

MP5

December 5, 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.

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

```
Requirement already satisfied: gym in c:\users\computer\anaconda3\lib\site-
packages (0.17.3)
Requirement already satisfied: pyvirtualdisplay in
c:\users\computer\anaconda3\lib\site-packages (1.3.2)
Requirement already satisfied: scipy in c:\users\computer\anaconda3\lib\site-
packages (from gym) (1.5.0)
Requirement already satisfied: cloudpickle<1.7.0,>=1.2.0 in
c:\users\computer\anaconda3\lib\site-packages (from gym) (1.5.0)
Requirement already satisfied: pygamelet<=1.5.0,>=1.4.0 in
c:\users\computer\anaconda3\lib\site-packages (from gym) (1.5.0)
Requirement already satisfied: numpy>=1.10.4 in
c:\users\computer\anaconda3\lib\site-packages (from gym) (1.18.5)
Requirement already satisfied: EasyProcess in
c:\users\computer\anaconda3\lib\site-packages (from pyvirtualdisplay) (0.3)
Requirement already satisfied: future in c:\users\computer\anaconda3\lib\site-
packages (from pygamelet<=1.5.0,>=1.4.0->gym) (0.18.2)

'sudo' is not recognized as an internal or external command,
operable program or batch file.
```

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

```
Requirement already up-to-date: setuptools in
c:\users\computer\anaconda3\lib\site-packages (50.3.2)
Requirement already satisfied: ez_setup in c:\users\computer\anaconda3\lib\site-
packages (0.9)
Requirement already satisfied: gym[atari] in
c:\users\computer\anaconda3\lib\site-packages (0.17.3)
```

Requirement already satisfied: pygame<=1.5.0,>=1.4.0 in
 c:\users\computer\anaconda3\lib\site-packages (from gym[atari]) (1.5.0)
 Requirement already satisfied: scipy in c:\users\computer\anaconda3\lib\site-
 packages (from gym[atari]) (1.5.0)
 Requirement already satisfied: numpy>=1.10.4 in
 c:\users\computer\anaconda3\lib\site-packages (from gym[atari]) (1.18.5)
 Requirement already satisfied: cloudpickle<1.7.0,>=1.2.0 in
 c:\users\computer\anaconda3\lib\site-packages (from gym[atari]) (1.5.0)
 Requirement already satisfied: opencv-python; extra == "atari" in
 c:\users\computer\anaconda3\lib\site-packages (from gym[atari]) (4.4.0.46)
 Requirement already satisfied: atari-py~0.2.0; extra == "atari" in
 c:\users\computer\anaconda3\lib\site-packages (from gym[atari]) (0.2.6)
 Requirement already satisfied: Pillow; extra == "atari" in
 c:\users\computer\anaconda3\lib\site-packages (from gym[atari]) (7.2.0)
 Requirement already satisfied: future in c:\users\computer\anaconda3\lib\site-
 packages (from pygame<=1.5.0,>=1.4.0->gym[atari]) (0.18.2)
 Requirement already satisfied: six in c:\users\computer\anaconda3\lib\site-
 packages (from atari-py~0.2.0; extra == "atari"->gym[atari]) (1.15.0)

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**.

```
[1]: !sudo apt-get install -y xvfb python-opengl ffmpeg
```

'sudo' is not recognized as an internal or external command,
 operable program or batch file.

```
[1]: %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 matplotlib.pyplot as plt
```

```
# %load_ext autoreload
# %autoreload 2
```

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.

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

```
[3]: 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.

```
[4]: double_dqn = True # 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”

```

[5]: rewards, episodes = [], []
best_eval_reward = 0
for e in range(EPIISODES):
    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[: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_ddqn.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_ddqn.pth")
            best_eval_reward = np.mean(evaluation_reward)

```

```

episode: 0   score: 1.0   memory length: 151   epsilon: 1.0   steps: 151   lr:
0.0001      evaluation reward: 1.0
episode: 1   score: 4.0   memory length: 465   epsilon: 1.0   steps: 314   lr:
0.0001      evaluation reward: 2.5
episode: 2   score: 1.0   memory length: 636   epsilon: 1.0   steps: 171   lr:
0.0001      evaluation reward: 2.0
episode: 3   score: 0.0   memory length: 758   epsilon: 1.0   steps: 122   lr:
0.0001      evaluation reward: 1.5
episode: 4   score: 1.0   memory length: 929   epsilon: 1.0   steps: 171   lr:
0.0001      evaluation reward: 1.4
episode: 5   score: 1.0   memory length: 1080  epsilon: 1.0   steps: 151
lr: 0.0001   evaluation reward: 1.3333333333333333
episode: 6   score: 2.0   memory length: 1260  epsilon: 1.0   steps: 180
lr: 0.0001   evaluation reward: 1.4285714285714286
episode: 7   score: 0.0   memory length: 1383  epsilon: 1.0   steps: 123
lr: 0.0001   evaluation reward: 1.25
episode: 8   score: 2.0   memory length: 1600  epsilon: 1.0   steps: 217
lr: 0.0001   evaluation reward: 1.3333333333333333
episode: 9   score: 1.0   memory length: 1751  epsilon: 1.0   steps: 151
lr: 0.0001   evaluation reward: 1.3
episode: 10  score: 1.0   memory length: 1902  epsilon: 1.0   steps: 151

```

lr: 0.0001	evaluation reward: 1.2727272727272727			
episode: 11	score: 0.0	memory length: 2024	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.1666666666666667			
episode: 12	score: 2.0	memory length: 2223	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.2307692307692308			
episode: 13	score: 0.0	memory length: 2346	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.1428571428571428			
episode: 14	score: 1.0	memory length: 2497	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.1333333333333333			
episode: 15	score: 4.0	memory length: 2790	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.3125			
episode: 16	score: 1.0	memory length: 2959	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.2941176470588236			
episode: 17	score: 0.0	memory length: 3082	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2222222222222223			
episode: 18	score: 4.0	memory length: 3354	epsilon: 1.0	steps: 272
lr: 0.0001	evaluation reward: 1.368421052631579			
episode: 19	score: 0.0	memory length: 3477	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3			
episode: 20	score: 0.0	memory length: 3600	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2380952380952381			
episode: 21	score: 3.0	memory length: 3827	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.3181818181818181			
episode: 22	score: 0.0	memory length: 3949	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.2608695652173914			
episode: 23	score: 4.0	memory length: 4245	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.375			
episode: 24	score: 1.0	memory length: 4414	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.36			
episode: 25	score: 1.0	memory length: 4565	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3461538461538463			
episode: 26	score: 0.0	memory length: 4688	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2962962962962963			
episode: 27	score: 2.0	memory length: 4868	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.3214285714285714			
episode: 28	score: 0.0	memory length: 4991	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2758620689655173			
episode: 29	score: 2.0	memory length: 5207	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.3			
episode: 30	score: 0.0	memory length: 5330	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2580645161290323			
episode: 31	score: 1.0	memory length: 5481	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.25			
episode: 32	score: 0.0	memory length: 5604	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2121212121212122			
episode: 33	score: 2.0	memory length: 5824	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.2352941176470589			
episode: 34	score: 1.0	memory length: 5993	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.2285714285714286			
episode: 35	score: 1.0	memory length: 6143	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.2222222222222223			
episode: 36	score: 3.0	memory length: 6386	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.2702702702702702			
episode: 37	score: 3.0	memory length: 6612	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.3157894736842106			
episode: 38	score: 5.0	memory length: 6940	epsilon: 1.0	steps: 328
lr: 0.0001	evaluation reward: 1.4102564102564104			
episode: 39	score: 4.0	memory length: 7253	epsilon: 1.0	steps: 313
lr: 0.0001	evaluation reward: 1.475			
episode: 40	score: 2.0	memory length: 7451	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4878048780487805			
episode: 41	score: 1.0	memory length: 7619	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.4761904761904763			
episode: 42	score: 2.0	memory length: 7839	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.4883720930232558			
episode: 43	score: 3.0	memory length: 8083	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.5227272727272727			
episode: 44	score: 1.0	memory length: 8253	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5111111111111111			
episode: 45	score: 2.0	memory length: 8450	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.5217391304347827			
episode: 46	score: 0.0	memory length: 8572	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4893617021276595			
episode: 47	score: 0.0	memory length: 8695	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4583333333333333			
episode: 48	score: 0.0	memory length: 8818	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4285714285714286			
episode: 49	score: 0.0	memory length: 8941	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 50	score: 0.0	memory length: 9064	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3725490196078431			
episode: 51	score: 4.0	memory length: 9341	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.4230769230769231			
episode: 52	score: 2.0	memory length: 9539	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4339622641509433			
episode: 53	score: 0.0	memory length: 9662	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4074074074074074			
episode: 54	score: 0.0	memory length: 9785	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3818181818181818			
episode: 55	score: 0.0	memory length: 9908	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3571428571428572			
episode: 56	score: 2.0	memory length: 10106	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.368421052631579			
episode: 57	score: 3.0	memory length: 10372	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.396551724137931			
episode: 58	score: 1.0	memory length: 10543	epsilon: 1.0	steps: 171

lr: 0.0001	evaluation reward: 1.3898305084745763			
episode: 59	score: 2.0	memory length: 10762	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.4			
episode: 60	score: 0.0	memory length: 10885	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3770491803278688			
episode: 61	score: 4.0	memory length: 11161	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.4193548387096775			
episode: 62	score: 0.0	memory length: 11283	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3968253968253967			
episode: 63	score: 0.0	memory length: 11405	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.375			
episode: 64	score: 1.0	memory length: 11556	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3692307692307693			
episode: 65	score: 3.0	memory length: 11784	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.393939393939394			
episode: 66	score: 2.0	memory length: 12003	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.4029850746268657			
episode: 67	score: 1.0	memory length: 12171	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.3970588235294117			
episode: 68	score: 0.0	memory length: 12294	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3768115942028984			
episode: 69	score: 2.0	memory length: 12492	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.3857142857142857			
episode: 70	score: 1.0	memory length: 12643	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.380281690140845			
episode: 71	score: 2.0	memory length: 12859	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.3888888888888888			
episode: 72	score: 1.0	memory length: 13028	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.3835616438356164			
episode: 73	score: 5.0	memory length: 13336	epsilon: 1.0	steps: 308
lr: 0.0001	evaluation reward: 1.4324324324324325			
episode: 74	score: 3.0	memory length: 13564	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.4533333333333334			
episode: 75	score: 2.0	memory length: 13743	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.4605263157894737			
episode: 76	score: 2.0	memory length: 13940	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.4675324675324675			
episode: 77	score: 4.0	memory length: 14213	epsilon: 1.0	steps: 273
lr: 0.0001	evaluation reward: 1.5			
episode: 78	score: 2.0	memory length: 14411	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5063291139240507			
episode: 79	score: 3.0	memory length: 14658	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.525			
episode: 80	score: 1.0	memory length: 14809	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.5185185185185186			
episode: 81	score: 4.0	memory length: 15105	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.548780487804878			
episode: 82	score: 1.0	memory length: 15273	epsilon: 1.0	steps: 168

lr: 0.0001	evaluation reward: 1.5421686746987953		
episode: 83	score: 0.0	memory length: 15396	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5238095238095237		
episode: 84	score: 4.0	memory length: 15691	epsilon: 1.0 steps: 295
lr: 0.0001	evaluation reward: 1.5529411764705883		
episode: 85	score: 0.0	memory length: 15814	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5348837209302326		
episode: 86	score: 3.0	memory length: 16063	epsilon: 1.0 steps: 249
lr: 0.0001	evaluation reward: 1.5517241379310345		
episode: 87	score: 1.0	memory length: 16233	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.5454545454545454		
episode: 88	score: 2.0	memory length: 16450	epsilon: 1.0 steps: 217
lr: 0.0001	evaluation reward: 1.550561797752809		
episode: 89	score: 0.0	memory length: 16572	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.5333333333333334		
episode: 90	score: 2.0	memory length: 16769	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.5384615384615385		
episode: 91	score: 1.0	memory length: 16937	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.5326086956521738		
episode: 92	score: 0.0	memory length: 17059	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.5161290322580645		
episode: 93	score: 1.0	memory length: 17229	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.5106382978723405		
episode: 94	score: 0.0	memory length: 17351	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.4947368421052631		
episode: 95	score: 0.0	memory length: 17474	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4791666666666667		
episode: 96	score: 1.0	memory length: 17625	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.4742268041237114		
episode: 97	score: 2.0	memory length: 17823	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.4795918367346939		
episode: 98	score: 1.0	memory length: 17992	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.4747474747474747		
episode: 99	score: 2.0	memory length: 18190	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.48		
episode: 100	score: 0.0	memory length: 18313	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 101	score: 1.0	memory length: 18481	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.44		
episode: 102	score: 2.0	memory length: 18679	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.45		
episode: 103	score: 5.0	memory length: 19024	epsilon: 1.0 steps: 345
lr: 0.0001	evaluation reward: 1.5		
episode: 104	score: 4.0	memory length: 19344	epsilon: 1.0 steps: 320
lr: 0.0001	evaluation reward: 1.53		
episode: 105	score: 0.0	memory length: 19466	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.52		
episode: 106	score: 4.0	memory length: 19741	epsilon: 1.0 steps: 275

lr: 0.0001	evaluation reward: 1.54			
episode: 107	score: 0.0	memory length: 19864	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.54			
episode: 108	score: 3.0	memory length: 20127	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.55			
episode: 109	score: 1.0	memory length: 20278	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.55			
episode: 110	score: 0.0	memory length: 20401	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.54			
episode: 111	score: 1.0	memory length: 20572	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.55			
episode: 112	score: 0.0	memory length: 20695	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53			
episode: 113	score: 6.0	memory length: 21035	epsilon: 1.0	steps: 340
lr: 0.0001	evaluation reward: 1.59			
episode: 114	score: 1.0	memory length: 21204	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.59			
episode: 115	score: 1.0	memory length: 21355	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.56			
episode: 116	score: 1.0	memory length: 21507	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.56			
episode: 117	score: 4.0	memory length: 21823	epsilon: 1.0	steps: 316
lr: 0.0001	evaluation reward: 1.6			
episode: 118	score: 1.0	memory length: 21974	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.57			
episode: 119	score: 2.0	memory length: 22155	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.59			
episode: 120	score: 0.0	memory length: 22278	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 121	score: 2.0	memory length: 22475	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.58			
episode: 122	score: 2.0	memory length: 22673	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6			
episode: 123	score: 4.0	memory length: 22928	epsilon: 1.0	steps: 255
lr: 0.0001	evaluation reward: 1.6			
episode: 124	score: 0.0	memory length: 23051	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 125	score: 2.0	memory length: 23249	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6			
episode: 126	score: 2.0	memory length: 23451	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.62			
episode: 127	score: 2.0	memory length: 23649	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.62			
episode: 128	score: 1.0	memory length: 23818	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.63			
episode: 129	score: 0.0	memory length: 23941	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61			
episode: 130	score: 0.0	memory length: 24064	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.61			
episode: 131	score: 3.0	memory length: 24296	epsilon: 1.0	steps: 232
lr: 0.0001	evaluation reward: 1.63			
episode: 132	score: 2.0	memory length: 24494	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65			
episode: 133	score: 3.0	memory length: 24743	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.66			
episode: 134	score: 3.0	memory length: 24995	epsilon: 1.0	steps: 252
lr: 0.0001	evaluation reward: 1.68			
episode: 135	score: 1.0	memory length: 25146	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.68			
episode: 136	score: 2.0	memory length: 25326	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.67			
episode: 137	score: 4.0	memory length: 25620	epsilon: 1.0	steps: 294
lr: 0.0001	evaluation reward: 1.68			
episode: 138	score: 6.0	memory length: 25980	epsilon: 1.0	steps: 360
lr: 0.0001	evaluation reward: 1.69			
episode: 139	score: 0.0	memory length: 26102	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.65			
episode: 140	score: 0.0	memory length: 26224	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.63			
episode: 141	score: 0.0	memory length: 26347	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62			
episode: 142	score: 0.0	memory length: 26469	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.6			
episode: 143	score: 3.0	memory length: 26695	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.6			
episode: 144	score: 2.0	memory length: 26912	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.61			
episode: 145	score: 0.0	memory length: 27035	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59			
episode: 146	score: 2.0	memory length: 27253	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.61			
episode: 147	score: 2.0	memory length: 27471	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.63			
episode: 148	score: 1.0	memory length: 27622	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.64			
episode: 149	score: 1.0	memory length: 27791	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.65			
episode: 150	score: 0.0	memory length: 27914	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 151	score: 4.0	memory length: 28192	epsilon: 1.0	steps: 278
lr: 0.0001	evaluation reward: 1.65			
episode: 152	score: 1.0	memory length: 28361	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.64			
episode: 153	score: 0.0	memory length: 28484	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 154	score: 2.0	memory length: 28683	epsilon: 1.0	steps: 199

lr: 0.0001	evaluation reward: 1.66			
episode: 155	score: 0.0	memory length: 28806	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66			
episode: 156	score: 2.0	memory length: 29004	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.66			
episode: 157	score: 2.0	memory length: 29202	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65			
episode: 158	score: 0.0	memory length: 29325	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 159	score: 4.0	memory length: 29597	epsilon: 1.0	steps: 272
lr: 0.0001	evaluation reward: 1.66			
episode: 160	score: 0.0	memory length: 29720	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66			
episode: 161	score: 4.0	memory length: 29994	epsilon: 1.0	steps: 274
lr: 0.0001	evaluation reward: 1.66			
episode: 162	score: 1.0	memory length: 30166	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.67			
episode: 163	score: 6.0	memory length: 30562	epsilon: 1.0	steps: 396
lr: 0.0001	evaluation reward: 1.73			
episode: 164	score: 2.0	memory length: 30781	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.74			
episode: 165	score: 4.0	memory length: 31077	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.75			
episode: 166	score: 4.0	memory length: 31352	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.77			
episode: 167	score: 0.0	memory length: 31474	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.76			
episode: 168	score: 2.0	memory length: 31692	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.78			
episode: 169	score: 1.0	memory length: 31861	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.77			
episode: 170	score: 1.0	memory length: 32033	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.77			
episode: 171	score: 3.0	memory length: 32263	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.78			
episode: 172	score: 1.0	memory length: 32434	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.78			
episode: 173	score: 0.0	memory length: 32557	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.73			
episode: 174	score: 2.0	memory length: 32755	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.72			
episode: 175	score: 2.0	memory length: 32955	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.72			
episode: 176	score: 2.0	memory length: 33170	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.72			
episode: 177	score: 2.0	memory length: 33370	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.7			
episode: 178	score: 0.0	memory length: 33493	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.68			
episode: 179	score: 1.0	memory length: 33665	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.66			
episode: 180	score: 0.0	memory length: 33788	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 181	score: 2.0	memory length: 33970	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.63			
episode: 182	score: 3.0	memory length: 34198	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.65			
episode: 183	score: 1.0	memory length: 34367	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.66			
episode: 184	score: 0.0	memory length: 34490	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62			
episode: 185	score: 2.0	memory length: 34688	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.64			
episode: 186	score: 3.0	memory length: 34950	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.64			
episode: 187	score: 12.0	memory length: 35441	epsilon: 1.0	steps: 491
lr: 0.0001	evaluation reward: 1.75			
episode: 188	score: 0.0	memory length: 35564	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.73			
episode: 189	score: 0.0	memory length: 35687	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.73			
episode: 190	score: 2.0	memory length: 35884	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.73			
episode: 191	score: 2.0	memory length: 36065	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.74			
episode: 192	score: 1.0	memory length: 36217	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.75			
episode: 193	score: 0.0	memory length: 36339	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.74			
episode: 194	score: 2.0	memory length: 36556	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.76			
episode: 195	score: 0.0	memory length: 36679	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.76			
episode: 196	score: 2.0	memory length: 36877	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.77			
episode: 197	score: 0.0	memory length: 37000	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.75			
episode: 198	score: 0.0	memory length: 37122	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.74			
episode: 199	score: 1.0	memory length: 37273	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.73			
episode: 200	score: 0.0	memory length: 37396	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.73			
episode: 201	score: 2.0	memory length: 37595	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.74			
episode: 202	score: 2.0	memory length: 37792	epsilon: 1.0	steps: 197

