 334  lr: 6.4000000000000001e-06  evaluation reward: 8.39
episode: 1812  score: 10.0  memory length: 100000  epsilon:
0.2640439000093616  steps: 517  lr: 6.4000000000000001e-06  evaluation
reward: 8.37
episode: 1813  score: 6.0  memory length: 100000  epsilon: 0.2633905000093575
steps: 330  lr: 6.4000000000000001e-06  evaluation reward: 8.35
episode: 1814  score: 3.0  memory length: 100000  epsilon: 0.2629014400093544
steps: 247  lr: 6.4000000000000001e-06  evaluation reward: 8.34
episode: 1815  score: 3.0  memory length: 100000  epsilon: 0.2624143600093513
steps: 246  lr: 6.4000000000000001e-06  evaluation reward: 8.26
episode: 1816  score: 3.0  memory length: 100000  epsilon: 0.2619253000093482
steps: 247  lr: 6.4000000000000001e-06  evaluation reward: 8.23
episode: 1817  score: 10.0  memory length: 100000  epsilon:
0.2609927200093423  steps: 471  lr: 6.4000000000000001e-06  evaluation
reward: 8.26
episode: 1818  score: 13.0  memory length: 100000  epsilon:
0.2598185800093349  steps: 593  lr: 6.4000000000000001e-06  evaluation
reward: 8.31
episode: 1819  score: 6.0  memory length: 100000  epsilon: 0.2591137000093304

```

```

steps: 356    lr: 6.4000000000000001e-06    evaluation reward: 8.17
episode: 1820    score: 9.0    memory length: 100000    epsilon:
0.25821874000932477    steps: 452    lr: 6.4000000000000001e-06    evaluation
reward: 8.18
episode: 1821    score: 7.0    memory length: 100000    epsilon: 0.2574465400093199
steps: 390    lr: 6.4000000000000001e-06    evaluation reward: 8.2
episode: 1822    score: 4.0    memory length: 100000    epsilon: 0.2569277800093166
steps: 262    lr: 6.4000000000000001e-06    evaluation reward: 8.2
episode: 1823    score: 7.0    memory length: 100000    epsilon: 0.2561852800093119
steps: 375    lr: 6.4000000000000001e-06    evaluation reward: 8.17
episode: 1824    score: 4.0    memory length: 100000    epsilon:
0.25559128000930814    steps: 300    lr: 6.4000000000000001e-06    evaluation
reward: 8.16
episode: 1825    score: 4.0    memory length: 100000    epsilon: 0.2550824200093049
steps: 257    lr: 6.4000000000000001e-06    evaluation reward: 8.12
episode: 1826    score: 11.0    memory length: 100000    epsilon:
0.25406074000929846    steps: 516    lr: 6.4000000000000001e-06    evaluation
reward: 8.18
episode: 1827    score: 12.0    memory length: 100000    epsilon:
0.2529321400092913    steps: 570    lr: 6.4000000000000001e-06    evaluation
reward: 8.19
episode: 1828    score: 6.0    memory length: 100000    epsilon: 0.2522351800092869
steps: 352    lr: 6.4000000000000001e-06    evaluation reward: 8.17
episode: 1829    score: 8.0    memory length: 100000    epsilon: 0.2514471400092819
steps: 398    lr: 6.4000000000000001e-06    evaluation reward: 8.21
episode: 1830    score: 3.0    memory length: 100000    epsilon: 0.250987780009279
steps: 232    lr: 6.4000000000000001e-06    evaluation reward: 8.17
episode: 1831    score: 16.0    memory length: 100000    epsilon:
0.24958594000927015    steps: 708    lr: 6.4000000000000001e-06    evaluation
reward: 8.25
episode: 1832    score: 13.0    memory length: 100000    epsilon:
0.24860188000926392    steps: 497    lr: 6.4000000000000001e-06    evaluation
reward: 8.33
episode: 1833    score: 8.0    memory length: 100000    epsilon:
0.24774256000925848    steps: 434    lr: 6.4000000000000001e-06    evaluation
reward: 8.32
episode: 1834    score: 19.0    memory length: 100000    epsilon:
0.24633280000924956    steps: 712    lr: 6.4000000000000001e-06    evaluation
reward: 8.36
episode: 1835    score: 12.0    memory length: 100000    epsilon:
0.24520420000924242    steps: 570    lr: 6.4000000000000001e-06    evaluation
reward: 8.35
episode: 1836    score: 8.0    memory length: 100000    epsilon: 0.2443943800092373
steps: 409    lr: 6.4000000000000001e-06    evaluation reward: 8.35
episode: 1837    score: 6.0    memory length: 100000    epsilon: 0.2437291000092331
steps: 336    lr: 6.4000000000000001e-06    evaluation reward: 8.3
episode: 1838    score: 17.0    memory length: 100000    epsilon:
0.24236686000922447    steps: 688    lr: 6.4000000000000001e-06    evaluation

```

```

reward: 8.4
episode: 1839  score: 15.0  memory length: 100000  epsilon:
0.2413907200092183  steps: 493  lr: 6.4000000000000001e-06  evaluation
reward: 8.49
episode: 1840  score: 10.0  memory length: 100000  epsilon:
0.24035122000921172  steps: 525  lr: 6.4000000000000001e-06  evaluation
reward: 8.48
episode: 1841  score: 7.0  memory length: 100000  epsilon:
0.23951566000920643  steps: 422  lr: 6.4000000000000001e-06  evaluation
reward: 8.5
episode: 1842  score: 6.0  memory length: 100000  epsilon:
0.23877712000920176  steps: 373  lr: 6.4000000000000001e-06  evaluation
reward: 8.46
episode: 1843  score: 16.0  memory length: 100000  epsilon:
0.2376465400091946  steps: 571  lr: 6.4000000000000001e-06  evaluation
reward: 8.58
episode: 1844  score: 7.0  memory length: 100000  epsilon: 0.2368882000091898
steps: 383  lr: 6.4000000000000001e-06  evaluation reward: 8.58
episode: 1845  score: 6.0  memory length: 100000  epsilon:
0.23618332000918535  steps: 356  lr: 6.4000000000000001e-06  evaluation
reward: 8.57
episode: 1846  score: 13.0  memory length: 100000  epsilon:
0.23493394000917744  steps: 631  lr: 6.4000000000000001e-06  evaluation
reward: 8.59
episode: 1847  score: 9.0  memory length: 100000  epsilon:
0.23402908000917172  steps: 457  lr: 6.4000000000000001e-06  evaluation
reward: 8.5
episode: 1848  score: 4.0  memory length: 100000  epsilon:
0.23344498000916802  steps: 295  lr: 6.4000000000000001e-06  evaluation
reward: 8.42
episode: 1849  score: 6.0  memory length: 100000  epsilon:
0.23274208000916358  steps: 355  lr: 6.4000000000000001e-06  evaluation
reward: 8.4
episode: 1850  score: 10.0  memory length: 100000  epsilon:
0.23172436000915714  steps: 514  lr: 6.4000000000000001e-06  evaluation
reward: 8.42
episode: 1851  score: 9.0  memory length: 100000  epsilon:
0.23077594000915114  steps: 479  lr: 6.4000000000000001e-06  evaluation
reward: 8.43
episode: 1852  score: 8.0  memory length: 100000  epsilon:
0.22994038000914585  steps: 422  lr: 6.4000000000000001e-06  evaluation
reward: 8.41
episode: 1853  score: 9.0  memory length: 100000  epsilon: 0.22901770000914
steps: 466  lr: 6.4000000000000001e-06  evaluation reward: 8.42
episode: 1854  score: 7.0  memory length: 100000  epsilon:
0.22826134000913523  steps: 382  lr: 6.4000000000000001e-06  evaluation
reward: 8.45
episode: 1855  score: 12.0  memory length: 100000  epsilon:

```

0.22709908000912787	steps: 587	lr: 6.4000000000000001e-06	evaluation
reward: 8.49			
episode: 1856	score: 11.0	memory length: 100000	epsilon:
0.22603384000912113	steps: 538	lr: 6.4000000000000001e-06	evaluation
reward: 8.57			
episode: 1857	score: 6.0	memory length: 100000	epsilon:
0.22537450000911696	steps: 333	lr: 6.4000000000000001e-06	evaluation
reward: 8.56			
episode: 1858	score: 5.0	memory length: 100000	epsilon:
0.22472506000911285	steps: 328	lr: 6.4000000000000001e-06	evaluation
reward: 8.51			
episode: 1859	score: 10.0	memory length: 100000	epsilon:
0.22388158000910752	steps: 426	lr: 6.4000000000000001e-06	evaluation
reward: 8.56			
episode: 1860	score: 7.0	memory length: 100000	epsilon:
0.22317472000910304	steps: 357	lr: 6.4000000000000001e-06	evaluation
reward: 8.59			
episode: 1861	score: 6.0	memory length: 100000	epsilon: 0.2224262800090983
steps: 378	lr: 6.4000000000000001e-06	evaluation	reward: 8.51
episode: 1862	score: 12.0	memory length: 100000	epsilon:
0.22150558000909248	steps: 465	lr: 6.4000000000000001e-06	evaluation
reward: 8.57			
episode: 1863	score: 7.0	memory length: 100000	epsilon:
0.22076110000908777	steps: 376	lr: 6.4000000000000001e-06	evaluation
reward: 8.5			
episode: 1864	score: 7.0	memory length: 100000	epsilon:
0.21996712000908275	steps: 401	lr: 6.4000000000000001e-06	evaluation
reward: 8.49			
episode: 1865	score: 6.0	memory length: 100000	epsilon:
0.21919096000907784	steps: 392	lr: 6.4000000000000001e-06	evaluation
reward: 8.49			
episode: 1866	score: 7.0	memory length: 100000	epsilon:
0.21845044000907315	steps: 374	lr: 6.4000000000000001e-06	evaluation
reward: 8.49			
episode: 1867	score: 14.0	memory length: 100000	epsilon:
0.21739312000906647	steps: 534	lr: 6.4000000000000001e-06	evaluation
reward: 8.58			
episode: 1868	score: 14.0	memory length: 100000	epsilon:
0.21609226000905823	steps: 657	lr: 6.4000000000000001e-06	evaluation
reward: 8.63			
episode: 1869	score: 9.0	memory length: 100000	epsilon:
0.21518938000905252	steps: 456	lr: 6.4000000000000001e-06	evaluation
reward: 8.58			
episode: 1870	score: 15.0	memory length: 100000	epsilon:
0.21411622000904573	steps: 542	lr: 6.4000000000000001e-06	evaluation
reward: 8.61			
episode: 1871	score: 11.0	memory length: 100000	epsilon:
0.21309454000903927	steps: 516	lr: 6.4000000000000001e-06	evaluation