lr: 0.0001	evaluation reward: 1.74			
episode: 203	score: 2.0	memory length: 37989	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.71			
episode: 204	score: 0.0	memory length: 38112	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67			
episode: 205	score: 1.0	memory length: 38281	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.68			
episode: 206	score: 0.0	memory length: 38403	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.64			
episode: 207	score: 0.0	memory length: 38526	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 208	score: 1.0	memory length: 38696	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62			
episode: 209	score: 2.0	memory length: 38894	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.63			
episode: 210	score: 0.0	memory length: 39016	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.63			
episode: 211	score: 0.0	memory length: 39139	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62			
episode: 212	score: 2.0	memory length: 39337	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.64			
episode: 213	score: 3.0	memory length: 39569	epsilon: 1.0	steps: 232
lr: 0.0001	evaluation reward: 1.61			
episode: 214	score: 2.0	memory length: 39767	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.62			
episode: 215	score: 1.0	memory length: 39937	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62			
episode: 216	score: 3.0	memory length: 40205	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.64			
episode: 217	score: 1.0	memory length: 40375	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.61			
episode: 218	score: 0.0	memory length: 40498	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6			
episode: 219	score: 0.0	memory length: 40620	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.58			
episode: 220	score: 2.0	memory length: 40817	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.6			
episode: 221	score: 2.0	memory length: 41015	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6			
episode: 222	score: 2.0	memory length: 41231	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.6			
episode: 223	score: 0.0	memory length: 41354	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56			
episode: 224	score: 3.0	memory length: 41581	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.59			
episode: 225	score: 2.0	memory length: 41799	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.59			
episode: 226	score: 0.0	memory length: 41922	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.57			
episode: 227	score: 2.0	memory length: 42122	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.57			
episode: 228	score: 0.0	memory length: 42244	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.56			
episode: 229	score: 0.0	memory length: 42366	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.56			
episode: 230	score: 2.0	memory length: 42564	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.58			
episode: 231	score: 0.0	memory length: 42687	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55			
episode: 232	score: 1.0	memory length: 42856	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.54			
episode: 233	score: 3.0	memory length: 43103	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.54			
episode: 234	score: 1.0	memory length: 43254	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.52			
episode: 235	score: 2.0	memory length: 43454	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.53			
episode: 236	score: 0.0	memory length: 43576	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.51			
episode: 237	score: 0.0	memory length: 43698	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.47			
episode: 238	score: 0.0	memory length: 43821	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41			
episode: 239	score: 1.0	memory length: 43990	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.42			
episode: 240	score: 8.0	memory length: 44339	epsilon: 1.0	steps: 349
lr: 0.0001	evaluation reward: 1.5			
episode: 241	score: 0.0	memory length: 44462	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5			
episode: 242	score: 1.0	memory length: 44630	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.51			
episode: 243	score: 0.0	memory length: 44753	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 244	score: 1.0	memory length: 44922	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.47			
episode: 245	score: 1.0	memory length: 45073	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48			
episode: 246	score: 2.0	memory length: 45253	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.48			
episode: 247	score: 2.0	memory length: 45469	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.48			
episode: 248	score: 2.0	memory length: 45687	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.49			
episode: 249	score: 0.0	memory length: 45810	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 250	score: 2.0	memory length: 46028	epsilon: 1.0	steps: 218

lr: 0.0001	evaluation reward: 1.5			
episode: 251	score: 1.0	memory length: 46178	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.47			
episode: 252	score: 1.0	memory length: 46347	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.47			
episode: 253	score: 1.0	memory length: 46498	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48			
episode: 254	score: 4.0	memory length: 46774	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.5			
episode: 255	score: 0.0	memory length: 46897	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5			
episode: 256	score: 1.0	memory length: 47048	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.49			
episode: 257	score: 1.0	memory length: 47219	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.48			
episode: 258	score: 2.0	memory length: 47400	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.5			
episode: 259	score: 1.0	memory length: 47568	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.47			
episode: 260	score: 1.0	memory length: 47719	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48			
episode: 261	score: 1.0	memory length: 47891	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.45			
episode: 262	score: 0.0	memory length: 48014	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 263	score: 2.0	memory length: 48194	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.4			
episode: 264	score: 2.0	memory length: 48414	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.4			
episode: 265	score: 2.0	memory length: 48629	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.38			
episode: 266	score: 7.0	memory length: 48949	epsilon: 1.0	steps: 320
lr: 0.0001	evaluation reward: 1.41			
episode: 267	score: 2.0	memory length: 49166	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.43			
episode: 268	score: 2.0	memory length: 49364	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.43			
episode: 269	score: 3.0	memory length: 49633	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.45			
episode: 270	score: 2.0	memory length: 49831	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.46			
episode: 271	score: 3.0	memory length: 50075	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.46			
episode: 272	score: 2.0	memory length: 50256	epsilon: 1.0	steps: 181
lr: 0.0001	evaluation reward: 1.47			
episode: 273	score: 1.0	memory length: 50424	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.48			
episode: 274	score: 0.0	memory length: 50546	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.46			
episode: 275	score: 0.0	memory length: 50668	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44			
episode: 276	score: 2.0	memory length: 50865	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.44			
episode: 277	score: 4.0	memory length: 51141	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.46			
episode: 278	score: 0.0	memory length: 51263	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.46			
episode: 279	score: 3.0	memory length: 51496	epsilon: 1.0	steps: 233
lr: 0.0001	evaluation reward: 1.48			
episode: 280	score: 2.0	memory length: 51676	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.5			
episode: 281	score: 3.0	memory length: 51922	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.51			
episode: 282	score: 3.0	memory length: 52167	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.51			
episode: 283	score: 1.0	memory length: 52338	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.51			
episode: 284	score: 0.0	memory length: 52460	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.51			
episode: 285	score: 2.0	memory length: 52676	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.51			
episode: 286	score: 3.0	memory length: 52944	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.51			
episode: 287	score: 0.0	memory length: 53067	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39			
episode: 288	score: 1.0	memory length: 53217	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4			
episode: 289	score: 0.0	memory length: 53339	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4			
episode: 290	score: 1.0	memory length: 53508	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39			
episode: 291	score: 2.0	memory length: 53725	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.39			
episode: 292	score: 1.0	memory length: 53894	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39			
episode: 293	score: 4.0	memory length: 54190	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.43			
episode: 294	score: 2.0	memory length: 54407	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.43			
episode: 295	score: 2.0	memory length: 54591	epsilon: 1.0	steps: 184
lr: 0.0001	evaluation reward: 1.45			
episode: 296	score: 1.0	memory length: 54741	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.44			
episode: 297	score: 0.0	memory length: 54864	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 298	score: 0.0	memory length: 54986	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.44			
episode: 299	score: 2.0	memory length: 55166	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.45			
episode: 300	score: 2.0	memory length: 55364	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47			
episode: 301	score: 2.0	memory length: 55581	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.47			
episode: 302	score: 1.0	memory length: 55749	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.46			
episode: 303	score: 0.0	memory length: 55872	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 304	score: 1.0	memory length: 56041	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.45			
episode: 305	score: 1.0	memory length: 56212	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.45			
episode: 306	score: 0.0	memory length: 56335	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45			
episode: 307	score: 1.0	memory length: 56507	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.46			
episode: 308	score: 0.0	memory length: 56629	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45			
episode: 309	score: 0.0	memory length: 56752	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43			
episode: 310	score: 1.0	memory length: 56903	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.44			
episode: 311	score: 3.0	memory length: 57151	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.47			
episode: 312	score: 2.0	memory length: 57369	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.47			
episode: 313	score: 1.0	memory length: 57537	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.45			
episode: 314	score: 0.0	memory length: 57659	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43			
episode: 315	score: 0.0	memory length: 57781	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.42			
episode: 316	score: 1.0	memory length: 57952	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.4			
episode: 317	score: 2.0	memory length: 58172	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.41			
episode: 318	score: 0.0	memory length: 58295	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41			
episode: 319	score: 2.0	memory length: 58492	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.43			
episode: 320	score: 0.0	memory length: 58614	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41			
episode: 321	score: 1.0	memory length: 58783	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.4			
episode: 322	score: 1.0	memory length: 58934	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.39			
episode: 323	score: 1.0	memory length: 59102	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.4			
episode: 324	score: 1.0	memory length: 59271	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.38			
episode: 325	score: 0.0	memory length: 59394	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36			
episode: 326	score: 2.0	memory length: 59592	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.38			
episode: 327	score: 3.0	memory length: 59820	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.39			
episode: 328	score: 1.0	memory length: 59992	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.4			
episode: 329	score: 1.0	memory length: 60164	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.41			
episode: 330	score: 3.0	memory length: 60395	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.42			
episode: 331	score: 0.0	memory length: 60517	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.42			
episode: 332	score: 0.0	memory length: 60639	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41			
episode: 333	score: 1.0	memory length: 60790	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.39			
episode: 334	score: 0.0	memory length: 60913	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38			
episode: 335	score: 0.0	memory length: 61036	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36			
episode: 336	score: 0.0	memory length: 61159	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36			
episode: 337	score: 1.0	memory length: 61309	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.37			
episode: 338	score: 0.0	memory length: 61432	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 339	score: 1.0	memory length: 61601	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.37			
episode: 340	score: 3.0	memory length: 61829	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.32			
episode: 341	score: 4.0	memory length: 62092	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.36			
episode: 342	score: 0.0	memory length: 62215	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35			
episode: 343	score: 2.0	memory length: 62413	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.37			
episode: 344	score: 0.0	memory length: 62535	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36			
episode: 345	score: 3.0	memory length: 62779	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.38			
episode: 346	score: 2.0	memory length: 62959	epsilon: 1.0	steps: 180

lr: 0.0001	evaluation reward: 1.38			
episode: 347	score: 4.0	memory length: 63273	epsilon: 1.0	steps: 314
lr: 0.0001	evaluation reward: 1.4			
episode: 348	score: 4.0	memory length: 63548	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.42			
episode: 349	score: 3.0	memory length: 63798	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.45			
episode: 350	score: 1.0	memory length: 63970	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.44			
episode: 351	score: 2.0	memory length: 64168	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 352	score: 1.0	memory length: 64319	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.45			
episode: 353	score: 0.0	memory length: 64442	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 354	score: 0.0	memory length: 64564	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4			
episode: 355	score: 4.0	memory length: 64839	epsilon: 1.0	steps: 275
lr: 0.0001	evaluation reward: 1.44			
episode: 356	score: 2.0	memory length: 65037	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 357	score: 1.0	memory length: 65207	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.45			
episode: 358	score: 2.0	memory length: 65405	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 359	score: 1.0	memory length: 65556	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.45			
episode: 360	score: 0.0	memory length: 65679	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 361	score: 2.0	memory length: 65898	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.45			
episode: 362	score: 2.0	memory length: 66096	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47			
episode: 363	score: 1.0	memory length: 66246	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.46			
episode: 364	score: 2.0	memory length: 66444	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.46			
episode: 365	score: 3.0	memory length: 66691	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.47			
episode: 366	score: 2.0	memory length: 66889	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.42			
episode: 367	score: 1.0	memory length: 67060	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.41			
episode: 368	score: 1.0	memory length: 67211	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.4			
episode: 369	score: 5.0	memory length: 67506	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.42			
episode: 370	score: 1.0	memory length: 67657	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.41			
episode: 371	score: 1.0	memory length: 67808	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.39			
episode: 372	score: 2.0	memory length: 67988	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.39			
episode: 373	score: 2.0	memory length: 68204	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.4			
episode: 374	score: 4.0	memory length: 68521	epsilon: 1.0	steps: 317
lr: 0.0001	evaluation reward: 1.44			
episode: 375	score: 2.0	memory length: 68740	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.46			
episode: 376	score: 3.0	memory length: 68969	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.47			
episode: 377	score: 0.0	memory length: 69092	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43			
episode: 378	score: 2.0	memory length: 69311	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.45			
episode: 379	score: 1.0	memory length: 69483	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.43			
episode: 380	score: 0.0	memory length: 69606	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41			
episode: 381	score: 1.0	memory length: 69756	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.39			
episode: 382	score: 3.0	memory length: 69985	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.39			
episode: 383	score: 1.0	memory length: 70154	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39			
episode: 384	score: 2.0	memory length: 70370	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.41			
episode: 385	score: 1.0	memory length: 70521	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.4			
episode: 386	score: 0.0	memory length: 70644	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 387	score: 1.0	memory length: 70794	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.38			
episode: 388	score: 0.0	memory length: 70917	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 389	score: 0.0	memory length: 71039	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37			
episode: 390	score: 0.0	memory length: 71162	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36			
episode: 391	score: 1.0	memory length: 71331	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.35			
episode: 392	score: 1.0	memory length: 71500	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.35			
episode: 393	score: 1.0	memory length: 71650	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.32			
episode: 394	score: 0.0	memory length: 71772	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.3			
episode: 395	score: 0.0	memory length: 71895	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.28			
episode: 396	score: 1.0	memory length: 72046	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.28			
episode: 397	score: 2.0	memory length: 72264	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.3			
episode: 398	score: 2.0	memory length: 72483	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.32			
episode: 399	score: 1.0	memory length: 72652	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.31			
episode: 400	score: 0.0	memory length: 72775	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.29			
episode: 401	score: 1.0	memory length: 72927	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.28			
episode: 402	score: 0.0	memory length: 73050	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.27			
episode: 403	score: 1.0	memory length: 73218	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.28			
episode: 404	score: 2.0	memory length: 73415	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.29			
episode: 405	score: 0.0	memory length: 73538	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.28			
episode: 406	score: 2.0	memory length: 73736	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.3			
episode: 407	score: 1.0	memory length: 73905	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.3			
episode: 408	score: 0.0	memory length: 74028	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3			
episode: 409	score: 3.0	memory length: 74297	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.33			
episode: 410	score: 1.0	memory length: 74468	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.33			
episode: 411	score: 1.0	memory length: 74619	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.31			
episode: 412	score: 2.0	memory length: 74837	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.31			
episode: 413	score: 2.0	memory length: 75055	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.32			
episode: 414	score: 0.0	memory length: 75178	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32			
episode: 415	score: 3.0	memory length: 75442	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.35			
episode: 416	score: 2.0	memory length: 75640	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.36			
episode: 417	score: 3.0	memory length: 75867	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.37			
episode: 418	score: 0.0	memory length: 75989	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.37			
episode: 419	score: 2.0	memory length: 76205	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.37			
episode: 420	score: 3.0	memory length: 76452	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.4			
episode: 421	score: 2.0	memory length: 76650	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41			
episode: 422	score: 0.0	memory length: 76773	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 423	score: 1.0	memory length: 76923	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4			
episode: 424	score: 1.0	memory length: 77091	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.4			
episode: 425	score: 0.0	memory length: 77214	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 426	score: 0.0	memory length: 77336	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38			
episode: 427	score: 5.0	memory length: 77673	epsilon: 1.0	steps: 337
lr: 0.0001	evaluation reward: 1.4			
episode: 428	score: 2.0	memory length: 77894	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.41			
episode: 429	score: 0.0	memory length: 78017	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 430	score: 2.0	memory length: 78236	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.39			
episode: 431	score: 2.0	memory length: 78433	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.41			
episode: 432	score: 0.0	memory length: 78555	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41			
episode: 433	score: 3.0	memory length: 78801	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.43			
episode: 434	score: 0.0	memory length: 78924	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43			
episode: 435	score: 1.0	memory length: 79096	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.44			
episode: 436	score: 0.0	memory length: 79219	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 437	score: 2.0	memory length: 79436	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.45			
episode: 438	score: 3.0	memory length: 79682	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.48			
episode: 439	score: 2.0	memory length: 79880	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49			
episode: 440	score: 2.0	memory length: 80095	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.48			
episode: 441	score: 0.0	memory length: 80218	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 442	score: 3.0	memory length: 80470	epsilon: 1.0	steps: 252

lr: 0.0001	evaluation reward: 1.47			
episode: 443	score: 2.0	memory length: 80670	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.47			
episode: 444	score: 1.0	memory length: 80839	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.48			
episode: 445	score: 5.0	memory length: 81166	epsilon: 1.0	steps: 327
lr: 0.0001	evaluation reward: 1.5			
episode: 446	score: 3.0	memory length: 81413	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.51			
episode: 447	score: 2.0	memory length: 81629	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.49			
episode: 448	score: 6.0	memory length: 81979	epsilon: 1.0	steps: 350
lr: 0.0001	evaluation reward: 1.51			
episode: 449	score: 0.0	memory length: 82102	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 450	score: 3.0	memory length: 82367	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.5			
episode: 451	score: 0.0	memory length: 82490	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 452	score: 0.0	memory length: 82612	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.47			
episode: 453	score: 1.0	memory length: 82763	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48			
episode: 454	score: 0.0	memory length: 82885	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.48			
episode: 455	score: 4.0	memory length: 83161	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.48			
episode: 456	score: 0.0	memory length: 83284	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46			
episode: 457	score: 3.0	memory length: 83494	epsilon: 1.0	steps: 210
lr: 0.0001	evaluation reward: 1.48			
episode: 458	score: 2.0	memory length: 83710	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.48			
episode: 459	score: 0.0	memory length: 83833	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 460	score: 0.0	memory length: 83956	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 461	score: 0.0	memory length: 84078	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45			
episode: 462	score: 4.0	memory length: 84371	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.47			
episode: 463	score: 0.0	memory length: 84494	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46			
episode: 464	score: 0.0	memory length: 84617	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 465	score: 1.0	memory length: 84768	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.42			
episode: 466	score: 1.0	memory length: 84919	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.41			
episode: 467	score: 2.0	memory length: 85136	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.42			
episode: 468	score: 2.0	memory length: 85334	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.43			
episode: 469	score: 1.0	memory length: 85485	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.39			
episode: 470	score: 0.0	memory length: 85608	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38			
episode: 471	score: 2.0	memory length: 85806	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.39			
episode: 472	score: 0.0	memory length: 85929	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 473	score: 0.0	memory length: 86051	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.35			
episode: 474	score: 0.0	memory length: 86174	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31			
episode: 475	score: 1.0	memory length: 86343	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.3			
episode: 476	score: 5.0	memory length: 86692	epsilon: 1.0	steps: 349
lr: 0.0001	evaluation reward: 1.32			
episode: 477	score: 1.0	memory length: 86861	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.33			
episode: 478	score: 3.0	memory length: 87109	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.34			
episode: 479	score: 0.0	memory length: 87231	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.33			
episode: 480	score: 2.0	memory length: 87429	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.35			
episode: 481	score: 0.0	memory length: 87552	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.34			
episode: 482	score: 0.0	memory length: 87675	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31			
episode: 483	score: 4.0	memory length: 87992	epsilon: 1.0	steps: 317
lr: 0.0001	evaluation reward: 1.34			
episode: 484	score: 1.0	memory length: 88161	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.33			
episode: 485	score: 2.0	memory length: 88359	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.34			
episode: 486	score: 2.0	memory length: 88557	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.36			
episode: 487	score: 0.0	memory length: 88680	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35			
episode: 488	score: 2.0	memory length: 88898	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.37			
episode: 489	score: 2.0	memory length: 89096	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.39			
episode: 490	score: 1.0	memory length: 89265	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.4			
episode: 491	score: 2.0	memory length: 89484	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.41			
episode: 492	score: 4.0	memory length: 89764	epsilon: 1.0	steps: 280
lr: 0.0001	evaluation reward: 1.44			
episode: 493	score: 2.0	memory length: 89983	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.45			
episode: 494	score: 0.0	memory length: 90106	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45			
episode: 495	score: 1.0	memory length: 90274	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.46			
episode: 496	score: 0.0	memory length: 90397	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45			
episode: 497	score: 0.0	memory length: 90519	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43			
episode: 498	score: 0.0	memory length: 90641	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41			
episode: 499	score: 0.0	memory length: 90764	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 500	score: 1.0	memory length: 90914	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.41			
episode: 501	score: 1.0	memory length: 91083	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.41			
episode: 502	score: 3.0	memory length: 91296	epsilon: 1.0	steps: 213
lr: 0.0001	evaluation reward: 1.44			
episode: 503	score: 2.0	memory length: 91494	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 504	score: 2.0	memory length: 91692	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 505	score: 0.0	memory length: 91815	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45			
episode: 506	score: 1.0	memory length: 91984	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.44			
episode: 507	score: 2.0	memory length: 92181	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.45			
episode: 508	score: 2.0	memory length: 92399	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.47			
episode: 509	score: 1.0	memory length: 92569	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.45			
episode: 510	score: 2.0	memory length: 92769	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.46			
episode: 511	score: 0.0	memory length: 92891	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45			
episode: 512	score: 0.0	memory length: 93014	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.43			
episode: 513	score: 0.0	memory length: 93136	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41			
episode: 514	score: 1.0	memory length: 93287	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.42			
episode: 515	score: 1.0	memory length: 93456	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.4			
episode: 516	score: 1.0	memory length: 93607	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.39			
episode: 517	score: 1.0	memory length: 93758	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.37			
episode: 518	score: 0.0	memory length: 93881	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 519	score: 0.0	memory length: 94003	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.35			
episode: 520	score: 0.0	memory length: 94125	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.32			
episode: 521	score: 2.0	memory length: 94323	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.32			
episode: 522	score: 0.0	memory length: 94445	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.32			
episode: 523	score: 0.0	memory length: 94568	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31			
episode: 524	score: 2.0	memory length: 94789	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.32			
episode: 525	score: 0.0	memory length: 94911	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.32			
episode: 526	score: 2.0	memory length: 95129	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.34			
episode: 527	score: 0.0	memory length: 95251	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.29			
episode: 528	score: 0.0	memory length: 95373	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.27			
episode: 529	score: 3.0	memory length: 95606	epsilon: 1.0	steps: 233
lr: 0.0001	evaluation reward: 1.3			
episode: 530	score: 3.0	memory length: 95833	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.31			
episode: 531	score: 1.0	memory length: 96002	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.3			
episode: 532	score: 0.0	memory length: 96125	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3			
episode: 533	score: 0.0	memory length: 96248	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.27			
episode: 534	score: 2.0	memory length: 96445	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.29			
episode: 535	score: 0.0	memory length: 96568	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.28			
episode: 536	score: 4.0	memory length: 96884	epsilon: 1.0	steps: 316
lr: 0.0001	evaluation reward: 1.32			
episode: 537	score: 3.0	memory length: 97131	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.33			
episode: 538	score: 3.0	memory length: 97396	epsilon: 1.0	steps: 265

```

lr: 0.0001      evaluation reward: 1.33
episode: 539    score: 2.0    memory length: 97575    epsilon: 1.0    steps: 179
lr: 0.0001      evaluation reward: 1.33
episode: 540    score: 2.0    memory length: 97793    epsilon: 1.0    steps: 218
lr: 0.0001      evaluation reward: 1.33
episode: 541    score: 0.0    memory length: 97916    epsilon: 1.0    steps: 123
lr: 0.0001      evaluation reward: 1.33
episode: 542    score: 2.0    memory length: 98132    epsilon: 1.0    steps: 216
lr: 0.0001      evaluation reward: 1.32
episode: 543    score: 2.0    memory length: 98330    epsilon: 1.0    steps: 198
lr: 0.0001      evaluation reward: 1.32
episode: 544    score: 4.0    memory length: 98626    epsilon: 1.0    steps: 296
lr: 0.0001      evaluation reward: 1.35
episode: 545    score: 0.0    memory length: 98749    epsilon: 1.0    steps: 123
lr: 0.0001      evaluation reward: 1.3
episode: 546    score: 1.0    memory length: 98918    epsilon: 1.0    steps: 169
lr: 0.0001      evaluation reward: 1.28
episode: 547    score: 1.0    memory length: 99086    epsilon: 1.0    steps: 168
lr: 0.0001      evaluation reward: 1.27
episode: 548    score: 0.0    memory length: 99209    epsilon: 1.0    steps: 123
lr: 0.0001      evaluation reward: 1.21
episode: 549    score: 2.0    memory length: 99410    epsilon: 1.0    steps: 201
lr: 0.0001      evaluation reward: 1.23
episode: 550    score: 6.0    memory length: 99786    epsilon: 1.0    steps: 376
lr: 0.0001      evaluation reward: 1.26
episode: 551    score: 1.0    memory length: 99955    epsilon: 1.0    steps: 169
lr: 0.0001      evaluation reward: 1.27

```

C:\Users\Computer\Desktop\fall2020\cs498dl\assignment5\memory.py:30:

VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray

```
sample = np.array(sample)
```

C:\Users\Computer\Desktop\fall2020\cs498dl\assignment5\agent_double.py:74:

VisibleDeprecationWarning: Creating an ndarray from ragged nested sequences (which is a list-or-tuple of lists-or-tuples-or ndarrays with different lengths or shapes) is deprecated. If you meant to do this, you must specify 'dtype=object' when creating the ndarray

```
mini_batch = np.array(mini_batch).transpose()
```

```

episode: 552    score: 1.0    memory length: 100124    epsilon: 0.9997525000000054
steps: 169      lr: 0.0001      evaluation reward: 1.28
episode: 553    score: 0.0    memory length: 100246    epsilon: 0.9995109400000106
steps: 122      lr: 0.0001      evaluation reward: 1.27
episode: 554    score: 2.0    memory length: 100465    epsilon: 0.999077320000002
steps: 219      lr: 0.0001      evaluation reward: 1.29
episode: 555    score: 1.0    memory length: 100633    epsilon: 0.9987446800000273
steps: 168      lr: 0.0001      evaluation reward: 1.26

```

episode: 556	score: 0.0	memory length: 100755	epsilon: 0.9985031200000325
steps: 122	lr: 0.0001	evaluation reward: 1.26	
episode: 557	score: 0.0	memory length: 100877	epsilon: 0.9982615600000377
steps: 122	lr: 0.0001	evaluation reward: 1.23	
episode: 558	score: 1.0	memory length: 101028	epsilon: 0.9979625800000442
steps: 151	lr: 0.0001	evaluation reward: 1.22	
episode: 559	score: 0.0	memory length: 101151	epsilon: 0.9977190400000495
steps: 123	lr: 0.0001	evaluation reward: 1.22	
episode: 560	score: 2.0	memory length: 101367	epsilon: 0.9972913600000588
steps: 216	lr: 0.0001	evaluation reward: 1.24	
episode: 561	score: 4.0	memory length: 101662	epsilon: 0.9967072600000715
steps: 295	lr: 0.0001	evaluation reward: 1.28	
episode: 562	score: 0.0	memory length: 101785	epsilon: 0.9964637200000768
steps: 123	lr: 0.0001	evaluation reward: 1.24	
episode: 563	score: 3.0	memory length: 102032	epsilon: 0.9959746600000874
steps: 247	lr: 0.0001	evaluation reward: 1.27	
episode: 564	score: 0.0	memory length: 102154	epsilon: 0.9957331000000926
steps: 122	lr: 0.0001	evaluation reward: 1.27	
episode: 565	score: 1.0	memory length: 102322	epsilon: 0.9954004600000999
steps: 168	lr: 0.0001	evaluation reward: 1.27	
episode: 566	score: 4.0	memory length: 102600	epsilon: 0.9948500200001118
steps: 278	lr: 0.0001	evaluation reward: 1.3	
episode: 567	score: 0.0	memory length: 102723	epsilon: 0.9946064800001171
steps: 123	lr: 0.0001	evaluation reward: 1.28	
episode: 568	score: 0.0	memory length: 102846	epsilon: 0.9943629400001224
steps: 123	lr: 0.0001	evaluation reward: 1.26	
episode: 569	score: 0.0	memory length: 102968	epsilon: 0.9941213800001276
steps: 122	lr: 0.0001	evaluation reward: 1.25	
episode: 570	score: 0.0	memory length: 103091	epsilon: 0.9938778400001329
steps: 123	lr: 0.0001	evaluation reward: 1.25	
episode: 571	score: 1.0	memory length: 103242	epsilon: 0.9935788600001394
steps: 151	lr: 0.0001	evaluation reward: 1.24	
episode: 572	score: 1.0	memory length: 103392	epsilon: 0.9932818600001458
steps: 150	lr: 0.0001	evaluation reward: 1.25	
episode: 573	score: 2.0	memory length: 103607	epsilon: 0.9928561600001551
steps: 215	lr: 0.0001	evaluation reward: 1.27	
episode: 574	score: 1.0	memory length: 103775	epsilon: 0.9925235200001623
steps: 168	lr: 0.0001	evaluation reward: 1.28	
episode: 575	score: 3.0	memory length: 104023	epsilon: 0.992032480000173
steps: 248	lr: 0.0001	evaluation reward: 1.3	
episode: 576	score: 1.0	memory length: 104192	epsilon: 0.9916978600001802
steps: 169	lr: 0.0001	evaluation reward: 1.26	
episode: 577	score: 1.0	memory length: 104361	epsilon: 0.9913632400001875
steps: 169	lr: 0.0001	evaluation reward: 1.26	
episode: 578	score: 3.0	memory length: 104605	epsilon: 0.990880120000198
steps: 244	lr: 0.0001	evaluation reward: 1.26	
episode: 579	score: 5.0	memory length: 104968	epsilon: 0.9901613800002136
steps: 363	lr: 0.0001	evaluation reward: 1.31	

episode: 580	score: 2.0	memory length: 105165	epsilon: 0.989771320000222
steps: 197	lr: 0.0001	evaluation reward: 1.31	
episode: 581	score: 1.0	memory length: 105336	epsilon: 0.9894327400002294
steps: 171	lr: 0.0001	evaluation reward: 1.32	
episode: 582	score: 0.0	memory length: 105459	epsilon: 0.9891892000002347
steps: 123	lr: 0.0001	evaluation reward: 1.32	
episode: 583	score: 0.0	memory length: 105581	epsilon: 0.9889476400002399
steps: 122	lr: 0.0001	evaluation reward: 1.28	
episode: 584	score: 2.0	memory length: 105799	epsilon: 0.9885160000002493
steps: 218	lr: 0.0001	evaluation reward: 1.29	
episode: 585	score: 3.0	memory length: 106066	epsilon: 0.9879873400002608
steps: 267	lr: 0.0001	evaluation reward: 1.3	
episode: 586	score: 1.0	memory length: 106234	epsilon: 0.987654700000268
steps: 168	lr: 0.0001	evaluation reward: 1.29	
episode: 587	score: 2.0	memory length: 106431	epsilon: 0.9872646400002765
steps: 197	lr: 0.0001	evaluation reward: 1.31	
episode: 588	score: 0.0	memory length: 106554	epsilon: 0.9870211000002818
steps: 123	lr: 0.0001	evaluation reward: 1.29	
episode: 589	score: 2.0	memory length: 106752	epsilon: 0.9866290600002903
steps: 198	lr: 0.0001	evaluation reward: 1.29	
episode: 590	score: 2.0	memory length: 106950	epsilon: 0.9862370200002988
steps: 198	lr: 0.0001	evaluation reward: 1.3	
episode: 591	score: 2.0	memory length: 107148	epsilon: 0.9858449800003073
steps: 198	lr: 0.0001	evaluation reward: 1.3	
episode: 592	score: 1.0	memory length: 107317	epsilon: 0.9855103600003146
steps: 169	lr: 0.0001	evaluation reward: 1.27	
episode: 593	score: 5.0	memory length: 107627	epsilon: 0.9848965600003279
steps: 310	lr: 0.0001	evaluation reward: 1.3	
episode: 594	score: 1.0	memory length: 107778	epsilon: 0.9845975800003344
steps: 151	lr: 0.0001	evaluation reward: 1.31	
episode: 595	score: 3.0	memory length: 108005	epsilon: 0.9841481200003441
steps: 227	lr: 0.0001	evaluation reward: 1.33	
episode: 596	score: 0.0	memory length: 108127	epsilon: 0.9839065600003494
steps: 122	lr: 0.0001	evaluation reward: 1.33	
episode: 597	score: 2.0	memory length: 108346	epsilon: 0.9834729400003588
steps: 219	lr: 0.0001	evaluation reward: 1.35	
episode: 598	score: 1.0	memory length: 108514	epsilon: 0.983140300000366
steps: 168	lr: 0.0001	evaluation reward: 1.36	
episode: 599	score: 1.0	memory length: 108664	epsilon: 0.9828433000003725
steps: 150	lr: 0.0001	evaluation reward: 1.37	
episode: 600	score: 0.0	memory length: 108786	epsilon: 0.9826017400003777
steps: 122	lr: 0.0001	evaluation reward: 1.36	
episode: 601	score: 0.0	memory length: 108909	epsilon: 0.982358200000383
steps: 123	lr: 0.0001	evaluation reward: 1.35	
episode: 602	score: 1.0	memory length: 109060	epsilon: 0.9820592200003895
steps: 151	lr: 0.0001	evaluation reward: 1.33	
episode: 603	score: 0.0	memory length: 109182	epsilon: 0.9818176600003947
steps: 122	lr: 0.0001	evaluation reward: 1.31	

episode: 604	score: 0.0	memory length: 109304	epsilon: 0.98157610000004
steps: 122	lr: 0.0001	evaluation reward: 1.29	
episode: 605	score: 1.0	memory length: 109455	epsilon: 0.98127712000004065
steps: 151	lr: 0.0001	evaluation reward: 1.3	
episode: 606	score: 2.0	memory length: 109652	epsilon: 0.98088706000004149
steps: 197	lr: 0.0001	evaluation reward: 1.31	
episode: 607	score: 1.0	memory length: 109820	epsilon: 0.98055442000004221
steps: 168	lr: 0.0001	evaluation reward: 1.3	
episode: 608	score: 0.0	memory length: 109943	epsilon: 0.98031088000004274
steps: 123	lr: 0.0001	evaluation reward: 1.28	
episode: 609	score: 0.0	memory length: 110065	epsilon: 0.98006932000004327
steps: 122	lr: 0.0001	evaluation reward: 1.27	
episode: 610	score: 3.0	memory length: 110293	epsilon: 0.97961788000004425
steps: 228	lr: 0.0001	evaluation reward: 1.28	
episode: 611	score: 1.0	memory length: 110445	epsilon: 0.9793169200000449
steps: 152	lr: 0.0001	evaluation reward: 1.29	
episode: 612	score: 0.0	memory length: 110568	epsilon: 0.97907338000004543
steps: 123	lr: 0.0001	evaluation reward: 1.29	
episode: 613	score: 1.0	memory length: 110740	epsilon: 0.97873282000004617
steps: 172	lr: 0.0001	evaluation reward: 1.3	
episode: 614	score: 4.0	memory length: 111015	epsilon: 0.97818832000004735
steps: 275	lr: 0.0001	evaluation reward: 1.33	
episode: 615	score: 1.0	memory length: 111165	epsilon: 0.977891320000048
steps: 150	lr: 0.0001	evaluation reward: 1.33	
episode: 616	score: 0.0	memory length: 111288	epsilon: 0.97764778000004852
steps: 123	lr: 0.0001	evaluation reward: 1.32	
episode: 617	score: 1.0	memory length: 111457	epsilon: 0.97731316000004925
steps: 169	lr: 0.0001	evaluation reward: 1.32	
episode: 618	score: 2.0	memory length: 111675	epsilon: 0.97688152000005019
steps: 218	lr: 0.0001	evaluation reward: 1.34	
episode: 619	score: 1.0	memory length: 111843	epsilon: 0.97654888000005091
steps: 168	lr: 0.0001	evaluation reward: 1.35	
episode: 620	score: 2.0	memory length: 112059	epsilon: 0.97612120000005184
steps: 216	lr: 0.0001	evaluation reward: 1.37	
episode: 621	score: 2.0	memory length: 112257	epsilon: 0.97572916000005269
steps: 198	lr: 0.0001	evaluation reward: 1.37	
episode: 622	score: 0.0	memory length: 112380	epsilon: 0.97548562000005322
steps: 123	lr: 0.0001	evaluation reward: 1.37	
episode: 623	score: 3.0	memory length: 112645	epsilon: 0.97496092000005436
steps: 265	lr: 0.0001	evaluation reward: 1.4	
episode: 624	score: 0.0	memory length: 112767	epsilon: 0.97471936000005488
steps: 122	lr: 0.0001	evaluation reward: 1.38	
episode: 625	score: 5.0	memory length: 113055	epsilon: 0.97414912000005612
steps: 288	lr: 0.0001	evaluation reward: 1.43	
episode: 626	score: 6.0	memory length: 113453	epsilon: 0.97336108000005783
steps: 398	lr: 0.0001	evaluation reward: 1.47	
episode: 627	score: 1.0	memory length: 113622	epsilon: 0.97302646000005856
steps: 169	lr: 0.0001	evaluation reward: 1.48	

episode: 628	score: 1.0	memory length: 113791	epsilon: 0.9726918400005928
steps: 169	lr: 0.0001	evaluation reward: 1.49	
episode: 629	score: 0.0	memory length: 113913	epsilon: 0.9724502800005981
steps: 122	lr: 0.0001	evaluation reward: 1.46	
episode: 630	score: 1.0	memory length: 114064	epsilon: 0.9721513000006046
steps: 151	lr: 0.0001	evaluation reward: 1.44	
episode: 631	score: 2.0	memory length: 114281	epsilon: 0.9717216400006139
steps: 217	lr: 0.0001	evaluation reward: 1.45	
episode: 632	score: 3.0	memory length: 114548	epsilon: 0.9711929800006254
steps: 267	lr: 0.0001	evaluation reward: 1.48	
episode: 633	score: 2.0	memory length: 114746	epsilon: 0.9708009400006339
steps: 198	lr: 0.0001	evaluation reward: 1.5	
episode: 634	score: 1.0	memory length: 114918	epsilon: 0.9704603800006413
steps: 172	lr: 0.0001	evaluation reward: 1.49	
episode: 635	score: 0.0	memory length: 115041	epsilon: 0.9702168400006466
steps: 123	lr: 0.0001	evaluation reward: 1.49	
episode: 636	score: 1.0	memory length: 115192	epsilon: 0.969917860000653
steps: 151	lr: 0.0001	evaluation reward: 1.46	
episode: 637	score: 1.0	memory length: 115343	epsilon: 0.9696188800006595
steps: 151	lr: 0.0001	evaluation reward: 1.44	
episode: 638	score: 1.0	memory length: 115512	epsilon: 0.9692842600006668
steps: 169	lr: 0.0001	evaluation reward: 1.42	
episode: 639	score: 2.0	memory length: 115710	epsilon: 0.9688922200006753
steps: 198	lr: 0.0001	evaluation reward: 1.42	
episode: 640	score: 0.0	memory length: 115833	epsilon: 0.9686486800006806
steps: 123	lr: 0.0001	evaluation reward: 1.4	
episode: 641	score: 3.0	memory length: 116076	epsilon: 0.968167540000691
steps: 243	lr: 0.0001	evaluation reward: 1.43	
episode: 642	score: 2.0	memory length: 116274	epsilon: 0.9677755000006996
steps: 198	lr: 0.0001	evaluation reward: 1.43	
episode: 643	score: 2.0	memory length: 116456	epsilon: 0.9674151400007074
steps: 182	lr: 0.0001	evaluation reward: 1.43	
episode: 644	score: 1.0	memory length: 116607	epsilon: 0.9671161600007139
steps: 151	lr: 0.0001	evaluation reward: 1.4	
episode: 645	score: 3.0	memory length: 116857	epsilon: 0.9666211600007246
steps: 250	lr: 0.0001	evaluation reward: 1.43	
episode: 646	score: 0.0	memory length: 116979	epsilon: 0.9663796000007299
steps: 122	lr: 0.0001	evaluation reward: 1.42	
episode: 647	score: 5.0	memory length: 117315	epsilon: 0.9657143200007443
steps: 336	lr: 0.0001	evaluation reward: 1.46	
episode: 648	score: 1.0	memory length: 117486	epsilon: 0.9653757400007517
steps: 171	lr: 0.0001	evaluation reward: 1.47	
episode: 649	score: 1.0	memory length: 117636	epsilon: 0.9650787400007581
steps: 150	lr: 0.0001	evaluation reward: 1.46	
episode: 650	score: 3.0	memory length: 117881	epsilon: 0.9645936400007686
steps: 245	lr: 0.0001	evaluation reward: 1.43	
episode: 651	score: 3.0	memory length: 118126	epsilon: 0.9641085400007792
steps: 245	lr: 0.0001	evaluation reward: 1.45	

episode: 652	score: 2.0	memory length: 118344	epsilon: 0.9636769000007885
steps: 218	lr: 0.0001	evaluation reward: 1.46	
episode: 653	score: 3.0	memory length: 118572	epsilon: 0.9632254600007983
steps: 228	lr: 0.0001	evaluation reward: 1.49	
episode: 654	score: 4.0	memory length: 118866	epsilon: 0.962643340000811
steps: 294	lr: 0.0001	evaluation reward: 1.51	
episode: 655	score: 1.0	memory length: 119016	epsilon: 0.9623463400008174
steps: 150	lr: 0.0001	evaluation reward: 1.51	
episode: 656	score: 0.0	memory length: 119139	epsilon: 0.9621028000008227
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 657	score: 4.0	memory length: 119437	epsilon: 0.9615127600008355
steps: 298	lr: 0.0001	evaluation reward: 1.55	
episode: 658	score: 2.0	memory length: 119654	epsilon: 0.9610831000008448
steps: 217	lr: 0.0001	evaluation reward: 1.56	
episode: 659	score: 1.0	memory length: 119823	epsilon: 0.9607484800008521
steps: 169	lr: 0.0001	evaluation reward: 1.57	
episode: 660	score: 3.0	memory length: 120053	epsilon: 0.960293080000862
steps: 230	lr: 0.0001	evaluation reward: 1.58	
episode: 661	score: 1.0	memory length: 120225	epsilon: 0.9599525200008694
steps: 172	lr: 0.0001	evaluation reward: 1.55	
episode: 662	score: 1.0	memory length: 120394	epsilon: 0.9596179000008767
steps: 169	lr: 0.0001	evaluation reward: 1.56	
episode: 663	score: 1.0	memory length: 120562	epsilon: 0.9592852600008839
steps: 168	lr: 0.0001	evaluation reward: 1.54	
episode: 664	score: 1.0	memory length: 120713	epsilon: 0.9589862800008904
steps: 151	lr: 0.0001	evaluation reward: 1.55	
episode: 665	score: 4.0	memory length: 120988	epsilon: 0.9584417800009022
steps: 275	lr: 0.0001	evaluation reward: 1.58	
episode: 666	score: 2.0	memory length: 121186	epsilon: 0.9580497400009107
steps: 198	lr: 0.0001	evaluation reward: 1.56	
episode: 667	score: 1.0	memory length: 121355	epsilon: 0.957715120000918
steps: 169	lr: 0.0001	evaluation reward: 1.57	
episode: 668	score: 0.0	memory length: 121477	epsilon: 0.9574735600009232
steps: 122	lr: 0.0001	evaluation reward: 1.57	
episode: 669	score: 1.0	memory length: 121630	epsilon: 0.9571706200009298
steps: 153	lr: 0.0001	evaluation reward: 1.58	
episode: 670	score: 0.0	memory length: 121752	epsilon: 0.956929060000935
steps: 122	lr: 0.0001	evaluation reward: 1.58	
episode: 671	score: 1.0	memory length: 121921	epsilon: 0.9565944400009423
steps: 169	lr: 0.0001	evaluation reward: 1.58	
episode: 672	score: 2.0	memory length: 122121	epsilon: 0.9561984400009509
steps: 200	lr: 0.0001	evaluation reward: 1.59	
episode: 673	score: 0.0	memory length: 122243	epsilon: 0.9559568800009561
steps: 122	lr: 0.0001	evaluation reward: 1.57	
episode: 674	score: 1.0	memory length: 122414	epsilon: 0.9556183000009635
steps: 171	lr: 0.0001	evaluation reward: 1.57	
episode: 675	score: 3.0	memory length: 122681	epsilon: 0.955089640000975
steps: 267	lr: 0.0001	evaluation reward: 1.57	