```

reward: 8.69
episode: 1872  score: 9.0  memory length: 100000  epsilon:
0.21209464000903294  steps: 505  lr: 6.4000000000000001e-06  evaluation
reward: 8.69
episode: 1873  score: 7.0  memory length: 100000  epsilon:
0.21133630000902814  steps: 383  lr: 6.4000000000000001e-06  evaluation
reward: 8.69
episode: 1874  score: 10.0  memory length: 100000  epsilon:
0.21037798000902208  steps: 484  lr: 6.4000000000000001e-06  evaluation
reward: 8.7
episode: 1875  score: 6.0  memory length: 100000  epsilon:
0.20974240000901806  steps: 321  lr: 6.4000000000000001e-06  evaluation
reward: 8.7
episode: 1876  score: 7.0  memory length: 100000  epsilon: 0.2090058400090134
steps: 372  lr: 6.4000000000000001e-06  evaluation reward: 8.71
episode: 1877  score: 7.0  memory length: 100000  epsilon:
0.20823562000900853  steps: 389  lr: 6.4000000000000001e-06  evaluation
reward: 8.67
episode: 1878  score: 13.0  memory length: 100000  epsilon:
0.20709118000900129  steps: 578  lr: 2.5600000000000001e-06  evaluation
reward: 8.7
episode: 1879  score: 11.0  memory length: 100000  epsilon:
0.20630908000899634  steps: 395  lr: 2.5600000000000001e-06  evaluation
reward: 8.73
episode: 1880  score: 5.0  memory length: 100000  epsilon:
0.20569924000899248  steps: 308  lr: 2.5600000000000001e-06  evaluation
reward: 8.72
episode: 1881  score: 7.0  memory length: 100000  epsilon:
0.20499436000898802  steps: 356  lr: 2.5600000000000001e-06  evaluation
reward: 8.72
episode: 1882  score: 16.0  memory length: 100000  epsilon:
0.20375290000898016  steps: 627  lr: 2.5600000000000001e-06  evaluation
reward: 8.88
episode: 1883  score: 9.0  memory length: 100000  epsilon:
0.20290150000897478  steps: 430  lr: 2.5600000000000001e-06  evaluation
reward: 8.89
episode: 1884  score: 6.0  memory length: 100000  epsilon:
0.20218474000897024  steps: 362  lr: 2.5600000000000001e-06  evaluation
reward: 8.88
episode: 1885  score: 18.0  memory length: 100000  epsilon:
0.20077498000896132  steps: 712  lr: 2.5600000000000001e-06  evaluation
reward: 8.94
episode: 1886  score: 18.0  memory length: 100000  epsilon:
0.19938304000895252  steps: 703  lr: 2.5600000000000001e-06  evaluation
reward: 9.08
episode: 1887  score: 10.0  memory length: 100000  epsilon:
0.19830988000894573  steps: 542  lr: 2.5600000000000001e-06  evaluation
reward: 9.07

```

episode: 1888 score: 13.0 memory length: 100000 epsilon: 0.1971357400089383 steps: 593 lr: 2.5600000000000001e-06 evaluation
 reward: 9.11
 episode: 1889 score: 10.0 memory length: 100000 epsilon: 0.1961259400089319 steps: 510 lr: 2.5600000000000001e-06 evaluation
 reward: 9.09
 episode: 1890 score: 9.0 memory length: 100000 epsilon: 0.19520128000892606 steps: 467 lr: 2.5600000000000001e-06 evaluation
 reward: 9.12
 episode: 1891 score: 8.0 memory length: 100000 epsilon: 0.19439542000892096 steps: 407 lr: 2.5600000000000001e-06 evaluation
 reward: 9.13
 episode: 1892 score: 18.0 memory length: 100000 epsilon: 0.19287874000891136 steps: 766 lr: 2.5600000000000001e-06 evaluation
 reward: 9.23
 episode: 1893 score: 13.0 memory length: 100000 epsilon: 0.19193032000890536 steps: 479 lr: 2.5600000000000001e-06 evaluation
 reward: 9.27
 episode: 1894 score: 9.0 memory length: 100000 epsilon: 0.1910036800088995 steps: 468 lr: 2.5600000000000001e-06 evaluation reward: 9.26
 episode: 1895 score: 10.0 memory length: 100000 epsilon: 0.1899582400088929 steps: 528 lr: 2.5600000000000001e-06 evaluation
 reward: 9.31
 episode: 1896 score: 13.0 memory length: 100000 epsilon: 0.18869896000888492 steps: 636 lr: 2.5600000000000001e-06 evaluation
 reward: 9.34
 episode: 1897 score: 12.0 memory length: 100000 epsilon: 0.18749710000887732 steps: 607 lr: 2.5600000000000001e-06 evaluation
 reward: 9.35
 episode: 1898 score: 9.0 memory length: 100000 epsilon: 0.18661600000887174 steps: 445 lr: 2.5600000000000001e-06 evaluation
 reward: 9.37
 episode: 1899 score: 17.0 memory length: 100000 epsilon: 0.18524386000886306 steps: 693 lr: 2.5600000000000001e-06 evaluation
 reward: 9.41
 episode: 1900 score: 18.0 memory length: 100000 epsilon: 0.1839232000088547 steps: 667 lr: 2.5600000000000001e-06 evaluation
 reward: 9.48
 episode: 1901 score: 2.0 memory length: 100000 epsilon: 0.18349156000885197 steps: 218 lr: 2.5600000000000001e-06 evaluation
 reward: 9.37
 episode: 1902 score: 4.0 memory length: 100000 epsilon: 0.18295102000884855 steps: 273 lr: 2.5600000000000001e-06 evaluation
 reward: 9.33
 episode: 1903 score: 11.0 memory length: 100000 epsilon: 0.18184618000884156 steps: 558 lr: 2.5600000000000001e-06 evaluation
 reward: 9.34
 episode: 1904 score: 10.0 memory length: 100000 epsilon:

0.1808561800088353 steps: 500 lr: 2.5600000000000001e-06 evaluation
 reward: 9.32
 episode: 1905 score: 11.0 memory length: 100000 epsilon:
 0.1796701600088278 steps: 599 lr: 2.5600000000000001e-06 evaluation
 reward: 9.36
 episode: 1906 score: 17.0 memory length: 100000 epsilon:
 0.17808022000881774 steps: 803 lr: 2.5600000000000001e-06 evaluation
 reward: 9.43
 episode: 1907 score: 17.0 memory length: 100000 epsilon:
 0.17699122000881085 steps: 550 lr: 2.5600000000000001e-06 evaluation
 reward: 9.52
 episode: 1908 score: 6.0 memory length: 100000 epsilon: 0.1763695000088069
 steps: 314 lr: 2.5600000000000001e-06 evaluation reward: 9.46
 episode: 1909 score: 8.0 memory length: 100000 epsilon: 0.1755478000088017
 steps: 415 lr: 2.5600000000000001e-06 evaluation reward: 9.48
 episode: 1910 score: 8.0 memory length: 100000 epsilon: 0.1747379800087966
 steps: 409 lr: 2.5600000000000001e-06 evaluation reward: 9.47
 episode: 1911 score: 10.0 memory length: 100000 epsilon:
 0.17373412000879024 steps: 507 lr: 2.5600000000000001e-06 evaluation
 reward: 9.51
 episode: 1912 score: 11.0 memory length: 100000 epsilon:
 0.1726708600087835 steps: 537 lr: 2.5600000000000001e-06 evaluation
 reward: 9.52
 episode: 1913 score: 8.0 memory length: 100000 epsilon:
 0.17177986000877787 steps: 450 lr: 2.5600000000000001e-06 evaluation
 reward: 9.54
 episode: 1914 score: 8.0 memory length: 100000 epsilon: 0.171011620008773
 steps: 388 lr: 2.5600000000000001e-06 evaluation reward: 9.59
 episode: 1915 score: 5.0 memory length: 100000 epsilon:
 0.17039782000876913 steps: 310 lr: 2.5600000000000001e-06 evaluation
 reward: 9.61
 episode: 1916 score: 6.0 memory length: 100000 epsilon:
 0.16972462000876487 steps: 340 lr: 2.5600000000000001e-06 evaluation
 reward: 9.64
 episode: 1917 score: 16.0 memory length: 100000 epsilon:
 0.16846138000875688 steps: 638 lr: 2.5600000000000001e-06 evaluation
 reward: 9.7
 episode: 1918 score: 7.0 memory length: 100000 epsilon:
 0.16765354000875177 steps: 408 lr: 2.5600000000000001e-06 evaluation
 reward: 9.64
 episode: 1919 score: 12.0 memory length: 100000 epsilon:
 0.16644772000874414 steps: 609 lr: 2.5600000000000001e-06 evaluation
 reward: 9.7
 episode: 1920 score: 12.0 memory length: 100000 epsilon:
 0.16526170000873663 steps: 599 lr: 2.5600000000000001e-06 evaluation
 reward: 9.73
 episode: 1921 score: 13.0 memory length: 100000 epsilon:
 0.16429942000873055 steps: 486 lr: 2.5600000000000001e-06 evaluation