episode: 676	score: 2.0	memory length: 122878	epsilon: 0.9546995800009834
steps: 197	lr: 0.0001	evaluation reward: 1.58	
episode: 677	score: 3.0	memory length: 123104	epsilon: 0.9542521000009931
steps: 226	lr: 0.0001	evaluation reward: 1.6	
episode: 678	score: 0.0	memory length: 123226	epsilon: 0.9540105400009984
steps: 122	lr: 0.0001	evaluation reward: 1.57	
episode: 679	score: 1.0	memory length: 123377	epsilon: 0.9537115600010049
steps: 151	lr: 0.0001	evaluation reward: 1.53	
episode: 680	score: 0.0	memory length: 123499	epsilon: 0.9534700000010101
steps: 122	lr: 0.0001	evaluation reward: 1.51	
episode: 681	score: 0.0	memory length: 123622	epsilon: 0.9532264600010154
steps: 123	lr: 0.0001	evaluation reward: 1.5	
episode: 682	score: 2.0	memory length: 123839	epsilon: 0.9527968000010247
steps: 217	lr: 0.0001	evaluation reward: 1.52	
episode: 683	score: 3.0	memory length: 124105	epsilon: 0.9522701200010362
steps: 266	lr: 0.0001	evaluation reward: 1.55	
episode: 684	score: 3.0	memory length: 124333	epsilon: 0.951818680001046
steps: 228	lr: 0.0001	evaluation reward: 1.56	
episode: 685	score: 0.0	memory length: 124456	epsilon: 0.9515751400010513
steps: 123	lr: 0.0001	evaluation reward: 1.53	
episode: 686	score: 0.0	memory length: 124579	epsilon: 0.9513316000010565
steps: 123	lr: 0.0001	evaluation reward: 1.52	
episode: 687	score: 1.0	memory length: 124748	epsilon: 0.9509969800010638
steps: 169	lr: 0.0001	evaluation reward: 1.51	
episode: 688	score: 3.0	memory length: 124977	epsilon: 0.9505435600010737
steps: 229	lr: 0.0001	evaluation reward: 1.54	
episode: 689	score: 1.0	memory length: 125146	epsilon: 0.9502089400010809
steps: 169	lr: 0.0001	evaluation reward: 1.53	
episode: 690	score: 5.0	memory length: 125438	epsilon: 0.9496307800010935
steps: 292	lr: 0.0001	evaluation reward: 1.56	
episode: 691	score: 4.0	memory length: 125724	epsilon: 0.9490645000011058
steps: 286	lr: 0.0001	evaluation reward: 1.58	
episode: 692	score: 1.0	memory length: 125893	epsilon: 0.9487298800011113
steps: 169	lr: 0.0001	evaluation reward: 1.58	
episode: 693	score: 2.0	memory length: 126091	epsilon: 0.9483378400011215
steps: 198	lr: 0.0001	evaluation reward: 1.55	
episode: 694	score: 1.0	memory length: 126241	epsilon: 0.948040840001128
steps: 150	lr: 0.0001	evaluation reward: 1.55	
episode: 695	score: 1.0	memory length: 126409	epsilon: 0.9477082000011352
steps: 168	lr: 0.0001	evaluation reward: 1.53	
episode: 696	score: 0.0	memory length: 126532	epsilon: 0.9474646600011405
steps: 123	lr: 0.0001	evaluation reward: 1.53	
episode: 697	score: 1.0	memory length: 126701	epsilon: 0.9471300400011478
steps: 169	lr: 0.0001	evaluation reward: 1.52	
episode: 698	score: 2.0	memory length: 126899	epsilon: 0.9467380000011563
steps: 198	lr: 0.0001	evaluation reward: 1.53	
episode: 699	score: 0.0	memory length: 127022	epsilon: 0.9464944600011616
steps: 123	lr: 0.0001	evaluation reward: 1.52	

episode: 700	score: 2.0	memory length: 127237	epsilon: 0.9460687600011708
steps: 215	lr: 0.0001	evaluation reward: 1.54	
episode: 701	score: 1.0	memory length: 127406	epsilon: 0.9457341400011781
steps: 169	lr: 0.0001	evaluation reward: 1.55	
episode: 702	score: 1.0	memory length: 127577	epsilon: 0.9453955600011854
steps: 171	lr: 0.0001	evaluation reward: 1.55	
episode: 703	score: 1.0	memory length: 127728	epsilon: 0.9450965800011919
steps: 151	lr: 0.0001	evaluation reward: 1.56	
episode: 704	score: 0.0	memory length: 127851	epsilon: 0.9448530400011972
steps: 123	lr: 0.0001	evaluation reward: 1.56	
episode: 705	score: 2.0	memory length: 128070	epsilon: 0.9444194200012066
steps: 219	lr: 0.0001	evaluation reward: 1.57	
episode: 706	score: 0.0	memory length: 128192	epsilon: 0.9441778600012118
steps: 122	lr: 0.0001	evaluation reward: 1.55	
episode: 707	score: 0.0	memory length: 128315	epsilon: 0.9439343200012171
steps: 123	lr: 0.0001	evaluation reward: 1.54	
episode: 708	score: 1.0	memory length: 128483	epsilon: 0.9436016800012244
steps: 168	lr: 0.0001	evaluation reward: 1.55	
episode: 709	score: 5.0	memory length: 128829	epsilon: 0.9429166000012392
steps: 346	lr: 0.0001	evaluation reward: 1.6	
episode: 710	score: 0.0	memory length: 128952	epsilon: 0.9426730600012445
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 711	score: 0.0	memory length: 129075	epsilon: 0.9424295200012498
steps: 123	lr: 0.0001	evaluation reward: 1.56	
episode: 712	score: 1.0	memory length: 129226	epsilon: 0.9421305400012563
steps: 151	lr: 0.0001	evaluation reward: 1.57	
episode: 713	score: 0.0	memory length: 129348	epsilon: 0.9418889800012615
steps: 122	lr: 0.0001	evaluation reward: 1.56	
episode: 714	score: 1.0	memory length: 129517	epsilon: 0.9415543600012688
steps: 169	lr: 0.0001	evaluation reward: 1.53	
episode: 715	score: 2.0	memory length: 129735	epsilon: 0.9411227200012782
steps: 218	lr: 0.0001	evaluation reward: 1.54	
episode: 716	score: 2.0	memory length: 129954	epsilon: 0.9406891000012876
steps: 219	lr: 0.0001	evaluation reward: 1.56	
episode: 717	score: 0.0	memory length: 130076	epsilon: 0.9404475400012928
steps: 122	lr: 0.0001	evaluation reward: 1.55	
episode: 718	score: 0.0	memory length: 130199	epsilon: 0.9402040000012981
steps: 123	lr: 0.0001	evaluation reward: 1.53	
episode: 719	score: 2.0	memory length: 130396	epsilon: 0.9398139400013066
steps: 197	lr: 0.0001	evaluation reward: 1.54	
episode: 720	score: 1.0	memory length: 130567	epsilon: 0.9394753600013139
steps: 171	lr: 0.0001	evaluation reward: 1.53	
episode: 721	score: 2.0	memory length: 130765	epsilon: 0.9390833200013224
steps: 198	lr: 0.0001	evaluation reward: 1.53	
episode: 722	score: 2.0	memory length: 130963	epsilon: 0.938691280001331
steps: 198	lr: 0.0001	evaluation reward: 1.55	
episode: 723	score: 4.0	memory length: 131258	epsilon: 0.9381071800013436
steps: 295	lr: 0.0001	evaluation reward: 1.56	

episode: 724	score: 5.0	memory length: 131601	epsilon: 0.9374280400013584
steps: 343	lr: 0.0001	evaluation reward: 1.61	
episode: 725	score: 2.0	memory length: 131803	epsilon: 0.9370280800013671
steps: 202	lr: 0.0001	evaluation reward: 1.58	
episode: 726	score: 3.0	memory length: 132051	epsilon: 0.9365370400013777
steps: 248	lr: 0.0001	evaluation reward: 1.55	
episode: 727	score: 3.0	memory length: 132277	epsilon: 0.9360895600013874
steps: 226	lr: 0.0001	evaluation reward: 1.57	
episode: 728	score: 2.0	memory length: 132475	epsilon: 0.9356975200013959
steps: 198	lr: 0.0001	evaluation reward: 1.58	
episode: 729	score: 5.0	memory length: 132785	epsilon: 0.9350837200014093
steps: 310	lr: 0.0001	evaluation reward: 1.63	
episode: 730	score: 0.0	memory length: 132907	epsilon: 0.9348421600014145
steps: 122	lr: 0.0001	evaluation reward: 1.62	
episode: 731	score: 0.0	memory length: 133030	epsilon: 0.9345986200014198
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 732	score: 1.0	memory length: 133200	epsilon: 0.9342620200014271
steps: 170	lr: 0.0001	evaluation reward: 1.58	
episode: 733	score: 0.0	memory length: 133322	epsilon: 0.9340204600014324
steps: 122	lr: 0.0001	evaluation reward: 1.56	
episode: 734	score: 1.0	memory length: 133475	epsilon: 0.9337175200014389
steps: 153	lr: 0.0001	evaluation reward: 1.56	
episode: 735	score: 2.0	memory length: 133673	epsilon: 0.9333254800014474
steps: 198	lr: 0.0001	evaluation reward: 1.58	
episode: 736	score: 3.0	memory length: 133922	epsilon: 0.9328324600014581
steps: 249	lr: 0.0001	evaluation reward: 1.6	
episode: 737	score: 0.0	memory length: 134045	epsilon: 0.9325889200014634
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 738	score: 2.0	memory length: 134225	epsilon: 0.9322325200014712
steps: 180	lr: 0.0001	evaluation reward: 1.6	
episode: 739	score: 5.0	memory length: 134570	epsilon: 0.931549420001486
steps: 345	lr: 0.0001	evaluation reward: 1.63	
episode: 740	score: 0.0	memory length: 134693	epsilon: 0.9313058800014913
steps: 123	lr: 0.0001	evaluation reward: 1.63	
episode: 741	score: 0.0	memory length: 134816	epsilon: 0.9310623400014966
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 742	score: 2.0	memory length: 135014	epsilon: 0.9306703000015051
steps: 198	lr: 0.0001	evaluation reward: 1.6	
episode: 743	score: 4.0	memory length: 135308	epsilon: 0.9300881800015177
steps: 294	lr: 0.0001	evaluation reward: 1.62	
episode: 744	score: 2.0	memory length: 135506	epsilon: 0.9296961400015262
steps: 198	lr: 0.0001	evaluation reward: 1.63	
episode: 745	score: 3.0	memory length: 135753	epsilon: 0.9292070800015368
steps: 247	lr: 0.0001	evaluation reward: 1.63	
episode: 746	score: 2.0	memory length: 135950	epsilon: 0.9288170200015453
steps: 197	lr: 0.0001	evaluation reward: 1.65	
episode: 747	score: 2.0	memory length: 136147	epsilon: 0.9284269600015538
steps: 197	lr: 0.0001	evaluation reward: 1.62	

episode: 748	score: 1.0	memory length: 136298	epsilon: 0.9281279800015603
steps: 151	lr: 0.0001	evaluation reward: 1.62	
episode: 749	score: 0.0	memory length: 136420	epsilon: 0.9278864200015655
steps: 122	lr: 0.0001	evaluation reward: 1.61	
episode: 750	score: 0.0	memory length: 136543	epsilon: 0.9276428800015708
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 751	score: 1.0	memory length: 136694	epsilon: 0.9273439000015773
steps: 151	lr: 0.0001	evaluation reward: 1.56	
episode: 752	score: 1.0	memory length: 136865	epsilon: 0.9270053200015846
steps: 171	lr: 0.0001	evaluation reward: 1.55	
episode: 753	score: 3.0	memory length: 137132	epsilon: 0.9264766600015961
steps: 267	lr: 0.0001	evaluation reward: 1.55	
episode: 754	score: 0.0	memory length: 137255	epsilon: 0.9262331200016014
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 755	score: 1.0	memory length: 137425	epsilon: 0.9258965200016087
steps: 170	lr: 0.0001	evaluation reward: 1.51	
episode: 756	score: 1.0	memory length: 137576	epsilon: 0.9255975400016152
steps: 151	lr: 0.0001	evaluation reward: 1.52	
episode: 757	score: 1.0	memory length: 137748	epsilon: 0.9252569800016226
steps: 172	lr: 0.0001	evaluation reward: 1.49	
episode: 758	score: 2.0	memory length: 137969	epsilon: 0.9248194000016321
steps: 221	lr: 0.0001	evaluation reward: 1.49	
episode: 759	score: 3.0	memory length: 138216	epsilon: 0.9243303400016427
steps: 247	lr: 0.0001	evaluation reward: 1.51	
episode: 760	score: 2.0	memory length: 138434	epsilon: 0.9238987000016521
steps: 218	lr: 0.0001	evaluation reward: 1.5	
episode: 761	score: 2.0	memory length: 138632	epsilon: 0.9235066600016606
steps: 198	lr: 0.0001	evaluation reward: 1.51	
episode: 762	score: 1.0	memory length: 138802	epsilon: 0.9231700600016679
steps: 170	lr: 0.0001	evaluation reward: 1.51	
episode: 763	score: 2.0	memory length: 138999	epsilon: 0.9227800000016764
steps: 197	lr: 0.0001	evaluation reward: 1.52	
episode: 764	score: 3.0	memory length: 139225	epsilon: 0.9223325200016861
steps: 226	lr: 0.0001	evaluation reward: 1.54	
episode: 765	score: 0.0	memory length: 139348	epsilon: 0.9220889800016914
steps: 123	lr: 0.0001	evaluation reward: 1.5	
episode: 766	score: 4.0	memory length: 139628	epsilon: 0.9215345800017034
steps: 280	lr: 0.0001	evaluation reward: 1.52	
episode: 767	score: 4.0	memory length: 139924	epsilon: 0.9209485000017161
steps: 296	lr: 0.0001	evaluation reward: 1.55	
episode: 768	score: 1.0	memory length: 140093	epsilon: 0.9206138800017234
steps: 169	lr: 0.0001	evaluation reward: 1.56	
episode: 769	score: 0.0	memory length: 140216	epsilon: 0.9203703400017287
steps: 123	lr: 0.0001	evaluation reward: 1.55	
episode: 770	score: 2.0	memory length: 140436	epsilon: 0.9199347400017381
steps: 220	lr: 0.0001	evaluation reward: 1.57	
episode: 771	score: 6.0	memory length: 140811	epsilon: 0.9191922400017543
steps: 375	lr: 0.0001	evaluation reward: 1.62	

episode: 772	score: 4.0	memory length: 141103	epsilon: 0.9186140800017668
steps: 292	lr: 0.0001	evaluation reward: 1.64	
episode: 773	score: 0.0	memory length: 141226	epsilon: 0.9183705400017721
steps: 123	lr: 0.0001	evaluation reward: 1.64	
episode: 774	score: 0.0	memory length: 141349	epsilon: 0.9181270000017774
steps: 123	lr: 0.0001	evaluation reward: 1.63	
episode: 775	score: 2.0	memory length: 141547	epsilon: 0.9177349600017859
steps: 198	lr: 0.0001	evaluation reward: 1.62	
episode: 776	score: 3.0	memory length: 141798	epsilon: 0.9172379800017967
steps: 251	lr: 0.0001	evaluation reward: 1.63	
episode: 777	score: 0.0	memory length: 141921	epsilon: 0.916994440001802
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 778	score: 2.0	memory length: 142119	epsilon: 0.9166024000018105
steps: 198	lr: 0.0001	evaluation reward: 1.62	
episode: 779	score: 2.0	memory length: 142316	epsilon: 0.916212340001819
steps: 197	lr: 0.0001	evaluation reward: 1.63	
episode: 780	score: 2.0	memory length: 142496	epsilon: 0.9158559400018267
steps: 180	lr: 0.0001	evaluation reward: 1.65	
episode: 781	score: 2.0	memory length: 142676	epsilon: 0.9154995400018344
steps: 180	lr: 0.0001	evaluation reward: 1.67	
episode: 782	score: 4.0	memory length: 142970	epsilon: 0.9149174200018471
steps: 294	lr: 0.0001	evaluation reward: 1.69	
episode: 783	score: 2.0	memory length: 143188	epsilon: 0.9144857800018564
steps: 218	lr: 0.0001	evaluation reward: 1.68	
episode: 784	score: 3.0	memory length: 143452	epsilon: 0.9139630600018678
steps: 264	lr: 0.0001	evaluation reward: 1.68	
episode: 785	score: 1.0	memory length: 143620	epsilon: 0.913630420001875
steps: 168	lr: 0.0001	evaluation reward: 1.69	
episode: 786	score: 1.0	memory length: 143791	epsilon: 0.9132918400018823
steps: 171	lr: 0.0001	evaluation reward: 1.7	
episode: 787	score: 4.0	memory length: 144110	epsilon: 0.9126602200018961
steps: 319	lr: 0.0001	evaluation reward: 1.73	
episode: 788	score: 4.0	memory length: 144385	epsilon: 0.9121157200019079
steps: 275	lr: 0.0001	evaluation reward: 1.74	
episode: 789	score: 1.0	memory length: 144536	epsilon: 0.9118167400019144
steps: 151	lr: 0.0001	evaluation reward: 1.74	
episode: 790	score: 3.0	memory length: 144781	epsilon: 0.9113316400019249
steps: 245	lr: 0.0001	evaluation reward: 1.72	
episode: 791	score: 3.0	memory length: 145025	epsilon: 0.9108485200019354
steps: 244	lr: 0.0001	evaluation reward: 1.71	
episode: 792	score: 4.0	memory length: 145300	epsilon: 0.9103040200019472
steps: 275	lr: 0.0001	evaluation reward: 1.74	
episode: 793	score: 4.0	memory length: 145597	epsilon: 0.90971596000196
steps: 297	lr: 0.0001	evaluation reward: 1.76	
episode: 794	score: 4.0	memory length: 145891	epsilon: 0.9091338400019726
steps: 294	lr: 0.0001	evaluation reward: 1.79	
episode: 795	score: 3.0	memory length: 146139	epsilon: 0.9086428000019833
steps: 248	lr: 0.0001	evaluation reward: 1.81	

episode: 796	score: 2.0	memory length: 146339	epsilon: 0.9082468000019919
steps: 200	lr: 0.0001	evaluation reward: 1.83	
episode: 797	score: 1.0	memory length: 146509	epsilon: 0.9079102000019992
steps: 170	lr: 0.0001	evaluation reward: 1.83	
episode: 798	score: 0.0	memory length: 146631	epsilon: 0.9076686400020044
steps: 122	lr: 0.0001	evaluation reward: 1.81	
episode: 799	score: 3.0	memory length: 146841	epsilon: 0.9072528400020134
steps: 210	lr: 0.0001	evaluation reward: 1.84	
episode: 800	score: 2.0	memory length: 147038	epsilon: 0.9068627800020219
steps: 197	lr: 0.0001	evaluation reward: 1.84	
episode: 801	score: 1.0	memory length: 147188	epsilon: 0.9065657800020284
steps: 150	lr: 0.0001	evaluation reward: 1.84	
episode: 802	score: 2.0	memory length: 147388	epsilon: 0.906169780002037
steps: 200	lr: 0.0001	evaluation reward: 1.85	
episode: 803	score: 2.0	memory length: 147606	epsilon: 0.9057381400020463
steps: 218	lr: 0.0001	evaluation reward: 1.86	
episode: 804	score: 4.0	memory length: 147864	epsilon: 0.9052273000020574
steps: 258	lr: 0.0001	evaluation reward: 1.9	
episode: 805	score: 0.0	memory length: 147987	epsilon: 0.9049837600020627
steps: 123	lr: 0.0001	evaluation reward: 1.88	
episode: 806	score: 4.0	memory length: 148279	epsilon: 0.9044056000020753
steps: 292	lr: 0.0001	evaluation reward: 1.92	
episode: 807	score: 0.0	memory length: 148402	epsilon: 0.9041620600020805
steps: 123	lr: 0.0001	evaluation reward: 1.92	
episode: 808	score: 1.0	memory length: 148570	epsilon: 0.9038294200020878
steps: 168	lr: 0.0001	evaluation reward: 1.92	
episode: 809	score: 1.0	memory length: 148739	epsilon: 0.903494800002095
steps: 169	lr: 0.0001	evaluation reward: 1.88	
episode: 810	score: 1.0	memory length: 148910	epsilon: 0.9031562200021024
steps: 171	lr: 0.0001	evaluation reward: 1.89	
episode: 811	score: 0.0	memory length: 149033	epsilon: 0.9029126800021077
steps: 123	lr: 0.0001	evaluation reward: 1.89	
episode: 812	score: 3.0	memory length: 149300	epsilon: 0.9023840200021191
steps: 267	lr: 0.0001	evaluation reward: 1.91	
episode: 813	score: 0.0	memory length: 149423	epsilon: 0.9021404800021244
steps: 123	lr: 0.0001	evaluation reward: 1.91	
episode: 814	score: 2.0	memory length: 149623	epsilon: 0.901744480002133
steps: 200	lr: 0.0001	evaluation reward: 1.92	
episode: 815	score: 0.0	memory length: 149746	epsilon: 0.9015009400021383
steps: 123	lr: 0.0001	evaluation reward: 1.9	
episode: 816	score: 3.0	memory length: 149992	epsilon: 0.9010138600021489
steps: 246	lr: 0.0001	evaluation reward: 1.91	
episode: 817	score: 1.0	memory length: 150161	epsilon: 0.9006792400021562
steps: 169	lr: 0.0001	evaluation reward: 1.92	
episode: 818	score: 2.0	memory length: 150381	epsilon: 0.9002436400021656
steps: 220	lr: 0.0001	evaluation reward: 1.94	
episode: 819	score: 1.0	memory length: 150549	epsilon: 0.8999110000021728
steps: 168	lr: 0.0001	evaluation reward: 1.93	

episode: 820	score: 3.0	memory length: 150801	epsilon: 0.8994120400021837
steps: 252	lr: 0.0001	evaluation reward: 1.95	
episode: 821	score: 1.0	memory length: 150970	epsilon: 0.8990774200021909
steps: 169	lr: 0.0001	evaluation reward: 1.94	
episode: 822	score: 1.0	memory length: 151142	epsilon: 0.8987368600021983
steps: 172	lr: 0.0001	evaluation reward: 1.93	
episode: 823	score: 2.0	memory length: 151360	epsilon: 0.8983052200022077
steps: 218	lr: 0.0001	evaluation reward: 1.91	
episode: 824	score: 3.0	memory length: 151588	epsilon: 0.8978537800022175
steps: 228	lr: 0.0001	evaluation reward: 1.89	
episode: 825	score: 2.0	memory length: 151785	epsilon: 0.897463720002226
steps: 197	lr: 0.0001	evaluation reward: 1.89	
episode: 826	score: 2.0	memory length: 151983	epsilon: 0.8970716800022345
steps: 198	lr: 0.0001	evaluation reward: 1.88	
episode: 827	score: 1.0	memory length: 152152	epsilon: 0.8967370600022417
steps: 169	lr: 0.0001	evaluation reward: 1.86	
episode: 828	score: 0.0	memory length: 152274	epsilon: 0.896495500002247
steps: 122	lr: 0.0001	evaluation reward: 1.84	
episode: 829	score: 0.0	memory length: 152396	epsilon: 0.8962539400022522
steps: 122	lr: 0.0001	evaluation reward: 1.79	
episode: 830	score: 2.0	memory length: 152593	epsilon: 0.8958638800022607
steps: 197	lr: 0.0001	evaluation reward: 1.81	
episode: 831	score: 2.0	memory length: 152794	epsilon: 0.8954659000022693
steps: 201	lr: 0.0001	evaluation reward: 1.83	
episode: 832	score: 2.0	memory length: 152991	epsilon: 0.8950758400022778
steps: 197	lr: 0.0001	evaluation reward: 1.84	
episode: 833	score: 0.0	memory length: 153114	epsilon: 0.8948323000022831
steps: 123	lr: 0.0001	evaluation reward: 1.84	
episode: 834	score: 3.0	memory length: 153357	epsilon: 0.8943511600022935
steps: 243	lr: 0.0001	evaluation reward: 1.86	
episode: 835	score: 2.0	memory length: 153558	epsilon: 0.8939531800023022
steps: 201	lr: 0.0001	evaluation reward: 1.86	
episode: 836	score: 4.0	memory length: 153837	epsilon: 0.8934007600023142
steps: 279	lr: 0.0001	evaluation reward: 1.87	
episode: 837	score: 3.0	memory length: 154064	epsilon: 0.8929513000023239
steps: 227	lr: 0.0001	evaluation reward: 1.9	
episode: 838	score: 0.0	memory length: 154187	epsilon: 0.8927077600023292
steps: 123	lr: 0.0001	evaluation reward: 1.88	
episode: 839	score: 0.0	memory length: 154309	epsilon: 0.8924662000023345
steps: 122	lr: 0.0001	evaluation reward: 1.83	
episode: 840	score: 1.0	memory length: 154481	epsilon: 0.8921256400023418
steps: 172	lr: 0.0001	evaluation reward: 1.84	
episode: 841	score: 2.0	memory length: 154700	epsilon: 0.8916920200023513
steps: 219	lr: 0.0001	evaluation reward: 1.86	
episode: 842	score: 5.0	memory length: 155006	epsilon: 0.8910861400023644
steps: 306	lr: 0.0001	evaluation reward: 1.89	
episode: 843	score: 1.0	memory length: 155156	epsilon: 0.8907891400023709
steps: 150	lr: 0.0001	evaluation reward: 1.86	

episode: 844	score: 5.0	memory length: 155479	epsilon: 0.8901496000023847
steps: 323	lr: 0.0001	evaluation reward: 1.89	
episode: 845	score: 1.0	memory length: 155630	epsilon: 0.8898506200023912
steps: 151	lr: 0.0001	evaluation reward: 1.87	
episode: 846	score: 0.0	memory length: 155753	epsilon: 0.8896070800023965
steps: 123	lr: 0.0001	evaluation reward: 1.85	
episode: 847	score: 3.0	memory length: 155999	epsilon: 0.8891200000024071
steps: 246	lr: 0.0001	evaluation reward: 1.86	
episode: 848	score: 0.0	memory length: 156122	epsilon: 0.8888764600024124
steps: 123	lr: 0.0001	evaluation reward: 1.85	
episode: 849	score: 4.0	memory length: 156419	epsilon: 0.8882884000024251
steps: 297	lr: 0.0001	evaluation reward: 1.89	
episode: 850	score: 3.0	memory length: 156686	epsilon: 0.8877597400024366
steps: 267	lr: 0.0001	evaluation reward: 1.92	
episode: 851	score: 0.0	memory length: 156809	epsilon: 0.8875162000024419
steps: 123	lr: 0.0001	evaluation reward: 1.91	
episode: 852	score: 2.0	memory length: 157007	epsilon: 0.8871241600024504
steps: 198	lr: 0.0001	evaluation reward: 1.92	
episode: 853	score: 4.0	memory length: 157300	epsilon: 0.886544020002463
steps: 293	lr: 0.0001	evaluation reward: 1.93	
episode: 854	score: 3.0	memory length: 157548	epsilon: 0.8860529800024737
steps: 248	lr: 0.0001	evaluation reward: 1.96	
episode: 855	score: 2.0	memory length: 157748	epsilon: 0.8856569800024823
steps: 200	lr: 0.0001	evaluation reward: 1.97	
episode: 856	score: 3.0	memory length: 157994	epsilon: 0.8851699000024928
steps: 246	lr: 0.0001	evaluation reward: 1.99	
episode: 857	score: 0.0	memory length: 158116	epsilon: 0.8849283400024981
steps: 122	lr: 0.0001	evaluation reward: 1.98	
episode: 858	score: 3.0	memory length: 158383	epsilon: 0.8843996800025096
steps: 267	lr: 0.0001	evaluation reward: 1.99	
episode: 859	score: 3.0	memory length: 158632	epsilon: 0.8839066600025203
steps: 249	lr: 0.0001	evaluation reward: 1.99	
episode: 860	score: 1.0	memory length: 158801	epsilon: 0.8835720400025275
steps: 169	lr: 0.0001	evaluation reward: 1.98	
episode: 861	score: 4.0	memory length: 159097	epsilon: 0.8829859600025403
steps: 296	lr: 0.0001	evaluation reward: 2.0	
episode: 862	score: 3.0	memory length: 159342	epsilon: 0.8825008600025508
steps: 245	lr: 0.0001	evaluation reward: 2.02	
episode: 863	score: 3.0	memory length: 159590	epsilon: 0.8820098200025615
steps: 248	lr: 0.0001	evaluation reward: 2.03	
episode: 864	score: 3.0	memory length: 159836	epsilon: 0.881522740002572
steps: 246	lr: 0.0001	evaluation reward: 2.03	
episode: 865	score: 1.0	memory length: 160005	epsilon: 0.8811881200025793
steps: 169	lr: 0.0001	evaluation reward: 2.04	
episode: 866	score: 2.0	memory length: 160184	epsilon: 0.880833700002587
steps: 179	lr: 0.0001	evaluation reward: 2.02	
episode: 867	score: 2.0	memory length: 160400	epsilon: 0.8804060200025963
steps: 216	lr: 0.0001	evaluation reward: 2.0	

episode: 868	score: 1.0	memory length: 160571	epsilon: 0.8800674400026036
steps: 171	lr: 0.0001	evaluation reward: 2.0	
episode: 869	score: 6.0	memory length: 160943	epsilon: 0.8793308800026196
steps: 372	lr: 0.0001	evaluation reward: 2.06	
episode: 870	score: 2.0	memory length: 161141	epsilon: 0.8789388400026281
steps: 198	lr: 0.0001	evaluation reward: 2.06	
episode: 871	score: 2.0	memory length: 161322	epsilon: 0.8785804600026359
steps: 181	lr: 0.0001	evaluation reward: 2.02	
episode: 872	score: 2.0	memory length: 161504	epsilon: 0.8782201000026437
steps: 182	lr: 0.0001	evaluation reward: 2.0	
episode: 873	score: 1.0	memory length: 161655	epsilon: 0.8779211200026502
steps: 151	lr: 0.0001	evaluation reward: 2.01	
episode: 874	score: 1.0	memory length: 161806	epsilon: 0.8776221400026567
steps: 151	lr: 0.0001	evaluation reward: 2.02	
episode: 875	score: 0.0	memory length: 161929	epsilon: 0.877378600002662
steps: 123	lr: 0.0001	evaluation reward: 2.0	
episode: 876	score: 2.0	memory length: 162127	epsilon: 0.8769865600026705
steps: 198	lr: 0.0001	evaluation reward: 1.99	
episode: 877	score: 3.0	memory length: 162372	epsilon: 0.876501460002681
steps: 245	lr: 0.0001	evaluation reward: 2.02	
episode: 878	score: 3.0	memory length: 162620	epsilon: 0.8760104200026917
steps: 248	lr: 0.0001	evaluation reward: 2.03	
episode: 879	score: 3.0	memory length: 162867	epsilon: 0.8755213600027023
steps: 247	lr: 0.0001	evaluation reward: 2.04	
episode: 880	score: 4.0	memory length: 163183	epsilon: 0.8748956800027159
steps: 316	lr: 0.0001	evaluation reward: 2.06	
episode: 881	score: 2.0	memory length: 163380	epsilon: 0.8745056200027244
steps: 197	lr: 0.0001	evaluation reward: 2.06	
episode: 882	score: 3.0	memory length: 163605	epsilon: 0.874060120002734
steps: 225	lr: 0.0001	evaluation reward: 2.05	
episode: 883	score: 1.0	memory length: 163756	epsilon: 0.8737611400027405
steps: 151	lr: 0.0001	evaluation reward: 2.04	
episode: 884	score: 3.0	memory length: 164005	epsilon: 0.8732681200027512
steps: 249	lr: 0.0001	evaluation reward: 2.04	
episode: 885	score: 3.0	memory length: 164274	epsilon: 0.8727355000027628
steps: 269	lr: 0.0001	evaluation reward: 2.06	
episode: 886	score: 1.0	memory length: 164425	epsilon: 0.8724365200027693
steps: 151	lr: 0.0001	evaluation reward: 2.06	
episode: 887	score: 3.0	memory length: 164651	epsilon: 0.871989040002779
steps: 226	lr: 0.0001	evaluation reward: 2.05	
episode: 888	score: 2.0	memory length: 164849	epsilon: 0.8715970000027875
steps: 198	lr: 0.0001	evaluation reward: 2.03	
episode: 889	score: 3.0	memory length: 165094	epsilon: 0.871111900002798
steps: 245	lr: 0.0001	evaluation reward: 2.05	
episode: 890	score: 0.0	memory length: 165217	epsilon: 0.8708683600028033
steps: 123	lr: 0.0001	evaluation reward: 2.02	
episode: 891	score: 3.0	memory length: 165460	epsilon: 0.8703872200028138
steps: 243	lr: 0.0001	evaluation reward: 2.02	

episode: 892	score: 1.0	memory length: 165611	epsilon: 0.8700882400028203
steps: 151	lr: 0.0001	evaluation reward: 1.99	
episode: 893	score: 1.0	memory length: 165779	epsilon: 0.8697556000028275
steps: 168	lr: 0.0001	evaluation reward: 1.96	
episode: 894	score: 1.0	memory length: 165930	epsilon: 0.869456620002834
steps: 151	lr: 0.0001	evaluation reward: 1.93	
episode: 895	score: 4.0	memory length: 166223	epsilon: 0.8688764800028466
steps: 293	lr: 0.0001	evaluation reward: 1.94	
episode: 896	score: 4.0	memory length: 166514	epsilon: 0.8683003000028591
steps: 291	lr: 0.0001	evaluation reward: 1.96	
episode: 897	score: 3.0	memory length: 166743	epsilon: 0.8678468800028689
steps: 229	lr: 0.0001	evaluation reward: 1.98	
episode: 898	score: 2.0	memory length: 166960	epsilon: 0.8674172200028782
steps: 217	lr: 0.0001	evaluation reward: 2.0	
episode: 899	score: 3.0	memory length: 167205	epsilon: 0.8669321200028888
steps: 245	lr: 0.0001	evaluation reward: 2.0	
episode: 900	score: 3.0	memory length: 167451	epsilon: 0.8664450400028993
steps: 246	lr: 0.0001	evaluation reward: 2.01	
episode: 901	score: 4.0	memory length: 167746	epsilon: 0.865860940002912
steps: 295	lr: 0.0001	evaluation reward: 2.04	
episode: 902	score: 1.0	memory length: 167918	epsilon: 0.8655203800029194
steps: 172	lr: 0.0001	evaluation reward: 2.03	
episode: 903	score: 3.0	memory length: 168144	epsilon: 0.8650729000029291
steps: 226	lr: 0.0001	evaluation reward: 2.04	
episode: 904	score: 1.0	memory length: 168295	epsilon: 0.8647739200029356
steps: 151	lr: 0.0001	evaluation reward: 2.01	
episode: 905	score: 0.0	memory length: 168417	epsilon: 0.8645323600029409
steps: 122	lr: 0.0001	evaluation reward: 2.01	
episode: 906	score: 2.0	memory length: 168634	epsilon: 0.8641027000029502
steps: 217	lr: 0.0001	evaluation reward: 1.99	
episode: 907	score: 4.0	memory length: 168953	epsilon: 0.8634710800029639
steps: 319	lr: 0.0001	evaluation reward: 2.03	
episode: 908	score: 2.0	memory length: 169135	epsilon: 0.8631107200029717
steps: 182	lr: 0.0001	evaluation reward: 2.04	
episode: 909	score: 2.0	memory length: 169314	epsilon: 0.8627563000029794
steps: 179	lr: 0.0001	evaluation reward: 2.05	
episode: 910	score: 2.0	memory length: 169511	epsilon: 0.8623662400029879
steps: 197	lr: 0.0001	evaluation reward: 2.06	
episode: 911	score: 1.0	memory length: 169680	epsilon: 0.8620316200029952
steps: 169	lr: 0.0001	evaluation reward: 2.07	
episode: 912	score: 2.0	memory length: 169882	epsilon: 0.8616316600030038
steps: 202	lr: 0.0001	evaluation reward: 2.06	
episode: 913	score: 3.0	memory length: 170107	epsilon: 0.8611861600030135
steps: 225	lr: 0.0001	evaluation reward: 2.09	
episode: 914	score: 0.0	memory length: 170230	epsilon: 0.8609426200030188
steps: 123	lr: 0.0001	evaluation reward: 2.07	
episode: 915	score: 8.0	memory length: 170545	epsilon: 0.8603189200030323
steps: 315	lr: 0.0001	evaluation reward: 2.15	

episode: 916	score: 2.0	memory length: 170743	epsilon: 0.8599268800030408
steps: 198	lr: 0.0001	evaluation reward: 2.14	
episode: 917	score: 2.0	memory length: 170962	epsilon: 0.8594932600030503
steps: 219	lr: 0.0001	evaluation reward: 2.15	
episode: 918	score: 3.0	memory length: 171229	epsilon: 0.8589646000030617
steps: 267	lr: 0.0001	evaluation reward: 2.16	
episode: 919	score: 1.0	memory length: 171399	epsilon: 0.858628000003069
steps: 170	lr: 0.0001	evaluation reward: 2.16	
episode: 920	score: 1.0	memory length: 171550	epsilon: 0.8583290200030755
steps: 151	lr: 0.0001	evaluation reward: 2.14	
episode: 921	score: 1.0	memory length: 171701	epsilon: 0.858030040003082
steps: 151	lr: 0.0001	evaluation reward: 2.14	
episode: 922	score: 0.0	memory length: 171824	epsilon: 0.8577865000030873
steps: 123	lr: 0.0001	evaluation reward: 2.13	
episode: 923	score: 2.0	memory length: 172006	epsilon: 0.8574261400030951
steps: 182	lr: 0.0001	evaluation reward: 2.13	
episode: 924	score: 3.0	memory length: 172271	epsilon: 0.8569014400031065
steps: 265	lr: 0.0001	evaluation reward: 2.13	
episode: 925	score: 4.0	memory length: 172587	epsilon: 0.8562757600031201
steps: 316	lr: 0.0001	evaluation reward: 2.15	
episode: 926	score: 4.0	memory length: 172884	epsilon: 0.8556877000031329
steps: 297	lr: 0.0001	evaluation reward: 2.17	
episode: 927	score: 4.0	memory length: 173141	epsilon: 0.8551788400031439
steps: 257	lr: 0.0001	evaluation reward: 2.2	
episode: 928	score: 5.0	memory length: 173485	epsilon: 0.8544977200031587
steps: 344	lr: 0.0001	evaluation reward: 2.25	
episode: 929	score: 2.0	memory length: 173703	epsilon: 0.8540660800031681
steps: 218	lr: 0.0001	evaluation reward: 2.27	
episode: 930	score: 4.0	memory length: 174000	epsilon: 0.8534780200031808
steps: 297	lr: 0.0001	evaluation reward: 2.29	
episode: 931	score: 1.0	memory length: 174151	epsilon: 0.8531790400031873
steps: 151	lr: 0.0001	evaluation reward: 2.28	
episode: 932	score: 1.0	memory length: 174301	epsilon: 0.8528820400031938
steps: 150	lr: 0.0001	evaluation reward: 2.27	
episode: 933	score: 1.0	memory length: 174473	epsilon: 0.8525414800032012
steps: 172	lr: 0.0001	evaluation reward: 2.28	
episode: 934	score: 5.0	memory length: 174794	epsilon: 0.851905900003215
steps: 321	lr: 0.0001	evaluation reward: 2.3	
episode: 935	score: 3.0	memory length: 175041	epsilon: 0.8514168400032256
steps: 247	lr: 0.0001	evaluation reward: 2.31	
episode: 936	score: 2.0	memory length: 175259	epsilon: 0.850985200003235
steps: 218	lr: 0.0001	evaluation reward: 2.29	
episode: 937	score: 0.0	memory length: 175382	epsilon: 0.8507416600032403
steps: 123	lr: 0.0001	evaluation reward: 2.26	
episode: 938	score: 2.0	memory length: 175580	epsilon: 0.8503496200032488
steps: 198	lr: 0.0001	evaluation reward: 2.28	
episode: 939	score: 0.0	memory length: 175703	epsilon: 0.850106080003254
steps: 123	lr: 0.0001	evaluation reward: 2.28	