```

reward: 9.79
episode: 1922  score: 11.0  memory length: 100000  epsilon: 0.163265860008724
steps: 522  lr: 2.5600000000000001e-06  evaluation reward: 9.86
episode: 1923  score: 7.0  memory length: 100000  epsilon:
0.16251544000871926  steps: 379  lr: 2.5600000000000001e-06  evaluation
reward: 9.86
episode: 1924  score: 11.0  memory length: 100000  epsilon: 0.161527420008713
steps: 499  lr: 2.5600000000000001e-06  evaluation reward: 9.93
episode: 1925  score: 7.0  memory length: 100000  epsilon:
0.16082848000870859  steps: 353  lr: 2.5600000000000001e-06  evaluation
reward: 9.96
episode: 1926  score: 4.0  memory length: 100000  epsilon:
0.16031170000870532  steps: 261  lr: 2.5600000000000001e-06  evaluation
reward: 9.89
episode: 1927  score: 5.0  memory length: 100000  epsilon:
0.15973750000870168  steps: 290  lr: 2.5600000000000001e-06  evaluation
reward: 9.82
episode: 1928  score: 11.0  memory length: 100000  epsilon:
0.15889996000869638  steps: 423  lr: 2.5600000000000001e-06  evaluation
reward: 9.87
episode: 1929  score: 14.0  memory length: 100000  epsilon:
0.15756346000868793  steps: 675  lr: 2.5600000000000001e-06  evaluation
reward: 9.93
episode: 1930  score: 12.0  memory length: 100000  epsilon:
0.1563754600086804  steps: 600  lr: 2.5600000000000001e-06  evaluation
reward: 10.02
episode: 1931  score: 19.0  memory length: 100000  epsilon:
0.15519538000867295  steps: 596  lr: 2.5600000000000001e-06  evaluation
reward: 10.05
episode: 1932  score: 15.0  memory length: 100000  epsilon:
0.15407470000866585  steps: 566  lr: 2.5600000000000001e-06  evaluation
reward: 10.07
episode: 1933  score: 19.0  memory length: 100000  epsilon:
0.15301342000865914  steps: 536  lr: 2.5600000000000001e-06  evaluation
reward: 10.18
episode: 1934  score: 6.0  memory length: 100000  epsilon:
0.15239566000865523  steps: 312  lr: 2.5600000000000001e-06  evaluation
reward: 10.05
episode: 1935  score: 11.0  memory length: 100000  epsilon:
0.15125518000864802  steps: 576  lr: 2.5600000000000001e-06  evaluation
reward: 10.04
episode: 1936  score: 7.0  memory length: 100000  epsilon: 0.1504473400086429
steps: 408  lr: 2.5600000000000001e-06  evaluation reward: 10.03
episode: 1937  score: 13.0  memory length: 100000  epsilon:
0.1494989200086369  steps: 479  lr: 2.5600000000000001e-06  evaluation
reward: 10.1
episode: 1938  score: 10.0  memory length: 100000  epsilon:
0.14846734000863038  steps: 521  lr: 2.5600000000000001e-06  evaluation

```