episode: 940	score: 1.0	memory length: 175873	epsilon: 0.8497694800032614
steps: 170	lr: 0.0001	evaluation reward: 2.28	
episode: 941	score: 6.0	memory length: 176245	epsilon: 0.8490329200032773
steps: 372	lr: 0.0001	evaluation reward: 2.32	
episode: 942	score: 2.0	memory length: 176462	epsilon: 0.8486032600032867
steps: 217	lr: 0.0001	evaluation reward: 2.29	
episode: 943	score: 2.0	memory length: 176660	epsilon: 0.8482112200032952
steps: 198	lr: 0.0001	evaluation reward: 2.3	
episode: 944	score: 4.0	memory length: 176955	epsilon: 0.8476271200033079
steps: 295	lr: 0.0001	evaluation reward: 2.29	
episode: 945	score: 3.0	memory length: 177184	epsilon: 0.8471737000033177
steps: 229	lr: 0.0001	evaluation reward: 2.31	
episode: 946	score: 3.0	memory length: 177396	epsilon: 0.8467539400033268
steps: 212	lr: 0.0001	evaluation reward: 2.34	
episode: 947	score: 2.0	memory length: 177594	epsilon: 0.8463619000033353
steps: 198	lr: 0.0001	evaluation reward: 2.33	
episode: 948	score: 2.0	memory length: 177792	epsilon: 0.8459698600033438
steps: 198	lr: 0.0001	evaluation reward: 2.35	
episode: 949	score: 3.0	memory length: 178058	epsilon: 0.8454431800033553
steps: 266	lr: 0.0001	evaluation reward: 2.34	
episode: 950	score: 3.0	memory length: 178284	epsilon: 0.844995700003365
steps: 226	lr: 0.0001	evaluation reward: 2.34	
episode: 951	score: 4.0	memory length: 178595	epsilon: 0.8443799200033784
steps: 311	lr: 0.0001	evaluation reward: 2.38	
episode: 952	score: 5.0	memory length: 178939	epsilon: 0.8436988000033931
steps: 344	lr: 0.0001	evaluation reward: 2.41	
episode: 953	score: 2.0	memory length: 179136	epsilon: 0.8433087400034016
steps: 197	lr: 0.0001	evaluation reward: 2.39	
episode: 954	score: 4.0	memory length: 179410	epsilon: 0.8427662200034134
steps: 274	lr: 0.0001	evaluation reward: 2.4	
episode: 955	score: 3.0	memory length: 179657	epsilon: 0.842277160003424
steps: 247	lr: 0.0001	evaluation reward: 2.41	
episode: 956	score: 5.0	memory length: 179960	epsilon: 0.841677220003437
steps: 303	lr: 0.0001	evaluation reward: 2.43	
episode: 957	score: 4.0	memory length: 180217	epsilon: 0.8411683600034481
steps: 257	lr: 0.0001	evaluation reward: 2.47	
episode: 958	score: 0.0	memory length: 180340	epsilon: 0.8409248200034534
steps: 123	lr: 0.0001	evaluation reward: 2.44	
episode: 959	score: 4.0	memory length: 180635	epsilon: 0.840340720003466
steps: 295	lr: 0.0001	evaluation reward: 2.45	
episode: 960	score: 2.0	memory length: 180832	epsilon: 0.8399506600034745
steps: 197	lr: 0.0001	evaluation reward: 2.46	
episode: 961	score: 2.0	memory length: 181012	epsilon: 0.8395942600034823
steps: 180	lr: 0.0001	evaluation reward: 2.44	
episode: 962	score: 1.0	memory length: 181181	epsilon: 0.8392596400034895
steps: 169	lr: 0.0001	evaluation reward: 2.42	
episode: 963	score: 2.0	memory length: 181397	epsilon: 0.8388319600034988
steps: 216	lr: 0.0001	evaluation reward: 2.41	

episode: 964	score: 1.0	memory length: 181549	epsilon: 0.8385310000035053
steps: 152	lr: 0.0001	evaluation reward: 2.39	
episode: 965	score: 2.0	memory length: 181750	epsilon: 0.838133020003514
steps: 201	lr: 0.0001	evaluation reward: 2.4	
episode: 966	score: 5.0	memory length: 182054	epsilon: 0.837531100003527
steps: 304	lr: 0.0001	evaluation reward: 2.43	
episode: 967	score: 0.0	memory length: 182176	epsilon: 0.8372895400035323
steps: 122	lr: 0.0001	evaluation reward: 2.41	
episode: 968	score: 3.0	memory length: 182440	epsilon: 0.8367668200035436
steps: 264	lr: 0.0001	evaluation reward: 2.43	
episode: 969	score: 3.0	memory length: 182666	epsilon: 0.8363193400035533
steps: 226	lr: 0.0001	evaluation reward: 2.4	
episode: 970	score: 0.0	memory length: 182788	epsilon: 0.8360777800035586
steps: 122	lr: 0.0001	evaluation reward: 2.38	
episode: 971	score: 3.0	memory length: 183018	epsilon: 0.8356223800035685
steps: 230	lr: 0.0001	evaluation reward: 2.39	
episode: 972	score: 1.0	memory length: 183186	epsilon: 0.8352897400035757
steps: 168	lr: 0.0001	evaluation reward: 2.38	
episode: 973	score: 3.0	memory length: 183413	epsilon: 0.8348402800035855
steps: 227	lr: 0.0001	evaluation reward: 2.4	
episode: 974	score: 6.0	memory length: 183789	epsilon: 0.8340958000036016
steps: 376	lr: 0.0001	evaluation reward: 2.45	
episode: 975	score: 3.0	memory length: 184015	epsilon: 0.8336483200036113
steps: 226	lr: 0.0001	evaluation reward: 2.48	
episode: 976	score: 3.0	memory length: 184263	epsilon: 0.833157280003622
steps: 248	lr: 0.0001	evaluation reward: 2.49	
episode: 977	score: 2.0	memory length: 184482	epsilon: 0.8327236600036314
steps: 219	lr: 0.0001	evaluation reward: 2.48	
episode: 978	score: 1.0	memory length: 184633	epsilon: 0.8324246800036379
steps: 151	lr: 0.0001	evaluation reward: 2.46	
episode: 979	score: 0.0	memory length: 184756	epsilon: 0.8321811400036432
steps: 123	lr: 0.0001	evaluation reward: 2.43	
episode: 980	score: 3.0	memory length: 185005	epsilon: 0.8316881200036539
steps: 249	lr: 0.0001	evaluation reward: 2.42	
episode: 981	score: 2.0	memory length: 185185	epsilon: 0.8313317200036616
steps: 180	lr: 0.0001	evaluation reward: 2.42	
episode: 982	score: 3.0	memory length: 185431	epsilon: 0.8308446400036722
steps: 246	lr: 0.0001	evaluation reward: 2.42	
episode: 983	score: 0.0	memory length: 185554	epsilon: 0.8306011000036775
steps: 123	lr: 0.0001	evaluation reward: 2.41	
episode: 984	score: 2.0	memory length: 185752	epsilon: 0.830209060003686
steps: 198	lr: 0.0001	evaluation reward: 2.4	
episode: 985	score: 3.0	memory length: 185977	epsilon: 0.8297635600036957
steps: 225	lr: 0.0001	evaluation reward: 2.4	
episode: 986	score: 3.0	memory length: 186221	epsilon: 0.8292804400037062
steps: 244	lr: 0.0001	evaluation reward: 2.42	
episode: 987	score: 2.0	memory length: 186421	epsilon: 0.8288844400037148
steps: 200	lr: 0.0001	evaluation reward: 2.41	

episode: 988	score: 3.0	memory length: 186685	epsilon: 0.8283617200037261
steps: 264	lr: 0.0001	evaluation reward: 2.42	
episode: 989	score: 2.0	memory length: 186903	epsilon: 0.8279300800037355
steps: 218	lr: 0.0001	evaluation reward: 2.41	
episode: 990	score: 1.0	memory length: 187054	epsilon: 0.827631100003742
steps: 151	lr: 0.0001	evaluation reward: 2.42	
episode: 991	score: 1.0	memory length: 187205	epsilon: 0.8273321200037485
steps: 151	lr: 0.0001	evaluation reward: 2.4	
episode: 992	score: 2.0	memory length: 187425	epsilon: 0.8268965200037579
steps: 220	lr: 0.0001	evaluation reward: 2.41	
episode: 993	score: 1.0	memory length: 187576	epsilon: 0.8265975400037644
steps: 151	lr: 0.0001	evaluation reward: 2.41	
episode: 994	score: 6.0	memory length: 187933	epsilon: 0.8258906800037797
steps: 357	lr: 0.0001	evaluation reward: 2.46	
episode: 995	score: 2.0	memory length: 188133	epsilon: 0.8254946800037883
steps: 200	lr: 0.0001	evaluation reward: 2.44	
episode: 996	score: 6.0	memory length: 188469	epsilon: 0.8248294000038028
steps: 336	lr: 0.0001	evaluation reward: 2.46	
episode: 997	score: 1.0	memory length: 188620	epsilon: 0.8245304200038093
steps: 151	lr: 0.0001	evaluation reward: 2.44	
episode: 998	score: 2.0	memory length: 188820	epsilon: 0.8241344200038179
steps: 200	lr: 0.0001	evaluation reward: 2.44	
episode: 999	score: 4.0	memory length: 189081	epsilon: 0.8236176400038291
steps: 261	lr: 0.0001	evaluation reward: 2.45	
episode: 1000	score: 6.0	memory length: 189474	epsilon: 0.822839500003846
steps: 393	lr: 0.0001	evaluation reward: 2.48	
episode: 1001	score: 3.0	memory length: 189700	epsilon: 0.8223920200038557
steps: 226	lr: 0.0001	evaluation reward: 2.47	
episode: 1002	score: 0.0	memory length: 189822	epsilon: 0.8221504600038609
steps: 122	lr: 0.0001	evaluation reward: 2.46	
episode: 1003	score: 4.0	memory length: 190076	epsilon: 0.8216475400038719
steps: 254	lr: 0.0001	evaluation reward: 2.47	
episode: 1004	score: 1.0	memory length: 190246	epsilon: 0.8213109400038792
steps: 170	lr: 0.0001	evaluation reward: 2.47	
episode: 1005	score: 3.0	memory length: 190516	epsilon: 0.8207763400038908
steps: 270	lr: 0.0001	evaluation reward: 2.5	
episode: 1006	score: 2.0	memory length: 190698	epsilon: 0.8204159800038986
steps: 182	lr: 0.0001	evaluation reward: 2.5	
episode: 1007	score: 6.0	memory length: 191024	epsilon: 0.8197705000039126
steps: 326	lr: 0.0001	evaluation reward: 2.52	
episode: 1008	score: 3.0	memory length: 191289	epsilon: 0.819245800003924
steps: 265	lr: 0.0001	evaluation reward: 2.53	
episode: 1009	score: 6.0	memory length: 191622	epsilon: 0.8185864600039383
steps: 333	lr: 0.0001	evaluation reward: 2.57	
episode: 1010	score: 4.0	memory length: 191922	epsilon: 0.8179924600039512
steps: 300	lr: 0.0001	evaluation reward: 2.59	
episode: 1011	score: 2.0	memory length: 192120	epsilon: 0.8176004200039597
steps: 198	lr: 0.0001	evaluation reward: 2.6	

episode: 1012	score: 1.0	memory length: 192288	epsilon: 0.8172677800039669
steps: 168	lr: 0.0001	evaluation reward: 2.59	
episode: 1013	score: 3.0	memory length: 192514	epsilon: 0.8168203000039767
steps: 226	lr: 0.0001	evaluation reward: 2.59	
episode: 1014	score: 4.0	memory length: 192773	epsilon: 0.8163074800039878
steps: 259	lr: 0.0001	evaluation reward: 2.63	
episode: 1015	score: 3.0	memory length: 193023	epsilon: 0.8158124800039985
steps: 250	lr: 0.0001	evaluation reward: 2.58	
episode: 1016	score: 3.0	memory length: 193269	epsilon: 0.8153254000040091
steps: 246	lr: 0.0001	evaluation reward: 2.59	
episode: 1017	score: 0.0	memory length: 193392	epsilon: 0.8150818600040144
steps: 123	lr: 0.0001	evaluation reward: 2.57	
episode: 1018	score: 1.0	memory length: 193563	epsilon: 0.8147432800040217
steps: 171	lr: 0.0001	evaluation reward: 2.55	
episode: 1019	score: 2.0	memory length: 193782	epsilon: 0.8143096600040312
steps: 219	lr: 0.0001	evaluation reward: 2.56	
episode: 1020	score: 2.0	memory length: 193980	epsilon: 0.8139176200040397
steps: 198	lr: 0.0001	evaluation reward: 2.57	
episode: 1021	score: 1.0	memory length: 194131	epsilon: 0.8136186400040462
steps: 151	lr: 0.0001	evaluation reward: 2.57	
episode: 1022	score: 8.0	memory length: 194585	epsilon: 0.8127197200040657
steps: 454	lr: 0.0001	evaluation reward: 2.65	
episode: 1023	score: 0.0	memory length: 194707	epsilon: 0.8124781600040709
steps: 122	lr: 0.0001	evaluation reward: 2.63	
episode: 1024	score: 6.0	memory length: 195086	epsilon: 0.8117277400040872
steps: 379	lr: 0.0001	evaluation reward: 2.66	
episode: 1025	score: 5.0	memory length: 195427	epsilon: 0.8110525600041019
steps: 341	lr: 0.0001	evaluation reward: 2.67	
episode: 1026	score: 2.0	memory length: 195624	epsilon: 0.8106625000041103
steps: 197	lr: 0.0001	evaluation reward: 2.65	
episode: 1027	score: 3.0	memory length: 195868	epsilon: 0.8101793800041208
steps: 244	lr: 0.0001	evaluation reward: 2.64	
episode: 1028	score: 1.0	memory length: 196038	epsilon: 0.8098427800041281
steps: 170	lr: 0.0001	evaluation reward: 2.6	
episode: 1029	score: 2.0	memory length: 196256	epsilon: 0.8094111400041375
steps: 218	lr: 0.0001	evaluation reward: 2.6	
episode: 1030	score: 4.0	memory length: 196531	epsilon: 0.8088666400041493
steps: 275	lr: 0.0001	evaluation reward: 2.6	
episode: 1031	score: 4.0	memory length: 196788	epsilon: 0.8083577800041604
steps: 257	lr: 0.0001	evaluation reward: 2.63	
episode: 1032	score: 3.0	memory length: 197032	epsilon: 0.8078746600041709
steps: 244	lr: 0.0001	evaluation reward: 2.65	
episode: 1033	score: 2.0	memory length: 197250	epsilon: 0.8074430200041802
steps: 218	lr: 0.0001	evaluation reward: 2.66	
episode: 1034	score: 2.0	memory length: 197448	epsilon: 0.8070509800041887
steps: 198	lr: 0.0001	evaluation reward: 2.63	
episode: 1035	score: 3.0	memory length: 197712	epsilon: 0.8065282600042001
steps: 264	lr: 0.0001	evaluation reward: 2.63	

episode: 1036	score: 1.0	memory length: 197880	epsilon: 0.8061956200042073
steps: 168	lr: 0.0001	evaluation reward: 2.62	
episode: 1037	score: 1.0	memory length: 198050	epsilon: 0.8058590200042146
steps: 170	lr: 0.0001	evaluation reward: 2.63	
episode: 1038	score: 4.0	memory length: 198347	epsilon: 0.8052709600042274
steps: 297	lr: 0.0001	evaluation reward: 2.65	
episode: 1039	score: 5.0	memory length: 198670	epsilon: 0.8046314200042413
steps: 323	lr: 0.0001	evaluation reward: 2.7	
episode: 1040	score: 2.0	memory length: 198869	epsilon: 0.8042374000042498
steps: 199	lr: 0.0001	evaluation reward: 2.71	
episode: 1041	score: 2.0	memory length: 199051	epsilon: 0.8038770400042576
steps: 182	lr: 0.0001	evaluation reward: 2.67	
episode: 1042	score: 2.0	memory length: 199232	epsilon: 0.8035186600042654
steps: 181	lr: 0.0001	evaluation reward: 2.67	
episode: 1043	score: 3.0	memory length: 199478	epsilon: 0.803031580004276
steps: 246	lr: 0.0001	evaluation reward: 2.68	
episode: 1044	score: 4.0	memory length: 199778	epsilon: 0.8024375800042889
steps: 300	lr: 0.0001	evaluation reward: 2.68	
episode: 1045	score: 4.0	memory length: 200037	epsilon: 0.8019247600043
steps: 259	lr: 0.0001	evaluation reward: 2.69	
episode: 1046	score: 1.0	memory length: 200188	epsilon: 0.8016257800043065
steps: 151	lr: 0.0001	evaluation reward: 2.67	
episode: 1047	score: 4.0	memory length: 200503	epsilon: 0.80100208000432
steps: 315	lr: 0.0001	evaluation reward: 2.69	
episode: 1048	score: 2.0	memory length: 200702	epsilon: 0.8006080600043286
steps: 199	lr: 0.0001	evaluation reward: 2.69	
episode: 1049	score: 2.0	memory length: 200920	epsilon: 0.800176420004338
steps: 218	lr: 0.0001	evaluation reward: 2.68	
episode: 1050	score: 8.0	memory length: 201331	epsilon: 0.7993626400043556
steps: 411	lr: 0.0001	evaluation reward: 2.73	
episode: 1051	score: 5.0	memory length: 201671	epsilon: 0.7986894400043703
steps: 340	lr: 0.0001	evaluation reward: 2.74	
episode: 1052	score: 3.0	memory length: 201898	epsilon: 0.79823998000438
steps: 227	lr: 0.0001	evaluation reward: 2.72	
episode: 1053	score: 4.0	memory length: 202157	epsilon: 0.7977271600043911
steps: 259	lr: 0.0001	evaluation reward: 2.74	
episode: 1054	score: 3.0	memory length: 202385	epsilon: 0.797275720004401
steps: 228	lr: 0.0001	evaluation reward: 2.73	
episode: 1055	score: 5.0	memory length: 202690	epsilon: 0.796671820004414
steps: 305	lr: 0.0001	evaluation reward: 2.75	
episode: 1056	score: 5.0	memory length: 203036	epsilon: 0.7959867400044289
steps: 346	lr: 0.0001	evaluation reward: 2.75	
episode: 1057	score: 3.0	memory length: 203282	epsilon: 0.7954996600044395
steps: 246	lr: 0.0001	evaluation reward: 2.74	
episode: 1058	score: 4.0	memory length: 203560	epsilon: 0.7949492200044515
steps: 278	lr: 0.0001	evaluation reward: 2.78	
episode: 1059	score: 3.0	memory length: 203788	epsilon: 0.7944977800044613
steps: 228	lr: 0.0001	evaluation reward: 2.77	

episode: 1060	score: 3.0	memory length: 204053	epsilon: 0.7939730800044726
steps: 265	lr: 0.0001	evaluation reward: 2.78	
episode: 1061	score: 4.0	memory length: 204307	epsilon: 0.7934701600044836
steps: 254	lr: 0.0001	evaluation reward: 2.8	
episode: 1062	score: 2.0	memory length: 204506	epsilon: 0.7930761400044921
steps: 199	lr: 0.0001	evaluation reward: 2.81	
episode: 1063	score: 4.0	memory length: 204764	epsilon: 0.7925653000045032
steps: 258	lr: 0.0001	evaluation reward: 2.83	
episode: 1064	score: 4.0	memory length: 205038	epsilon: 0.792022780004515
steps: 274	lr: 0.0001	evaluation reward: 2.86	
episode: 1065	score: 2.0	memory length: 205220	epsilon: 0.7916624200045228
steps: 182	lr: 0.0001	evaluation reward: 2.86	
episode: 1066	score: 3.0	memory length: 205464	epsilon: 0.7911793000045333
steps: 244	lr: 0.0001	evaluation reward: 2.84	
episode: 1067	score: 4.0	memory length: 205779	epsilon: 0.7905556000045468
steps: 315	lr: 0.0001	evaluation reward: 2.88	
episode: 1068	score: 2.0	memory length: 205959	epsilon: 0.7901992000045546
steps: 180	lr: 0.0001	evaluation reward: 2.87	
episode: 1069	score: 4.0	memory length: 206233	epsilon: 0.7896566800045663
steps: 274	lr: 0.0001	evaluation reward: 2.88	
episode: 1070	score: 1.0	memory length: 206401	epsilon: 0.7893240400045736
steps: 168	lr: 0.0001	evaluation reward: 2.89	
episode: 1071	score: 3.0	memory length: 206628	epsilon: 0.7888745800045833
steps: 227	lr: 0.0001	evaluation reward: 2.89	
episode: 1072	score: 1.0	memory length: 206779	epsilon: 0.7885756000045898
steps: 151	lr: 0.0001	evaluation reward: 2.89	
episode: 1073	score: 0.0	memory length: 206901	epsilon: 0.7883340400045951
steps: 122	lr: 0.0001	evaluation reward: 2.86	
episode: 1074	score: 2.0	memory length: 207100	epsilon: 0.7879400200046036
steps: 199	lr: 0.0001	evaluation reward: 2.82	
episode: 1075	score: 2.0	memory length: 207279	epsilon: 0.7875856000046113
steps: 179	lr: 0.0001	evaluation reward: 2.81	
episode: 1076	score: 4.0	memory length: 207596	epsilon: 0.7869579400046249
steps: 317	lr: 0.0001	evaluation reward: 2.82	
episode: 1077	score: 4.0	memory length: 207896	epsilon: 0.7863639400046378
steps: 300	lr: 0.0001	evaluation reward: 2.84	
episode: 1078	score: 2.0	memory length: 208114	epsilon: 0.7859323000046472
steps: 218	lr: 0.0001	evaluation reward: 2.85	
episode: 1079	score: 5.0	memory length: 208478	epsilon: 0.7852115800046628
steps: 364	lr: 0.0001	evaluation reward: 2.9	
episode: 1080	score: 2.0	memory length: 208678	epsilon: 0.7848155800046714
steps: 200	lr: 0.0001	evaluation reward: 2.89	
episode: 1081	score: 4.0	memory length: 208954	epsilon: 0.7842691000046833
steps: 276	lr: 0.0001	evaluation reward: 2.91	
episode: 1082	score: 3.0	memory length: 209183	epsilon: 0.7838156800046931
steps: 229	lr: 0.0001	evaluation reward: 2.91	
episode: 1083	score: 4.0	memory length: 209473	epsilon: 0.7832414800047056
steps: 290	lr: 0.0001	evaluation reward: 2.95	