```

reward: 10.03
episode: 1939  score: 19.0  memory length: 100000  epsilon:
0.14708530000862163  steps: 698  lr: 2.5600000000000001e-06  evaluation
reward: 10.07
episode: 1940  score: 6.0  memory length: 100000  epsilon:
0.14641210000861737  steps: 340  lr: 2.5600000000000001e-06  evaluation
reward: 10.03
episode: 1941  score: 6.0  memory length: 100000  epsilon:
0.14575474000861321  steps: 332  lr: 2.5600000000000001e-06  evaluation
reward: 10.02
episode: 1942  score: 16.0  memory length: 100000  epsilon:
0.14442220000860478  steps: 673  lr: 2.5600000000000001e-06  evaluation
reward: 10.12
episode: 1943  score: 7.0  memory length: 100000  epsilon:
0.14364208000859985  steps: 394  lr: 2.5600000000000001e-06  evaluation
reward: 10.03
episode: 1944  score: 10.0  memory length: 100000  epsilon:
0.1427332600085941  steps: 459  lr: 2.5600000000000001e-06  evaluation
reward: 10.06
episode: 1945  score: 10.0  memory length: 100000  epsilon:
0.14165614000858728  steps: 544  lr: 2.5600000000000001e-06  evaluation
reward: 10.1
episode: 1946  score: 14.0  memory length: 100000  epsilon:
0.14047012000857978  steps: 599  lr: 2.5600000000000001e-06  evaluation
reward: 10.11
episode: 1947  score: 8.0  memory length: 100000  epsilon:
0.13975930000857528  steps: 359  lr: 2.5600000000000001e-06  evaluation
reward: 10.1
episode: 1948  score: 17.0  memory length: 100000  epsilon:
0.1385772400085678  steps: 597  lr: 2.5600000000000001e-06  evaluation
reward: 10.23
episode: 1949  score: 8.0  memory length: 100000  epsilon: 0.1377872200085628
steps: 399  lr: 2.5600000000000001e-06  evaluation reward: 10.25
episode: 1950  score: 13.0  memory length: 100000  epsilon:
0.13683484000855678  steps: 481  lr: 2.5600000000000001e-06  evaluation
reward: 10.28
episode: 1951  score: 10.0  memory length: 100000  epsilon:
0.13586860000855067  steps: 488  lr: 2.5600000000000001e-06  evaluation
reward: 10.29
episode: 1952  score: 9.0  memory length: 100000  epsilon:
0.13497562000854502  steps: 451  lr: 2.5600000000000001e-06  evaluation
reward: 10.3
episode: 1953  score: 14.0  memory length: 100000  epsilon:
0.1339618600085386  steps: 512  lr: 2.5600000000000001e-06  evaluation
reward: 10.35
episode: 1954  score: 8.0  memory length: 100000  epsilon:
0.13311244000853323  steps: 429  lr: 2.5600000000000001e-06  evaluation
reward: 10.36

```

episode: 1955	score: 15.0	memory length: 100000	epsilon: 0.1318591000085253	steps: 633	lr: 2.5600000000000001e-06	evaluation
reward: 10.39						
episode: 1956	score: 14.0	memory length: 100000	epsilon: 0.13076218000851836	steps: 554	lr: 2.5600000000000001e-06	evaluation
reward: 10.42						
episode: 1957	score: 15.0	memory length: 100000	epsilon: 0.1294732000085102	steps: 651	lr: 2.5600000000000001e-06	evaluation
reward: 10.51						
episode: 1958	score: 5.0	memory length: 100000	epsilon: 0.12882772000850612	steps: 326	lr: 2.5600000000000001e-06	evaluation
reward: 10.51						
episode: 1959	score: 10.0	memory length: 100000	epsilon: 0.1278753400085001	steps: 481	lr: 2.5600000000000001e-06	evaluation
reward: 10.51						
episode: 1960	score: 10.0	memory length: 100000	epsilon: 0.12686950000849373	steps: 508	lr: 2.5600000000000001e-06	evaluation
reward: 10.54						
episode: 1961	score: 6.0	memory length: 100000	epsilon: 0.12616462000848927	steps: 356	lr: 2.5600000000000001e-06	evaluation
reward: 10.54						
episode: 1962	score: 10.0	memory length: 100000	epsilon: 0.12528154000848368	steps: 446	lr: 2.5600000000000001e-06	evaluation
reward: 10.52						
episode: 1963	score: 22.0	memory length: 100000	epsilon: 0.12402028000848257	steps: 637	lr: 2.5600000000000001e-06	evaluation
reward: 10.67						
episode: 1964	score: 10.0	memory length: 100000	epsilon: 0.12307384000848322	steps: 478	lr: 2.5600000000000001e-06	evaluation
reward: 10.7						
episode: 1965	score: 6.0	memory length: 100000	epsilon: 0.1223669800084837	steps: 357	lr: 2.5600000000000001e-06	evaluation
reward: 10.7						
episode: 1966	score: 12.0	memory length: 100000	epsilon: 0.12127798000848444	steps: 550	lr: 2.5600000000000001e-06	evaluation
reward: 10.75						
episode: 1967	score: 6.0	memory length: 100000	epsilon: 0.12056716000848493	steps: 359	lr: 2.5600000000000001e-06	evaluation
reward: 10.67						
episode: 1968	score: 8.0	memory length: 100000	epsilon: 0.11970190000848552	steps: 437	lr: 2.5600000000000001e-06	evaluation
reward: 10.61						
episode: 1969	score: 18.0	memory length: 100000	epsilon: 0.11859508000848627	steps: 559	lr: 2.5600000000000001e-06	evaluation
reward: 10.7						
episode: 1970	score: 13.0	memory length: 100000	epsilon: 0.11744272000848706	steps: 582	lr: 2.5600000000000001e-06	evaluation
reward: 10.68						
episode: 1971	score: 15.0	memory length: 100000	epsilon:			

0.11613394000848795	steps: 661	lr: 2.5600000000000001e-06	evaluation
reward: 10.72			
episode: 1972	score: 12.0	memory length: 100000	epsilon:
0.11521126000848858	steps: 466	lr: 2.5600000000000001e-06	evaluation
reward: 10.75			
episode: 1973	score: 11.0	memory length: 100000	epsilon:
0.1141480000084893	steps: 537	lr: 2.5600000000000001e-06	evaluation
reward: 10.79			
episode: 1974	score: 12.0	memory length: 100000	epsilon:
0.11323522000848993	steps: 461	lr: 2.5600000000000001e-06	evaluation
reward: 10.81			
episode: 1975	score: 7.0	memory length: 100000	epsilon: 0.1123937200084905
steps: 425	lr: 2.5600000000000001e-06	evaluation	reward: 10.82
episode: 1976	score: 13.0	memory length: 100000	epsilon:
0.11116810000849134	steps: 619	lr: 2.5600000000000001e-06	evaluation
reward: 10.88			
episode: 1977	score: 10.0	memory length: 100000	epsilon:
0.11009296000849207	steps: 543	lr: 2.5600000000000001e-06	evaluation
reward: 10.91			
episode: 1978	score: 12.0	memory length: 100000	epsilon:
0.10900792000849281	steps: 548	lr: 2.5600000000000001e-06	evaluation
reward: 10.9			
episode: 1979	score: 9.0	memory length: 100000	epsilon:
0.10811494000849342	steps: 451	lr: 2.5600000000000001e-06	evaluation
reward: 10.88			
episode: 1980	score: 9.0	memory length: 100000	epsilon:
0.10717840000849406	steps: 473	lr: 2.5600000000000001e-06	evaluation
reward: 10.92			
episode: 1981	score: 10.0	memory length: 100000	epsilon:
0.10629928000849466	steps: 444	lr: 2.5600000000000001e-06	evaluation
reward: 10.95			
episode: 1982	score: 19.0	memory length: 100000	epsilon:
0.10499842000849555	steps: 657	lr: 2.5600000000000001e-06	evaluation
reward: 10.98			
episode: 1983	score: 7.0	memory length: 100000	epsilon:
0.10422820000849607	steps: 389	lr: 2.5600000000000001e-06	evaluation
reward: 10.96			
episode: 1984	score: 12.0	memory length: 100000	epsilon:
0.10321642000849676	steps: 511	lr: 2.5600000000000001e-06	evaluation
reward: 11.02			
episode: 1985	score: 7.0	memory length: 100000	epsilon:
0.10238284000849733	steps: 421	lr: 2.5600000000000001e-06	evaluation
reward: 10.91			
episode: 1986	score: 18.0	memory length: 100000	epsilon:
0.10091962000849833	steps: 739	lr: 2.5600000000000001e-06	evaluation
reward: 10.91			
episode: 1987	score: 18.0	memory length: 100000	epsilon:
0.0994999600084993	steps: 717	lr: 2.5600000000000001e-06	evaluation