episode: 1084	score: 5.0	memory length: 209777	epsilon: 0.7826395600047187
steps: 304	lr: 0.0001	evaluation reward: 2.98	
episode: 1085	score: 3.0	memory length: 210042	epsilon: 0.7821148600047301
steps: 265	lr: 0.0001	evaluation reward: 2.98	
episode: 1086	score: 4.0	memory length: 210337	epsilon: 0.7815307600047428
steps: 295	lr: 0.0001	evaluation reward: 2.99	
episode: 1087	score: 4.0	memory length: 210612	epsilon: 0.7809862600047546
steps: 275	lr: 0.0001	evaluation reward: 3.01	
episode: 1088	score: 3.0	memory length: 210859	epsilon: 0.7804972000047652
steps: 247	lr: 0.0001	evaluation reward: 3.01	
episode: 1089	score: 2.0	memory length: 211074	epsilon: 0.7800715000047744
steps: 215	lr: 0.0001	evaluation reward: 3.01	
episode: 1090	score: 3.0	memory length: 211284	epsilon: 0.7796557000047835
steps: 210	lr: 0.0001	evaluation reward: 3.03	
episode: 1091	score: 3.0	memory length: 211531	epsilon: 0.7791666400047941
steps: 247	lr: 0.0001	evaluation reward: 3.05	
episode: 1092	score: 7.0	memory length: 211911	epsilon: 0.7784142400048104
steps: 380	lr: 0.0001	evaluation reward: 3.1	
episode: 1093	score: 4.0	memory length: 212206	epsilon: 0.7778301400048231
steps: 295	lr: 0.0001	evaluation reward: 3.13	
episode: 1094	score: 6.0	memory length: 212602	epsilon: 0.7770460600048401
steps: 396	lr: 0.0001	evaluation reward: 3.13	
episode: 1095	score: 3.0	memory length: 212836	epsilon: 0.7765827400048502
steps: 234	lr: 0.0001	evaluation reward: 3.14	
episode: 1096	score: 3.0	memory length: 213066	epsilon: 0.7761273400048601
steps: 230	lr: 0.0001	evaluation reward: 3.11	
episode: 1097	score: 3.0	memory length: 213293	epsilon: 0.7756778800048698
steps: 227	lr: 0.0001	evaluation reward: 3.13	
episode: 1098	score: 2.0	memory length: 213472	epsilon: 0.7753234600048775
steps: 179	lr: 0.0001	evaluation reward: 3.13	
episode: 1099	score: 1.0	memory length: 213641	epsilon: 0.7749888400048848
steps: 169	lr: 0.0001	evaluation reward: 3.1	
episode: 1100	score: 5.0	memory length: 213984	epsilon: 0.7743097000048995
steps: 343	lr: 0.0001	evaluation reward: 3.09	
episode: 1101	score: 3.0	memory length: 214251	epsilon: 0.773781040004911
steps: 267	lr: 0.0001	evaluation reward: 3.09	
episode: 1102	score: 1.0	memory length: 214401	epsilon: 0.7734840400049174
steps: 150	lr: 0.0001	evaluation reward: 3.1	
episode: 1103	score: 6.0	memory length: 214776	epsilon: 0.7727415400049336
steps: 375	lr: 0.0001	evaluation reward: 3.12	
episode: 1104	score: 4.0	memory length: 215072	epsilon: 0.7721554600049463
steps: 296	lr: 0.0001	evaluation reward: 3.15	
episode: 1105	score: 3.0	memory length: 215300	epsilon: 0.7717040200049561
steps: 228	lr: 0.0001	evaluation reward: 3.15	
episode: 1106	score: 5.0	memory length: 215591	epsilon: 0.7711278400049686
steps: 291	lr: 0.0001	evaluation reward: 3.18	
episode: 1107	score: 3.0	memory length: 215856	epsilon: 0.77060314000498
steps: 265	lr: 0.0001	evaluation reward: 3.15	

episode: 1108	score: 2.0	memory length: 216054	epsilon: 0.7702111000049885
steps: 198	lr: 0.0001	evaluation reward: 3.14	
episode: 1109	score: 3.0	memory length: 216280	epsilon: 0.7697636200049982
steps: 226	lr: 0.0001	evaluation reward: 3.11	
episode: 1110	score: 5.0	memory length: 216605	epsilon: 0.7691201200050122
steps: 325	lr: 0.0001	evaluation reward: 3.12	
episode: 1111	score: 5.0	memory length: 216948	epsilon: 0.7684409800050269
steps: 343	lr: 0.0001	evaluation reward: 3.15	
episode: 1112	score: 3.0	memory length: 217195	epsilon: 0.7679519200050375
steps: 247	lr: 0.0001	evaluation reward: 3.17	
episode: 1113	score: 3.0	memory length: 217428	epsilon: 0.7674905800050476
steps: 233	lr: 0.0001	evaluation reward: 3.17	
episode: 1114	score: 1.0	memory length: 217578	epsilon: 0.767193580005054
steps: 150	lr: 0.0001	evaluation reward: 3.14	
episode: 1115	score: 4.0	memory length: 217875	epsilon: 0.7666055200050668
steps: 297	lr: 0.0001	evaluation reward: 3.15	
episode: 1116	score: 1.0	memory length: 218026	epsilon: 0.7663065400050733
steps: 151	lr: 0.0001	evaluation reward: 3.13	
episode: 1117	score: 3.0	memory length: 218273	epsilon: 0.7658174800050839
steps: 247	lr: 0.0001	evaluation reward: 3.16	
episode: 1118	score: 3.0	memory length: 218499	epsilon: 0.7653700000050936
steps: 226	lr: 0.0001	evaluation reward: 3.18	
episode: 1119	score: 2.0	memory length: 218722	epsilon: 0.7649284600051032
steps: 223	lr: 0.0001	evaluation reward: 3.18	
episode: 1120	score: 2.0	memory length: 218940	epsilon: 0.7644968200051125
steps: 218	lr: 0.0001	evaluation reward: 3.18	
episode: 1121	score: 4.0	memory length: 219216	epsilon: 0.7639503400051244
steps: 276	lr: 0.0001	evaluation reward: 3.21	
episode: 1122	score: 1.0	memory length: 219366	epsilon: 0.7636533400051309
steps: 150	lr: 0.0001	evaluation reward: 3.14	
episode: 1123	score: 1.0	memory length: 219538	epsilon: 0.7633127800051382
steps: 172	lr: 0.0001	evaluation reward: 3.15	
episode: 1124	score: 5.0	memory length: 219863	epsilon: 0.7626692800051522
steps: 325	lr: 0.0001	evaluation reward: 3.14	
episode: 1125	score: 3.0	memory length: 220109	epsilon: 0.7621822000051628
steps: 246	lr: 0.0001	evaluation reward: 3.12	
episode: 1126	score: 2.0	memory length: 220289	epsilon: 0.7618258000051705
steps: 180	lr: 0.0001	evaluation reward: 3.12	
episode: 1127	score: 6.0	memory length: 220635	epsilon: 0.7611407200051854
steps: 346	lr: 0.0001	evaluation reward: 3.15	
episode: 1128	score: 0.0	memory length: 220758	epsilon: 0.7608971800051907
steps: 123	lr: 0.0001	evaluation reward: 3.14	
episode: 1129	score: 4.0	memory length: 221015	epsilon: 0.7603883200052017
steps: 257	lr: 0.0001	evaluation reward: 3.16	
episode: 1130	score: 6.0	memory length: 221318	epsilon: 0.7597883800052148
steps: 303	lr: 0.0001	evaluation reward: 3.18	
episode: 1131	score: 4.0	memory length: 221614	epsilon: 0.7592023000052275
steps: 296	lr: 0.0001	evaluation reward: 3.18	

episode: 1132	score: 0.0	memory length: 221737	epsilon: 0.7589587600052328
steps: 123	lr: 0.0001	evaluation reward: 3.15	
episode: 1133	score: 5.0	memory length: 222078	epsilon: 0.7582835800052474
steps: 341	lr: 0.0001	evaluation reward: 3.18	
episode: 1134	score: 3.0	memory length: 222325	epsilon: 0.757794520005258
steps: 247	lr: 0.0001	evaluation reward: 3.19	
episode: 1135	score: 6.0	memory length: 222713	epsilon: 0.7570262800052747
steps: 388	lr: 0.0001	evaluation reward: 3.22	
episode: 1136	score: 6.0	memory length: 223084	epsilon: 0.7562917000052907
steps: 371	lr: 0.0001	evaluation reward: 3.27	
episode: 1137	score: 4.0	memory length: 223376	epsilon: 0.7557135400053032
steps: 292	lr: 0.0001	evaluation reward: 3.3	
episode: 1138	score: 4.0	memory length: 223642	epsilon: 0.7551868600053147
steps: 266	lr: 0.0001	evaluation reward: 3.3	
episode: 1139	score: 5.0	memory length: 223986	epsilon: 0.7545057400053294
steps: 344	lr: 0.0001	evaluation reward: 3.3	
episode: 1140	score: 0.0	memory length: 224108	epsilon: 0.7542641800053347
steps: 122	lr: 0.0001	evaluation reward: 3.28	
episode: 1141	score: 3.0	memory length: 224339	epsilon: 0.7538068000053446
steps: 231	lr: 0.0001	evaluation reward: 3.29	
episode: 1142	score: 1.0	memory length: 224490	epsilon: 0.7535078200053511
steps: 151	lr: 0.0001	evaluation reward: 3.28	
episode: 1143	score: 2.0	memory length: 224708	epsilon: 0.7530761800053605
steps: 218	lr: 0.0001	evaluation reward: 3.27	
episode: 1144	score: 3.0	memory length: 224975	epsilon: 0.752547520005372
steps: 267	lr: 0.0001	evaluation reward: 3.26	
episode: 1145	score: 6.0	memory length: 225366	epsilon: 0.7517733400053888
steps: 391	lr: 0.0001	evaluation reward: 3.28	
episode: 1146	score: 5.0	memory length: 225673	epsilon: 0.751165480005402
steps: 307	lr: 0.0001	evaluation reward: 3.32	
episode: 1147	score: 6.0	memory length: 226031	epsilon: 0.7504566400054173
steps: 358	lr: 0.0001	evaluation reward: 3.34	
episode: 1148	score: 3.0	memory length: 226261	epsilon: 0.7500012400054272
steps: 230	lr: 0.0001	evaluation reward: 3.35	
episode: 1149	score: 2.0	memory length: 226441	epsilon: 0.749644840005435
steps: 180	lr: 0.0001	evaluation reward: 3.35	
episode: 1150	score: 1.0	memory length: 226610	epsilon: 0.7493102200054422
steps: 169	lr: 0.0001	evaluation reward: 3.28	
episode: 1151	score: 2.0	memory length: 226790	epsilon: 0.74895382000545
steps: 180	lr: 0.0001	evaluation reward: 3.25	
episode: 1152	score: 1.0	memory length: 226942	epsilon: 0.7486528600054565
steps: 152	lr: 0.0001	evaluation reward: 3.23	
episode: 1153	score: 9.0	memory length: 227286	epsilon: 0.7479717400054713
steps: 344	lr: 0.0001	evaluation reward: 3.28	
episode: 1154	score: 6.0	memory length: 227639	epsilon: 0.7472728000054865
steps: 353	lr: 0.0001	evaluation reward: 3.31	
episode: 1155	score: 4.0	memory length: 227916	epsilon: 0.7467243400054984
steps: 277	lr: 0.0001	evaluation reward: 3.3	

episode: 1156	score: 2.0	memory length: 228132	epsilon: 0.7462966600055077
steps: 216	lr: 0.0001	evaluation reward: 3.27	
episode: 1157	score: 4.0	memory length: 228404	epsilon: 0.7457581000055193
steps: 272	lr: 0.0001	evaluation reward: 3.28	
episode: 1158	score: 3.0	memory length: 228630	epsilon: 0.7453106200055291
steps: 226	lr: 0.0001	evaluation reward: 3.27	
episode: 1159	score: 7.0	memory length: 228995	epsilon: 0.7445879200055447
steps: 365	lr: 0.0001	evaluation reward: 3.31	
episode: 1160	score: 4.0	memory length: 229260	epsilon: 0.7440632200055561
steps: 265	lr: 0.0001	evaluation reward: 3.32	
episode: 1161	score: 1.0	memory length: 229410	epsilon: 0.7437662200055626
steps: 150	lr: 0.0001	evaluation reward: 3.29	
episode: 1162	score: 4.0	memory length: 229706	epsilon: 0.7431801400055753
steps: 296	lr: 0.0001	evaluation reward: 3.31	
episode: 1163	score: 0.0	memory length: 229829	epsilon: 0.7429366000055806
steps: 123	lr: 0.0001	evaluation reward: 3.27	
episode: 1164	score: 3.0	memory length: 230059	epsilon: 0.7424812000055905
steps: 230	lr: 0.0001	evaluation reward: 3.26	
episode: 1165	score: 5.0	memory length: 230405	epsilon: 0.7417961200056054
steps: 346	lr: 0.0001	evaluation reward: 3.29	
episode: 1166	score: 4.0	memory length: 230718	epsilon: 0.7411763800056188
steps: 313	lr: 0.0001	evaluation reward: 3.3	
episode: 1167	score: 2.0	memory length: 230918	epsilon: 0.7407803800056274
steps: 200	lr: 0.0001	evaluation reward: 3.28	
episode: 1168	score: 1.0	memory length: 231069	epsilon: 0.7404814000056339
steps: 151	lr: 0.0001	evaluation reward: 3.27	
episode: 1169	score: 4.0	memory length: 231364	epsilon: 0.7398973000056466
steps: 295	lr: 0.0001	evaluation reward: 3.27	
episode: 1170	score: 3.0	memory length: 231633	epsilon: 0.7393646800056581
steps: 269	lr: 0.0001	evaluation reward: 3.29	
episode: 1171	score: 6.0	memory length: 232008	epsilon: 0.7386221800056743
steps: 375	lr: 0.0001	evaluation reward: 3.32	
episode: 1172	score: 3.0	memory length: 232254	epsilon: 0.7381351000056848
steps: 246	lr: 0.0001	evaluation reward: 3.34	
episode: 1173	score: 4.0	memory length: 232551	epsilon: 0.7375470400056976
steps: 297	lr: 0.0001	evaluation reward: 3.38	
episode: 1174	score: 3.0	memory length: 232821	epsilon: 0.7370124400057092
steps: 270	lr: 0.0001	evaluation reward: 3.39	
episode: 1175	score: 5.0	memory length: 233127	epsilon: 0.7364065600057224
steps: 306	lr: 0.0001	evaluation reward: 3.42	
episode: 1176	score: 5.0	memory length: 233455	epsilon: 0.7357571200057365
steps: 328	lr: 0.0001	evaluation reward: 3.43	
episode: 1177	score: 4.0	memory length: 233730	epsilon: 0.7352126200057483
steps: 275	lr: 0.0001	evaluation reward: 3.43	
episode: 1178	score: 6.0	memory length: 234085	epsilon: 0.7345097200057635
steps: 355	lr: 0.0001	evaluation reward: 3.47	
episode: 1179	score: 5.0	memory length: 234394	epsilon: 0.7338979000057768
steps: 309	lr: 0.0001	evaluation reward: 3.47	

episode: 1180	score: 7.0	memory length: 234813	epsilon: 0.7330682800057948
steps: 419	lr: 0.0001	evaluation reward: 3.52	
episode: 1181	score: 7.0	memory length: 235198	epsilon: 0.7323059800058114
steps: 385	lr: 0.0001	evaluation reward: 3.55	
episode: 1182	score: 4.0	memory length: 235479	epsilon: 0.7317496000058235
steps: 281	lr: 0.0001	evaluation reward: 3.56	
episode: 1183	score: 2.0	memory length: 235677	epsilon: 0.731357560005832
steps: 198	lr: 0.0001	evaluation reward: 3.54	
episode: 1184	score: 2.0	memory length: 235874	epsilon: 0.7309675000058404
steps: 197	lr: 0.0001	evaluation reward: 3.51	
episode: 1185	score: 2.0	memory length: 236056	epsilon: 0.7306071400058483
steps: 182	lr: 0.0001	evaluation reward: 3.5	
episode: 1186	score: 7.0	memory length: 236410	epsilon: 0.7299062200058635
steps: 354	lr: 0.0001	evaluation reward: 3.53	
episode: 1187	score: 7.0	memory length: 236821	epsilon: 0.7290924400058811
steps: 411	lr: 0.0001	evaluation reward: 3.56	
episode: 1188	score: 5.0	memory length: 237164	epsilon: 0.7284133000058959
steps: 343	lr: 0.0001	evaluation reward: 3.58	
episode: 1189	score: 3.0	memory length: 237392	epsilon: 0.7279618600059057
steps: 228	lr: 0.0001	evaluation reward: 3.59	
episode: 1190	score: 3.0	memory length: 237636	epsilon: 0.7274787400059162
steps: 244	lr: 0.0001	evaluation reward: 3.59	
episode: 1191	score: 5.0	memory length: 237963	epsilon: 0.7268312800059302
steps: 327	lr: 0.0001	evaluation reward: 3.61	
episode: 1192	score: 4.0	memory length: 238258	epsilon: 0.7262471800059429
steps: 295	lr: 0.0001	evaluation reward: 3.58	
episode: 1193	score: 3.0	memory length: 238486	epsilon: 0.7257957400059527
steps: 228	lr: 0.0001	evaluation reward: 3.57	
episode: 1194	score: 5.0	memory length: 238832	epsilon: 0.7251106600059676
steps: 346	lr: 0.0001	evaluation reward: 3.56	
episode: 1195	score: 4.0	memory length: 239092	epsilon: 0.7245958600059788
steps: 260	lr: 0.0001	evaluation reward: 3.57	
episode: 1196	score: 5.0	memory length: 239439	epsilon: 0.7239088000059937
steps: 347	lr: 0.0001	evaluation reward: 3.59	
episode: 1197	score: 8.0	memory length: 239768	epsilon: 0.7232573800060078
steps: 329	lr: 0.0001	evaluation reward: 3.64	
episode: 1198	score: 4.0	memory length: 240028	epsilon: 0.722742580006019
steps: 260	lr: 0.0001	evaluation reward: 3.66	
episode: 1199	score: 2.0	memory length: 240226	epsilon: 0.7223505400060275
steps: 198	lr: 0.0001	evaluation reward: 3.67	
episode: 1200	score: 2.0	memory length: 240424	epsilon: 0.721958500006036
steps: 198	lr: 0.0001	evaluation reward: 3.64	
episode: 1201	score: 1.0	memory length: 240593	epsilon: 0.7216238800060433
steps: 169	lr: 0.0001	evaluation reward: 3.62	
episode: 1202	score: 1.0	memory length: 240744	epsilon: 0.7213249000060498
steps: 151	lr: 0.0001	evaluation reward: 3.62	
episode: 1203	score: 6.0	memory length: 241156	epsilon: 0.7205091400060675
steps: 412	lr: 0.0001	evaluation reward: 3.62	