```

reward: 10.99
episode: 1988   score: 14.0   memory length: 100000   epsilon:
0.09843670000850002   steps: 537   lr: 2.5600000000000001e-06   evaluation
reward: 11.0
episode: 1989   score: 11.0   memory length: 100000   epsilon:
0.09737938000850074   steps: 534   lr: 2.5600000000000001e-06   evaluation
reward: 11.01
episode: 1990   score: 8.0    memory length: 100000   epsilon:
0.09654382000850131   steps: 422   lr: 2.5600000000000001e-06   evaluation
reward: 11.0
episode: 1991   score: 11.0   memory length: 100000   epsilon:
0.09546076000850205   steps: 547   lr: 2.5600000000000001e-06   evaluation
reward: 11.03
episode: 1992   score: 19.0   memory length: 100000   epsilon:
0.09410446000850298   steps: 685   lr: 2.5600000000000001e-06   evaluation
reward: 11.04
episode: 1993   score: 6.0    memory length: 100000   epsilon:
0.09343126000850344   steps: 340   lr: 2.5600000000000001e-06   evaluation
reward: 10.97
episode: 1994   score: 16.0   memory length: 100000   epsilon:
0.09207298000850436   steps: 686   lr: 2.5600000000000001e-06   evaluation
reward: 11.04
episode: 1995   score: 8.0    memory length: 100000   epsilon:
0.09123544000850493   steps: 423   lr: 2.5600000000000001e-06   evaluation
reward: 11.02
episode: 1996   score: 17.0   memory length: 100000   epsilon:
0.09002368000850576   steps: 612   lr: 2.5600000000000001e-06   evaluation
reward: 11.06
episode: 1997   score: 13.0   memory length: 100000   epsilon:
0.08882974000850657   steps: 603   lr: 2.5600000000000001e-06   evaluation
reward: 11.07
episode: 1998   score: 19.0   memory length: 100000   epsilon:
0.08723782000850766   steps: 804   lr: 2.5600000000000001e-06   evaluation
reward: 11.17
episode: 1999   score: 6.0    memory length: 100000   epsilon: 0.0864557200085082
steps: 395   lr: 2.5600000000000001e-06   evaluation reward: 11.06
episode: 2000   score: 7.0    memory length: 100000   epsilon:
0.08575678000850867   steps: 353   lr: 2.5600000000000001e-06   evaluation
reward: 10.95
episode: 2001   score: 13.0   memory length: 100000   epsilon:
0.08480440000850932   steps: 481   lr: 2.5600000000000001e-06   evaluation
reward: 11.06
episode: 2002   score: 17.0   memory length: 100000   epsilon:
0.0833669200085103   steps: 726   lr: 2.5600000000000001e-06   evaluation
reward: 11.19
episode: 2003   score: 9.0    memory length: 100000   epsilon:
0.08250760000851089   steps: 434   lr: 2.5600000000000001e-06   evaluation
reward: 11.17

```

episode: 2004	score: 18.0	memory length: 100000	epsilon: 0.08124238000851175	steps: 639	lr: 2.5600000000000001e-06	evaluation
reward: 11.25						
episode: 2005	score: 10.0	memory length: 100000	epsilon: 0.08024644000851243	steps: 503	lr: 2.5600000000000001e-06	evaluation
reward: 11.24						
episode: 2006	score: 9.0	memory length: 100000	epsilon: 0.07929010000851308	steps: 483	lr: 2.5600000000000001e-06	evaluation
reward: 11.16						
episode: 2007	score: 12.0	memory length: 100000	epsilon: 0.07817338000851384	steps: 564	lr: 2.5600000000000001e-06	evaluation
reward: 11.11						
episode: 2008	score: 12.0	memory length: 100000	epsilon: 0.07728436000851445	steps: 449	lr: 2.5600000000000001e-06	evaluation
reward: 11.17						
episode: 2009	score: 12.0	memory length: 100000	epsilon: 0.07621318000851518	steps: 541	lr: 2.5600000000000001e-06	evaluation
reward: 11.21						
episode: 2010	score: 14.0	memory length: 100000	epsilon: 0.07491034000851607	steps: 658	lr: 2.5600000000000001e-06	evaluation
reward: 11.27						
episode: 2011	score: 11.0	memory length: 100000	epsilon: 0.07389262000851676	steps: 514	lr: 2.5600000000000001e-06	evaluation
reward: 11.28						
episode: 2012	score: 19.0	memory length: 100000	epsilon: 0.0725244400085177	steps: 691	lr: 2.5600000000000001e-06	evaluation
reward: 11.36						
episode: 2013	score: 6.0	memory length: 100000	epsilon: 0.07185124000851816	steps: 340	lr: 2.5600000000000001e-06	evaluation
reward: 11.34						
episode: 2014	score: 27.0	memory length: 100000	epsilon: 0.07005340000851938	steps: 908	lr: 2.5600000000000001e-06	evaluation
reward: 11.53						
episode: 2015	score: 10.0	memory length: 100000	epsilon: 0.06905548000852006	steps: 504	lr: 2.5600000000000001e-06	evaluation
reward: 11.58						
episode: 2016	score: 16.0	memory length: 100000	epsilon: 0.06785560000852088	steps: 606	lr: 2.5600000000000001e-06	evaluation
reward: 11.68						
episode: 2017	score: 19.0	memory length: 100000	epsilon: 0.06649336000852181	steps: 688	lr: 2.5600000000000001e-06	evaluation
reward: 11.71						
episode: 2018	score: 11.0	memory length: 100000	epsilon: 0.06550930000852248	steps: 497	lr: 2.5600000000000001e-06	evaluation
reward: 11.75						
episode: 2019	score: 9.0	memory length: 100000	epsilon: 0.06462622000852308	steps: 446	lr: 2.5600000000000001e-06	evaluation
reward: 11.72						

episode: 2020 score: 4.0 memory length: 100000 epsilon:
 0.06407380000852346 steps: 279 lr: 2.5600000000000001e-06 evaluation
 reward: 11.64
 episode: 2021 score: 6.0 memory length: 100000 epsilon: 0.0634342600085239
 steps: 323 lr: 2.5600000000000001e-06 evaluation reward: 11.57
 episode: 2022 score: 9.0 memory length: 100000 epsilon:
 0.06252148000852452 steps: 461 lr: 2.5600000000000001e-06 evaluation
 reward: 11.55
 episode: 2023 score: 15.0 memory length: 100000 epsilon:
 0.06136516000852531 steps: 584 lr: 2.5600000000000001e-06 evaluation
 reward: 11.63
 episode: 2024 score: 11.0 memory length: 100000 epsilon:
 0.06030586000852603 steps: 535 lr: 2.5600000000000001e-06 evaluation
 reward: 11.63
 episode: 2025 score: 10.0 memory length: 100000 epsilon:
 0.05944258000852662 steps: 436 lr: 2.5600000000000001e-06 evaluation
 reward: 11.66
 episode: 2026 score: 8.0 memory length: 100000 epsilon:
 0.058610980008527186 steps: 420 lr: 2.5600000000000001e-06 evaluation
 reward: 11.7
 episode: 2027 score: 14.0 memory length: 100000 epsilon:
 0.05753980000852792 steps: 541 lr: 2.5600000000000001e-06 evaluation
 reward: 11.79
 episode: 2028 score: 11.0 memory length: 100000 epsilon:
 0.05644486000852866 steps: 553 lr: 2.5600000000000001e-06 evaluation
 reward: 11.79
 episode: 2029 score: 11.0 memory length: 100000 epsilon:
 0.055357840008529405 steps: 549 lr: 2.5600000000000001e-06 evaluation
 reward: 11.76
 episode: 2030 score: 7.0 memory length: 100000 epsilon:
 0.05464306000852989 steps: 361 lr: 2.5600000000000001e-06 evaluation
 reward: 11.71
 episode: 2031 score: 10.0 memory length: 100000 epsilon:
 0.05368870000853054 steps: 482 lr: 2.5600000000000001e-06 evaluation
 reward: 11.62
 episode: 2032 score: 3.0 memory length: 100000 epsilon: 0.0531640000085309
 steps: 265 lr: 2.5600000000000001e-06 evaluation reward: 11.5
 episode: 2033 score: 8.0 memory length: 100000 epsilon:
 0.052338340008531464 steps: 417 lr: 2.5600000000000001e-06 evaluation
 reward: 11.39
 episode: 2034 score: 7.0 memory length: 100000 epsilon:
 0.051560200008531995 steps: 393 lr: 2.5600000000000001e-06 evaluation
 reward: 11.4
 episode: 2035 score: 10.0 memory length: 100000 epsilon:
 0.05049694000853272 steps: 537 lr: 2.5600000000000001e-06 evaluation
 reward: 11.39
 episode: 2036 score: 15.0 memory length: 100000 epsilon:
 0.049301020008533536 steps: 604 lr: 2.5600000000000001e-06 evaluation

reward: 11.47				
episode: 2037	score: 11.0	memory length: 100000	epsilon: 0.04823974000853426	
	steps: 536	lr: 2.5600000000000001e-06		evaluation
reward: 11.45				
episode: 2038	score: 14.0	memory length: 100000	epsilon: 0.04718242000853498	
	steps: 534	lr: 2.5600000000000001e-06		evaluation
reward: 11.49				
episode: 2039	score: 14.0	memory length: 100000	epsilon: 0.04615282000853568	
	steps: 520	lr: 2.5600000000000001e-06		evaluation
reward: 11.44				
episode: 2040	score: 17.0	memory length: 100000	epsilon: 0.0448103800085366	
	steps: 678	lr: 2.5600000000000001e-06		evaluation
reward: 11.55				
episode: 2041	score: 12.0	memory length: 100000	epsilon: 0.04353328000853747	
	steps: 645	lr: 2.5600000000000001e-06		evaluation
reward: 11.61				
episode: 2042	score: 6.0	memory length: 100000	epsilon: 0.04284424000853794	
	steps: 348	lr: 2.5600000000000001e-06		evaluation
reward: 11.51				
episode: 2043	score: 10.0	memory length: 100000	epsilon: 0.04175722000853868	
	steps: 549	lr: 2.5600000000000001e-06		evaluation
reward: 11.54				
episode: 2044	score: 7.0	memory length: 100000	epsilon: 0.04087612000853928	
	steps: 445	lr: 2.5600000000000001e-06		evaluation
reward: 11.51				
episode: 2045	score: 9.0	memory length: 100000	epsilon: 0.04001284000853987	
	steps: 436	lr: 2.5600000000000001e-06		evaluation
reward: 11.5				
episode: 2046	score: 15.0	memory length: 100000	epsilon: 0.03891988000854062	
	steps: 552	lr: 2.5600000000000001e-06		evaluation
reward: 11.51				
episode: 2047	score: 24.0	memory length: 100000	epsilon: 0.03731608000854171	
	steps: 810	lr: 2.5600000000000001e-06		evaluation
reward: 11.67				
episode: 2048	score: 13.0	memory length: 100000	epsilon: 0.03639538000854234	
	steps: 465	lr: 2.5600000000000001e-06		evaluation
reward: 11.63				
episode: 2049	score: 19.0	memory length: 100000	epsilon: 0.03514798000854319	
	steps: 630	lr: 2.5600000000000001e-06		evaluation
reward: 11.74				
episode: 2050	score: 13.0	memory length: 100000	epsilon: 0.034199560008543836	
	steps: 479	lr: 2.5600000000000001e-06		evaluation
reward: 11.74				
episode: 2051	score: 26.0	memory length: 100000	epsilon: 0.032720500008544845	
	steps: 747	lr: 2.5600000000000001e-06		evaluation
reward: 11.9				
episode: 2052	score: 7.0	memory length: 100000	epsilon: 0.031900780008545404	
	steps: 414	lr: 2.5600000000000001e-06		evaluation

reward: 11.88		
episode: 2053	score: 2.0 memory length: 100000 epsilon: 0.03149884000854568 steps: 203 lr: 2.5600000000000001e-06	evaluation
reward: 11.76		
episode: 2054	score: 4.0 memory length: 100000 epsilon: 0.030916720008546075 steps: 294 lr: 2.5600000000000001e-06	evaluation
reward: 11.72		
episode: 2055	score: 5.0 memory length: 100000 epsilon: 0.03026530000854652 steps: 329 lr: 2.5600000000000001e-06	evaluation
reward: 11.62		
episode: 2056	score: 12.0 memory length: 100000 epsilon: 0.029316880008547166 steps: 479 lr: 2.5600000000000001e-06	evaluation
reward: 11.6		
episode: 2057	score: 8.0 memory length: 100000 epsilon: 0.02847538000854774 steps: 425 lr: 2.5600000000000001e-06	evaluation
reward: 11.53		
episode: 2058	score: 12.0 memory length: 100000 epsilon: 0.02727352000854856 steps: 607 lr: 2.5600000000000001e-06	evaluation
reward: 11.6		
episode: 2059	score: 14.0 memory length: 100000 epsilon: 0.025931080008549476 steps: 678 lr: 2.5600000000000001e-06	evaluation
reward: 11.64		
episode: 2060	score: 13.0 memory length: 100000 epsilon: 0.024756940008550277 steps: 593 lr: 2.5600000000000001e-06	evaluation
reward: 11.67		
episode: 2061	score: 7.0 memory length: 100000 epsilon: 0.023966920008550816 steps: 399 lr: 2.5600000000000001e-06	evaluation
reward: 11.68		
episode: 2062	score: 7.0 memory length: 100000 epsilon: 0.02319868000855134 steps: 388 lr: 2.5600000000000001e-06	evaluation
reward: 11.65		
episode: 2063	score: 9.0 memory length: 100000 epsilon: 0.022388860008551892 steps: 409 lr: 2.5600000000000001e-06	evaluation
reward: 11.52		
episode: 2064	score: 7.0 memory length: 100000 epsilon: 0.021517660008552486 steps: 440 lr: 2.5600000000000001e-06	evaluation
reward: 11.49		
episode: 2065	score: 15.0 memory length: 100000 epsilon: 0.020086120008553462 steps: 723 lr: 2.5600000000000001e-06	evaluation
reward: 11.58		
episode: 2066	score: 10.0 memory length: 100000 epsilon: 0.01919512000855407 steps: 450 lr: 2.5600000000000001e-06	evaluation
reward: 11.56		
episode: 2067	score: 7.0 memory length: 100000 epsilon: 0.01846054000855457 steps: 371 lr: 2.5600000000000001e-06	evaluation
reward: 11.57		
episode: 2068	score: 10.0 memory length: 100000 epsilon: 0.017470540008555246 steps: 500 lr: 2.5600000000000001e-06	evaluation

reward: 11.59				
episode: 2069	score: 11.0	memory length: 100000	epsilon: 0.01661518000855583	steps: 432
			lr: 2.5600000000000001e-06	evaluation
reward: 11.52				
episode: 2070	score: 7.0	memory length: 100000	epsilon: 0.015856840008556347	steps: 383
			lr: 2.5600000000000001e-06	evaluation
reward: 11.46				
episode: 2071	score: 11.0	memory length: 100000	epsilon: 0.014841100008556353	steps: 513
			lr: 2.5600000000000001e-06	evaluation
reward: 11.42				
episode: 2072	score: 12.0	memory length: 100000	epsilon: 0.013696660008556131	steps: 578
			lr: 2.5600000000000001e-06	evaluation
reward: 11.42				
episode: 2073	score: 16.0	memory length: 100000	epsilon: 0.012486880008555896	steps: 611
			lr: 2.5600000000000001e-06	evaluation
reward: 11.47				
episode: 2074	score: 7.0	memory length: 100000	epsilon: 0.011637460008555731	steps: 429
			lr: 2.5600000000000001e-06	evaluation
reward: 11.42				
episode: 2075	score: 11.0	memory length: 100000	epsilon: 0.010762300008555561	steps: 442
			lr: 2.5600000000000001e-06	evaluation
reward: 11.46				
episode: 2076	score: 13.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 599
			lr: 1.0240000000000005e-06	evaluation
reward: 11.46				
episode: 2077	score: 17.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 741
			lr: 1.0240000000000005e-06	evaluation
reward: 11.53				
episode: 2078	score: 16.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 684
			lr: 1.0240000000000005e-06	evaluation
reward: 11.57				
episode: 2079	score: 5.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 363
			lr: 1.0240000000000005e-06	evaluation
reward: 11.53				
episode: 2080	score: 10.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 488
			lr: 1.0240000000000005e-06	evaluation
reward: 11.54				
episode: 2081	score: 12.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 542
			lr: 1.0240000000000005e-06	evaluation
reward: 11.56				
episode: 2082	score: 8.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 421
			lr: 1.0240000000000005e-06	evaluation
reward: 11.45				
episode: 2083	score: 8.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 439
			lr: 1.0240000000000005e-06	evaluation
reward: 11.46				
episode: 2084	score: 9.0	memory length: 100000	epsilon: 0.009998020008555413	steps: 447
			lr: 1.0240000000000005e-06	evaluation

```

reward: 11.43
episode: 2085    score: 19.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 650    lr: 1.0240000000000005e-06    evaluation
reward: 11.55
episode: 2086    score: 10.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 510    lr: 1.0240000000000005e-06    evaluation
reward: 11.47
episode: 2087    score: 17.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 651    lr: 1.0240000000000005e-06    evaluation
reward: 11.46
episode: 2088    score: 18.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 631    lr: 1.0240000000000005e-06    evaluation
reward: 11.5
episode: 2089    score: 10.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 523    lr: 1.0240000000000005e-06    evaluation
reward: 11.49
episode: 2090    score: 14.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 626    lr: 1.0240000000000005e-06    evaluation
reward: 11.55
episode: 2091    score: 8.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 421    lr: 1.0240000000000005e-06    evaluation
reward: 11.52
episode: 2092    score: 17.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 588    lr: 1.0240000000000005e-06    evaluation
reward: 11.5
episode: 2093    score: 4.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 295    lr: 1.0240000000000005e-06    evaluation
reward: 11.48
episode: 2094    score: 6.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 377    lr: 1.0240000000000005e-06    evaluation
reward: 11.38
episode: 2095    score: 6.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 340    lr: 1.0240000000000005e-06    evaluation
reward: 11.36
episode: 2096    score: 12.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 545    lr: 1.0240000000000005e-06    evaluation
reward: 11.31
episode: 2097    score: 14.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 724    lr: 1.0240000000000005e-06    evaluation
reward: 11.32
episode: 2098    score: 10.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 491    lr: 1.0240000000000005e-06    evaluation
reward: 11.23
episode: 2099    score: 13.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 648    lr: 1.0240000000000005e-06    evaluation
reward: 11.3
episode: 2100    score: 8.0    memory length: 100000    epsilon:
0.009998020008555413    steps: 490    lr: 1.0240000000000005e-06    evaluation

```