episode: 1204	score: 3.0	memory length: 241402	epsilon: 0.720022060006078
steps: 246	lr: 0.0001	evaluation reward: 3.61	
episode: 1205	score: 4.0	memory length: 241657	epsilon: 0.719517160006089
steps: 255	lr: 0.0001	evaluation reward: 3.62	
episode: 1206	score: 3.0	memory length: 241888	epsilon: 0.7190597800060989
steps: 231	lr: 0.0001	evaluation reward: 3.6	
episode: 1207	score: 1.0	memory length: 242039	epsilon: 0.7187608000061054
steps: 151	lr: 0.0001	evaluation reward: 3.58	
episode: 1208	score: 6.0	memory length: 242423	epsilon: 0.7180004800061219
steps: 384	lr: 0.0001	evaluation reward: 3.62	
episode: 1209	score: 7.0	memory length: 242688	epsilon: 0.7174757800061333
steps: 265	lr: 0.0001	evaluation reward: 3.66	
episode: 1210	score: 6.0	memory length: 243061	epsilon: 0.7167372400061494
steps: 373	lr: 0.0001	evaluation reward: 3.67	
episode: 1211	score: 3.0	memory length: 243289	epsilon: 0.7162858000061592
steps: 228	lr: 0.0001	evaluation reward: 3.65	
episode: 1212	score: 3.0	memory length: 243500	epsilon: 0.7158680200061682
steps: 211	lr: 0.0001	evaluation reward: 3.65	
episode: 1213	score: 3.0	memory length: 243731	epsilon: 0.7154106400061782
steps: 231	lr: 0.0001	evaluation reward: 3.65	
episode: 1214	score: 4.0	memory length: 244008	epsilon: 0.7148621800061901
steps: 277	lr: 0.0001	evaluation reward: 3.68	
episode: 1215	score: 4.0	memory length: 244266	epsilon: 0.7143513400062012
steps: 258	lr: 0.0001	evaluation reward: 3.68	
episode: 1216	score: 5.0	memory length: 244592	epsilon: 0.7137058600062152
steps: 326	lr: 0.0001	evaluation reward: 3.72	
episode: 1217	score: 2.0	memory length: 244809	epsilon: 0.7132762000062245
steps: 217	lr: 0.0001	evaluation reward: 3.71	
episode: 1218	score: 5.0	memory length: 245145	epsilon: 0.7126109200062389
steps: 336	lr: 0.0001	evaluation reward: 3.73	
episode: 1219	score: 2.0	memory length: 245361	epsilon: 0.7121832400062482
steps: 216	lr: 0.0001	evaluation reward: 3.73	
episode: 1220	score: 3.0	memory length: 245589	epsilon: 0.711731800006258
steps: 228	lr: 0.0001	evaluation reward: 3.74	
episode: 1221	score: 3.0	memory length: 245837	epsilon: 0.7112407600062687
steps: 248	lr: 0.0001	evaluation reward: 3.73	
episode: 1222	score: 7.0	memory length: 246268	epsilon: 0.7103873800062872
steps: 431	lr: 0.0001	evaluation reward: 3.79	
episode: 1223	score: 5.0	memory length: 246592	epsilon: 0.7097458600063011
steps: 324	lr: 0.0001	evaluation reward: 3.83	
episode: 1224	score: 2.0	memory length: 246807	epsilon: 0.7093201600063104
steps: 215	lr: 0.0001	evaluation reward: 3.8	
episode: 1225	score: 3.0	memory length: 247055	epsilon: 0.708829120006321
steps: 248	lr: 0.0001	evaluation reward: 3.8	
episode: 1226	score: 5.0	memory length: 247370	epsilon: 0.7082054200063346
steps: 315	lr: 0.0001	evaluation reward: 3.83	
episode: 1227	score: 5.0	memory length: 247716	epsilon: 0.7075203400063494
steps: 346	lr: 0.0001	evaluation reward: 3.82	

episode: 1228	score: 3.0	memory length: 247930	epsilon: 0.7070966200063586
steps: 214	lr: 0.0001	evaluation reward: 3.85	
episode: 1229	score: 5.0	memory length: 248235	epsilon: 0.7064927200063718
steps: 305	lr: 0.0001	evaluation reward: 3.86	
episode: 1230	score: 2.0	memory length: 248456	epsilon: 0.7060551400063813
steps: 221	lr: 0.0001	evaluation reward: 3.82	
episode: 1231	score: 2.0	memory length: 248654	epsilon: 0.7056631000063898
steps: 198	lr: 0.0001	evaluation reward: 3.8	
episode: 1232	score: 5.0	memory length: 248962	epsilon: 0.705053260006403
steps: 308	lr: 0.0001	evaluation reward: 3.85	
episode: 1233	score: 2.0	memory length: 249162	epsilon: 0.7046572600064116
steps: 200	lr: 0.0001	evaluation reward: 3.82	
episode: 1234	score: 1.0	memory length: 249312	epsilon: 0.704360260006418
steps: 150	lr: 0.0001	evaluation reward: 3.8	
episode: 1235	score: 6.0	memory length: 249676	epsilon: 0.7036395400064337
steps: 364	lr: 0.0001	evaluation reward: 3.8	
episode: 1236	score: 3.0	memory length: 249906	epsilon: 0.7031841400064436
steps: 230	lr: 0.0001	evaluation reward: 3.77	
episode: 1237	score: 4.0	memory length: 250201	epsilon: 0.7026000400064563
steps: 295	lr: 0.0001	evaluation reward: 3.77	
episode: 1238	score: 4.0	memory length: 250492	epsilon: 0.7020238600064688
steps: 291	lr: 0.0001	evaluation reward: 3.77	
episode: 1239	score: 5.0	memory length: 250837	epsilon: 0.7013407600064836
steps: 345	lr: 0.0001	evaluation reward: 3.77	
episode: 1240	score: 4.0	memory length: 251113	epsilon: 0.7007942800064955
steps: 276	lr: 0.0001	evaluation reward: 3.81	
episode: 1241	score: 4.0	memory length: 251368	epsilon: 0.7002893800065064
steps: 255	lr: 0.0001	evaluation reward: 3.82	
episode: 1242	score: 4.0	memory length: 251645	epsilon: 0.6997409200065183
steps: 277	lr: 0.0001	evaluation reward: 3.85	
episode: 1243	score: 3.0	memory length: 251874	epsilon: 0.6992875000065282
steps: 229	lr: 0.0001	evaluation reward: 3.86	
episode: 1244	score: 5.0	memory length: 252201	epsilon: 0.6986400400065422
steps: 327	lr: 0.0001	evaluation reward: 3.88	
episode: 1245	score: 2.0	memory length: 252399	epsilon: 0.6982480000065507
steps: 198	lr: 0.0001	evaluation reward: 3.84	
episode: 1246	score: 4.0	memory length: 252716	epsilon: 0.6976203400065644
steps: 317	lr: 0.0001	evaluation reward: 3.83	
episode: 1247	score: 4.0	memory length: 252977	epsilon: 0.6971035600065756
steps: 261	lr: 0.0001	evaluation reward: 3.81	
episode: 1248	score: 8.0	memory length: 253410	epsilon: 0.6962462200065942
steps: 433	lr: 0.0001	evaluation reward: 3.86	
episode: 1249	score: 5.0	memory length: 253734	epsilon: 0.6956047000066081
steps: 324	lr: 0.0001	evaluation reward: 3.89	
episode: 1250	score: 3.0	memory length: 253961	epsilon: 0.6951552400066179
steps: 227	lr: 0.0001	evaluation reward: 3.91	
episode: 1251	score: 0.0	memory length: 254084	epsilon: 0.6949117000066232
steps: 123	lr: 0.0001	evaluation reward: 3.89	

episode: 1252	score: 6.0	memory length: 254456	epsilon: 0.6941751400066392
steps: 372	lr: 0.0001	evaluation reward: 3.94	
episode: 1253	score: 5.0	memory length: 254781	epsilon: 0.6935316400066531
steps: 325	lr: 0.0001	evaluation reward: 3.9	
episode: 1254	score: 8.0	memory length: 255232	epsilon: 0.6926386600066725
steps: 451	lr: 0.0001	evaluation reward: 3.92	
episode: 1255	score: 2.0	memory length: 255431	epsilon: 0.6922446400066811
steps: 199	lr: 0.0001	evaluation reward: 3.9	
episode: 1256	score: 2.0	memory length: 255613	epsilon: 0.6918842800066889
steps: 182	lr: 0.0001	evaluation reward: 3.9	
episode: 1257	score: 7.0	memory length: 255998	epsilon: 0.6911219800067054
steps: 385	lr: 0.0001	evaluation reward: 3.93	
episode: 1258	score: 5.0	memory length: 256322	epsilon: 0.6904804600067194
steps: 324	lr: 0.0001	evaluation reward: 3.95	
episode: 1259	score: 4.0	memory length: 256581	epsilon: 0.6899676400067305
steps: 259	lr: 0.0001	evaluation reward: 3.92	
episode: 1260	score: 4.0	memory length: 256878	epsilon: 0.6893795800067433
steps: 297	lr: 0.0001	evaluation reward: 3.92	
episode: 1261	score: 10.0	memory length: 257375	epsilon:
0.6883955200067646	steps: 497	lr: 0.0001	evaluation reward: 4.01
episode: 1262	score: 1.0	memory length: 257546	epsilon: 0.688056940006772
steps: 171	lr: 0.0001	evaluation reward: 3.98	
episode: 1263	score: 5.0	memory length: 257854	epsilon: 0.6874471000067852
steps: 308	lr: 0.0001	evaluation reward: 4.03	
episode: 1264	score: 2.0	memory length: 258072	epsilon: 0.6870154600067946
steps: 218	lr: 0.0001	evaluation reward: 4.02	
episode: 1265	score: 9.0	memory length: 258503	epsilon: 0.6861620800068131
steps: 431	lr: 0.0001	evaluation reward: 4.06	
episode: 1266	score: 5.0	memory length: 258828	epsilon: 0.6855185800068271
steps: 325	lr: 0.0001	evaluation reward: 4.07	
episode: 1267	score: 1.0	memory length: 258978	epsilon: 0.6852215800068335
steps: 150	lr: 0.0001	evaluation reward: 4.06	
episode: 1268	score: 2.0	memory length: 259178	epsilon: 0.6848255800068421
steps: 200	lr: 0.0001	evaluation reward: 4.07	
episode: 1269	score: 2.0	memory length: 259376	epsilon: 0.6844335400068506
steps: 198	lr: 0.0001	evaluation reward: 4.05	
episode: 1270	score: 2.0	memory length: 259597	epsilon: 0.6839959600068601
steps: 221	lr: 0.0001	evaluation reward: 4.04	
episode: 1271	score: 4.0	memory length: 259874	epsilon: 0.683447500006872
steps: 277	lr: 0.0001	evaluation reward: 4.02	
episode: 1272	score: 5.0	memory length: 260219	epsilon: 0.6827644000068869
steps: 345	lr: 0.0001	evaluation reward: 4.04	
episode: 1273	score: 4.0	memory length: 260518	epsilon: 0.6821723800068997
steps: 299	lr: 0.0001	evaluation reward: 4.04	
episode: 1274	score: 2.0	memory length: 260736	epsilon: 0.6817407400069091
steps: 218	lr: 0.0001	evaluation reward: 4.03	
episode: 1275	score: 2.0	memory length: 260933	epsilon: 0.6813506800069176
steps: 197	lr: 0.0001	evaluation reward: 4.0	

episode: 1276	score: 3.0	memory length: 261180	epsilon: 0.6808616200069282
steps: 247	lr: 0.0001	evaluation reward: 3.98	
episode: 1277	score: 4.0	memory length: 261453	epsilon: 0.6803210800069399
steps: 273	lr: 0.0001	evaluation reward: 3.98	
episode: 1278	score: 1.0	memory length: 261625	epsilon: 0.6799805200069473
steps: 172	lr: 0.0001	evaluation reward: 3.93	
episode: 1279	score: 3.0	memory length: 261873	epsilon: 0.679489480006958
steps: 248	lr: 0.0001	evaluation reward: 3.91	
episode: 1280	score: 3.0	memory length: 262098	epsilon: 0.6790439800069676
steps: 225	lr: 0.0001	evaluation reward: 3.87	
episode: 1281	score: 5.0	memory length: 262405	epsilon: 0.6784361200069808
steps: 307	lr: 0.0001	evaluation reward: 3.85	
episode: 1282	score: 7.0	memory length: 262814	epsilon: 0.6776263000069984
steps: 409	lr: 0.0001	evaluation reward: 3.88	
episode: 1283	score: 3.0	memory length: 263040	epsilon: 0.6771788200070081
steps: 226	lr: 0.0001	evaluation reward: 3.89	
episode: 1284	score: 3.0	memory length: 263250	epsilon: 0.6767630200070172
steps: 210	lr: 0.0001	evaluation reward: 3.9	
episode: 1285	score: 7.0	memory length: 263666	epsilon: 0.675939340007035
steps: 416	lr: 0.0001	evaluation reward: 3.95	
episode: 1286	score: 4.0	memory length: 263957	epsilon: 0.6753631600070475
steps: 291	lr: 0.0001	evaluation reward: 3.92	
episode: 1287	score: 3.0	memory length: 264201	epsilon: 0.674880040007058
steps: 244	lr: 0.0001	evaluation reward: 3.88	
episode: 1288	score: 6.0	memory length: 264553	epsilon: 0.6741830800070732
steps: 352	lr: 0.0001	evaluation reward: 3.89	
episode: 1289	score: 7.0	memory length: 264965	epsilon: 0.6733673200070909
steps: 412	lr: 0.0001	evaluation reward: 3.93	
episode: 1290	score: 3.0	memory length: 265191	epsilon: 0.6729198400071006
steps: 226	lr: 0.0001	evaluation reward: 3.93	
episode: 1291	score: 2.0	memory length: 265388	epsilon: 0.6725297800071091
steps: 197	lr: 0.0001	evaluation reward: 3.9	
episode: 1292	score: 3.0	memory length: 265618	epsilon: 0.6720743800071189
steps: 230	lr: 0.0001	evaluation reward: 3.89	
episode: 1293	score: 2.0	memory length: 265815	epsilon: 0.6716843200071274
steps: 197	lr: 0.0001	evaluation reward: 3.88	
episode: 1294	score: 4.0	memory length: 266091	epsilon: 0.6711378400071393
steps: 276	lr: 0.0001	evaluation reward: 3.87	
episode: 1295	score: 4.0	memory length: 266346	epsilon: 0.6706329400071502
steps: 255	lr: 0.0001	evaluation reward: 3.87	
episode: 1296	score: 10.0	memory length: 266707	epsilon:
0.6699181600071658	steps: 361	lr: 0.0001	evaluation reward: 3.92
episode: 1297	score: 3.0	memory length: 266916	epsilon: 0.6695043400071747
steps: 209	lr: 0.0001	evaluation reward: 3.87	
episode: 1298	score: 5.0	memory length: 267284	epsilon: 0.6687757000071906
steps: 368	lr: 0.0001	evaluation reward: 3.88	
episode: 1299	score: 4.0	memory length: 267544	epsilon: 0.6682609000072017
steps: 260	lr: 0.0001	evaluation reward: 3.9	

episode: 1300	score: 6.0	memory length: 267902	epsilon: 0.6675520600072171
steps: 358	lr: 0.0001	evaluation reward: 3.94	
episode: 1301	score: 5.0	memory length: 268206	epsilon: 0.6669501400072302
steps: 304	lr: 0.0001	evaluation reward: 3.98	
episode: 1302	score: 5.0	memory length: 268531	epsilon: 0.6663066400072442
steps: 325	lr: 0.0001	evaluation reward: 4.02	
episode: 1303	score: 3.0	memory length: 268757	epsilon: 0.6658591600072539
steps: 226	lr: 0.0001	evaluation reward: 3.99	
episode: 1304	score: 5.0	memory length: 269078	epsilon: 0.6652235800072677
steps: 321	lr: 0.0001	evaluation reward: 4.01	
episode: 1305	score: 1.0	memory length: 269249	epsilon: 0.664885000007275
steps: 171	lr: 0.0001	evaluation reward: 3.98	
episode: 1306	score: 4.0	memory length: 269507	epsilon: 0.6643741600072861
steps: 258	lr: 0.0001	evaluation reward: 3.99	
episode: 1307	score: 4.0	memory length: 269769	epsilon: 0.6638554000072974
steps: 262	lr: 0.0001	evaluation reward: 4.02	
episode: 1308	score: 5.0	memory length: 270095	epsilon: 0.6632099200073114
steps: 326	lr: 0.0001	evaluation reward: 4.01	
episode: 1309	score: 3.0	memory length: 270325	epsilon: 0.6627545200073213
steps: 230	lr: 0.0001	evaluation reward: 3.97	
episode: 1310	score: 2.0	memory length: 270505	epsilon: 0.662398120007329
steps: 180	lr: 0.0001	evaluation reward: 3.93	
episode: 1311	score: 4.0	memory length: 270800	epsilon: 0.6618140200073417
steps: 295	lr: 0.0001	evaluation reward: 3.94	
episode: 1312	score: 5.0	memory length: 271107	epsilon: 0.6612061600073549
steps: 307	lr: 0.0001	evaluation reward: 3.96	
episode: 1313	score: 3.0	memory length: 271354	epsilon: 0.6607171000073655
steps: 247	lr: 0.0001	evaluation reward: 3.96	
episode: 1314	score: 5.0	memory length: 271645	epsilon: 0.660140920007378
steps: 291	lr: 0.0001	evaluation reward: 3.97	
episode: 1315	score: 1.0	memory length: 271814	epsilon: 0.6598063000073853
steps: 169	lr: 0.0001	evaluation reward: 3.94	
episode: 1316	score: 6.0	memory length: 272202	epsilon: 0.659038060007402
steps: 388	lr: 0.0001	evaluation reward: 3.95	
episode: 1317	score: 5.0	memory length: 272564	epsilon: 0.6583213000074175
steps: 362	lr: 0.0001	evaluation reward: 3.98	
episode: 1318	score: 5.0	memory length: 272907	epsilon: 0.6576421600074323
steps: 343	lr: 0.0001	evaluation reward: 3.98	
episode: 1319	score: 8.0	memory length: 273363	epsilon: 0.6567392800074519
steps: 456	lr: 0.0001	evaluation reward: 4.04	
episode: 1320	score: 3.0	memory length: 273590	epsilon: 0.6562898200074616
steps: 227	lr: 0.0001	evaluation reward: 4.04	
episode: 1321	score: 4.0	memory length: 273849	epsilon: 0.6557770000074727
steps: 259	lr: 0.0001	evaluation reward: 4.05	
episode: 1322	score: 7.0	memory length: 274260	epsilon: 0.6549632200074904
steps: 411	lr: 0.0001	evaluation reward: 4.05	
episode: 1323	score: 5.0	memory length: 274587	epsilon: 0.6543157600075045
steps: 327	lr: 0.0001	evaluation reward: 4.05	

episode: 1324	score: 4.0	memory length: 274884	epsilon: 0.6537277000075172
steps: 297	lr: 0.0001	evaluation reward: 4.07	
episode: 1325	score: 7.0	memory length: 275250	epsilon: 0.653003020007533
steps: 366	lr: 0.0001	evaluation reward: 4.11	
episode: 1326	score: 3.0	memory length: 275480	epsilon: 0.6525476200075429
steps: 230	lr: 0.0001	evaluation reward: 4.09	
episode: 1327	score: 2.0	memory length: 275677	epsilon: 0.6521575600075513
steps: 197	lr: 0.0001	evaluation reward: 4.06	
episode: 1328	score: 4.0	memory length: 275936	epsilon: 0.6516447400075625
steps: 259	lr: 0.0001	evaluation reward: 4.07	
episode: 1329	score: 5.0	memory length: 276264	epsilon: 0.6509953000075765
steps: 328	lr: 0.0001	evaluation reward: 4.07	
episode: 1330	score: 4.0	memory length: 276560	epsilon: 0.6504092200075893
steps: 296	lr: 0.0001	evaluation reward: 4.09	
episode: 1331	score: 2.0	memory length: 276760	epsilon: 0.6500132200075979
steps: 200	lr: 0.0001	evaluation reward: 4.09	
episode: 1332	score: 3.0	memory length: 276986	epsilon: 0.6495657400076076
steps: 226	lr: 0.0001	evaluation reward: 4.07	
episode: 1333	score: 1.0	memory length: 277137	epsilon: 0.6492667600076141
steps: 151	lr: 0.0001	evaluation reward: 4.06	
episode: 1334	score: 3.0	memory length: 277385	epsilon: 0.6487757200076247
steps: 248	lr: 0.0001	evaluation reward: 4.08	
episode: 1335	score: 7.0	memory length: 277810	epsilon: 0.647934220007643
steps: 425	lr: 0.0001	evaluation reward: 4.09	
episode: 1336	score: 2.0	memory length: 277991	epsilon: 0.6475758400076508
steps: 181	lr: 0.0001	evaluation reward: 4.08	
episode: 1337	score: 4.0	memory length: 278285	epsilon: 0.6469937200076634
steps: 294	lr: 0.0001	evaluation reward: 4.08	
episode: 1338	score: 5.0	memory length: 278611	epsilon: 0.6463482400076774
steps: 326	lr: 0.0001	evaluation reward: 4.09	
episode: 1339	score: 6.0	memory length: 278971	epsilon: 0.6456354400076929
steps: 360	lr: 0.0001	evaluation reward: 4.1	
episode: 1340	score: 9.0	memory length: 279330	epsilon: 0.6449246200077083
steps: 359	lr: 0.0001	evaluation reward: 4.