```

reward: 11.31
episode: 2101  score: 10.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 562  lr: 1.0240000000000005e-06  evaluation
reward: 11.28
episode: 2102  score: 10.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 498  lr: 1.0240000000000005e-06  evaluation
reward: 11.21
episode: 2103  score: 17.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 627  lr: 1.0240000000000005e-06  evaluation
reward: 11.29
episode: 2104  score: 10.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 542  lr: 1.0240000000000005e-06  evaluation
reward: 11.21
episode: 2105  score: 8.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 473  lr: 1.0240000000000005e-06  evaluation
reward: 11.19
episode: 2106  score: 8.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 470  lr: 1.0240000000000005e-06  evaluation
reward: 11.18
episode: 2107  score: 20.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 764  lr: 1.0240000000000005e-06  evaluation
reward: 11.26
episode: 2108  score: 19.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 728  lr: 1.0240000000000005e-06  evaluation
reward: 11.33
episode: 2109  score: 8.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 445  lr: 1.0240000000000005e-06  evaluation
reward: 11.29
episode: 2110  score: 10.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 527  lr: 1.0240000000000005e-06  evaluation
reward: 11.25
episode: 2111  score: 9.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 500  lr: 1.0240000000000005e-06  evaluation
reward: 11.23
episode: 2112  score: 14.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 605  lr: 1.0240000000000005e-06  evaluation
reward: 11.18
episode: 2113  score: 8.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 415  lr: 1.0240000000000005e-06  evaluation
reward: 11.2
episode: 2114  score: 26.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 816  lr: 1.0240000000000005e-06  evaluation
reward: 11.19
episode: 2115  score: 6.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 375  lr: 1.0240000000000005e-06  evaluation
reward: 11.15
episode: 2116  score: 16.0  memory length: 100000  epsilon:
0.009998020008555413  steps: 622  lr: 1.0240000000000005e-06  evaluation

```

reward: 11.15

```

└─
└─-----
KeyboardInterrupt                                Traceback (most recent call
└─last)

<ipython-input-10-8583aeff7434> in <module>()
    54         pylab.ylabel('Rewards')
    55         pylab.title('Episodes vs Reward')
--> 56         pylab.savefig("./save_graph/breakout_dqn.png") # save
└─graph for training visualization
    57
    58         # every episode, plot the play time

/usr/local/lib/python3.6/dist-packages/matplotlib/pyplot.py in
└─savefig(*args, **kwargs)
    722     fig = gcf()
    723     res = fig.savefig(*args, **kwargs)
--> 724     fig.canvas.draw_idle() # need this if 'transparent=True' to
└─reset colors
    725     return res
    726

/usr/local/lib/python3.6/dist-packages/matplotlib/backend_bases.py in
└─draw_idle(self, *args, **kwargs)
   1945         if not self._is_idle_drawing:
   1946             with self._idle_draw_cntx():
-> 1947                 self.draw(*args, **kwargs)
   1948
   1949         @cbook.deprecated("3.2")

/usr/local/lib/python3.6/dist-packages/matplotlib/backends/backend_agg.
└─py in draw(self)
    391         (self.toolbar._wait_cursor_for_draw_cm() if self.toolbar
    392         else nullcontext()):
--> 393         self.figure.draw(self.renderer)
    394         # A GUI class may be need to update a window using this
└─draw, so
    395         # don't forget to call the superclass.
```

```

/usr/local/lib/python3.6/dist-packages/matplotlib/artist.py in
↳ draw_wrapper(artist, renderer, *args, **kwargs)
    36         renderer.start_filter()
    37
---> 38         return draw(artist, renderer, *args, **kwargs)
    39     finally:
    40         if artist.get_agg_filter() is not None:

/usr/local/lib/python3.6/dist-packages/matplotlib/figure.py in
↳ draw(self, renderer)
    1734         self.patch.draw(renderer)
    1735         mimage._draw_list_compositing_images(
-> 1736             renderer, self, artists, self.suppressComposite)
    1737
    1738         renderer.close_group('figure')

/usr/local/lib/python3.6/dist-packages/matplotlib/image.py in
↳ _draw_list_compositing_images(renderer, parent, artists, suppress_composite)
    135     if not_composite or not has_images:
    136         for a in artists:
--> 137             a.draw(renderer)
    138     else:
    139         # Composite any adjacent images together

/usr/local/lib/python3.6/dist-packages/matplotlib/artist.py in
↳ draw_wrapper(artist, renderer, *args, **kwargs)
    36         renderer.start_filter()
    37
---> 38         return draw(artist, renderer, *args, **kwargs)
    39     finally:
    40         if artist.get_agg_filter() is not None:

/usr/local/lib/python3.6/dist-packages/matplotlib/axes/_base.py in
↳ draw(self, renderer, inframe)
    2628         renderer.stop_rasterizing()
    2629
-> 2630         mimage._draw_list_compositing_images(renderer, self, artists)
    2631
    2632         renderer.close_group('axes')

/usr/local/lib/python3.6/dist-packages/matplotlib/image.py in
↳ _draw_list_compositing_images(renderer, parent, artists, suppress_composite)

```

```

135     if not_composite or not has_images:
136         for a in artists:
--> 137             a.draw(renderer)
138     else:
139         # Composite any adjacent images together

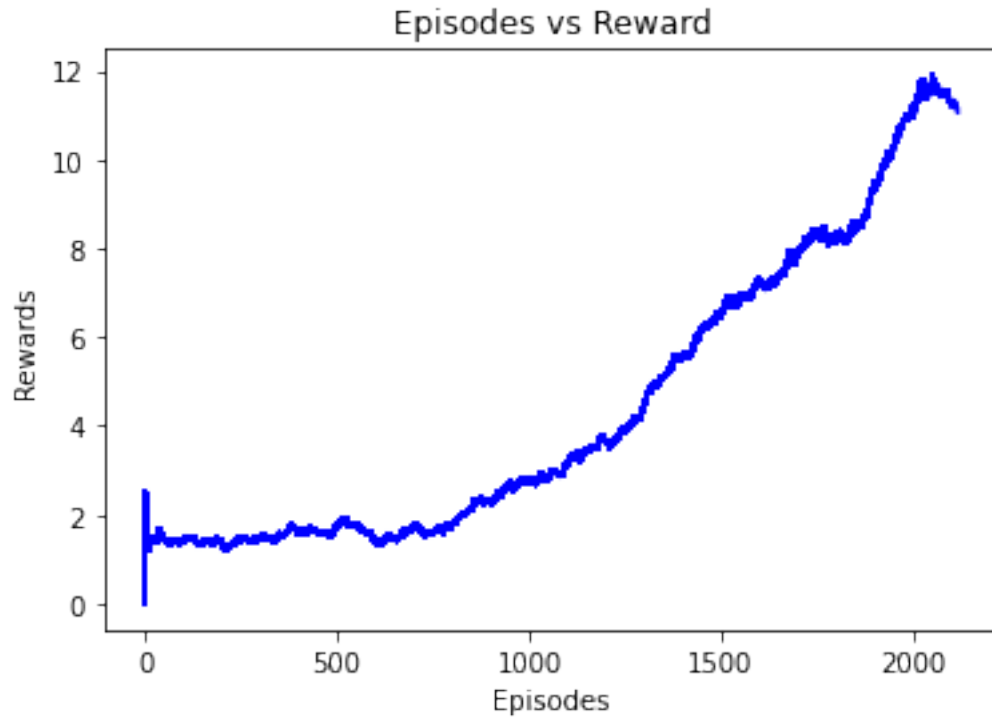
/usr/local/lib/python3.6/dist-packages/matplotlib/artist.py in
↳draw_wrapper(artist, renderer, *args, **kwargs)
    36         renderer.start_filter()
    37
---> 38         return draw(artist, renderer, *args, **kwargs)
    39     finally:
    40         if artist.get_agg_filter() is not None:

/usr/local/lib/python3.6/dist-packages/matplotlib/lines.py in draw(self,
↳renderer)
    800
    801         gc.set_dashes(self._dashOffset, self._dashSeq)
--> 802         renderer.draw_path(gc, tpath, affine.frozen())
    803         gc.restore()
    804

/usr/local/lib/python3.6/dist-packages/matplotlib/backends/backend_agg.
↳py in draw_path(self, gc, path, transform, rgbFace)
    150     else:
    151         try:
--> 152             self._renderer.draw_path(gc, path, transform,
↳rgbFace)
    153         except OverflowError:
    154             raise OverflowError("Exceeded cell block limit (set "

```

KeyboardInterrupt:



2 Visualize Agent Performance

BE AWARE THIS CODE BELOW MAY CRASH THE KERNEL IF YOU RUN THE SAME CELL TWICE.

Please save your model before running this portion of the code.

```
[ ]: torch.save(agent.policy_net, "./save_model/breakout_dqn_latest.pth")
```

```
[ ]: from gym.wrappers import Monitor
import glob
import io
import base64

from IPython.display import HTML
from IPython import display as ipythondisplay

from pyvirtualdisplay import Display

# Displaying the game live
def show_state(env, step=0, info=""):
    plt.figure(3)
    plt.clf()
    plt.imshow(env.render(mode='rgb_array'))
    plt.title("%s | Step: %d %s" % ("Agent Playing", step, info))
```

```

plt.axis('off')

ipythondisplay.clear_output(wait=True)
ipythondisplay.display(plt.gcf())

# Recording the game and replaying the game afterwards
def show_video():
    mp4list = glob.glob('video/*.mp4')
    if len(mp4list) > 0:
        mp4 = mp4list[0]
        video = io.open(mp4, 'r+b').read()
        encoded = base64.b64encode(video)
        ipythondisplay.display(HTML(data='''<video alt="test" autoplay
            loop controls style="height: 400px;">
            <source src="data:video/mp4;base64,{0}" type="video/mp4" />
            </video>''' .format(encoded.decode('ascii'))))
    else:
        print("Could not find video")

def wrap_env(env):
    env = Monitor(env, './video', force=True)
    return env

```

```

[:]: ### single DQN

display = Display(visible=0, size=(300, 200))
display.start()

# Load agent
agent.load_policy_net("./save_model/breakout_dqn.pth")
agent.epsilon = 0.0 # Set agent to only exploit the best action

env = gym.make('BreakoutDeterministic-v4')
env = wrap_env(env)

done = False
score = 0
step = 0
state = env.reset()
next_state = state
life = number_lives
history = np.zeros([5, 84, 84], dtype=np.uint8)
get_init_state(history, state)

while not done:

```



```

# Render breakout
# env.render()
show_state(env,step) # uncommenting this provides another way to visualize
→the game

step += 1
frame += 1

# Perform a fire action if ball is no longer on screen
if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
    action = 0
else:
    print(np.float32(np.float32(history[None, :4, :, :]).shape))
    action = agent.get_action((np.float32(history[None, :4, :, :]) / 255.)).
→cuda()
    # action = agent.get_action((np.float32(history[None, :4, :, :]) / 255.
→).cuda())
    state = next_state

next_state, reward, done, info = env.step(action + 1)

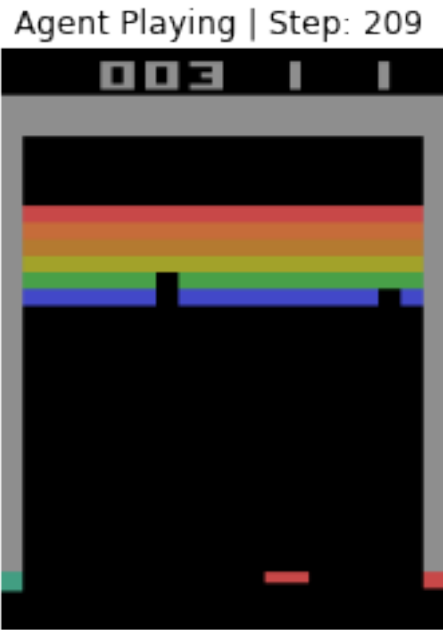
frame_next_state = get_frame(next_state)
history[4, :, :] = frame_next_state
terminal_state = check_live(life, info['ale.lives'])

life = info['ale.lives']
r = np.clip(reward, -1, 1)
r = reward

# Store the transition in memory
agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
# Start training after random sample generation
score += reward

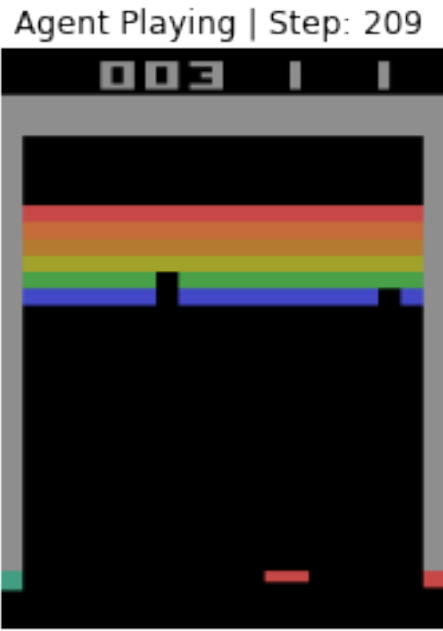
history[:4, :, :] = history[1:, :, :]
env.close()
show_video()
display.stop()

```



<IPython.core.display.HTML object>

[]: <pyvirtualdisplay.display.Display at 0x7f9d9dc341d0>



```

[:]: ### double DQN

display = Display(visible=0, size=(300, 200))
display.start()

# Load agent
agent.load_policy_net("./save_model/breakout_dqn.pth")
agent.epsilon = 0.0 # Set agent to only exploit the best action

env = gym.make('BreakoutDeterministic-v4')
env = wrap_env(env)

done = False
score = 0
step = 0
state = env.reset()
next_state = state
life = number_lives
history = np.zeros([5, 84, 84], dtype=np.uint8)
get_init_state(history, state)

while not done:

    # Render breakout
    # env.render()
    show_state(env, step) # uncommenting this provides another way to visualize
    → the game

    step += 1
    frame += 1

    # Perform a fire action if ball is no longer on screen
    if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
        action = 0
    else:
        print(np.float32(np.float32(history[None, :4, :, :]).shape))
        action = agent.get_action((np.float32(history[None, :4, :, :]) / 255.)).
        → cuda()
        # action = agent.get_action((np.float32(history[None, :4, :, :]) / 255.
        → ).cuda())
        state = next_state

    next_state, reward, done, info = env.step(action + 1)

    frame_next_state = get_frame(next_state)
    history[4, :, :] = frame_next_state
    terminal_state = check_live(life, info['ale.lives'])

```

```

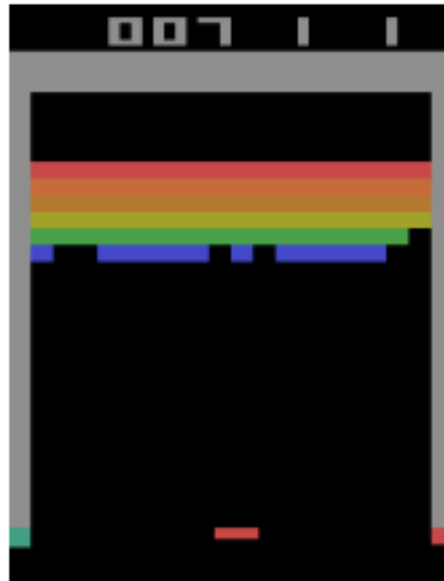
life = info['ale.lives']
r = np.clip(reward, -1, 1)
r = reward

# Store the transition in memory
agent.memory.push(deepcopy(frame_next_state), action, r, terminal_state)
# Start training after random sample generation
score += reward

history[:4, :, :] = history[1:, :, :]
env.close()
show_video()
display.stop()

```

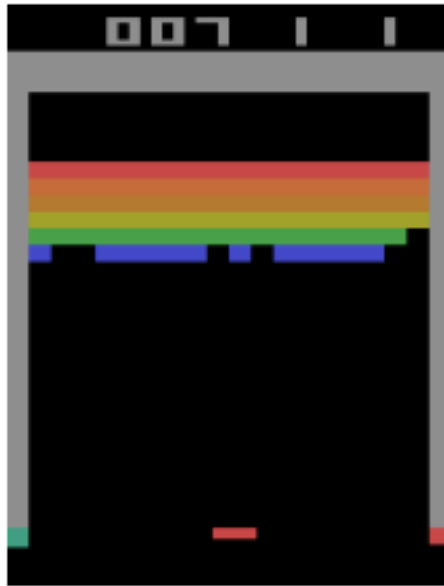
Agent Playing | Step: 421



<IPython.core.display.HTML object>

[]: <pyvirtualdisplay.display.Display at 0x7fe80e12f9b0>

Agent Playing | Step: 421



[]: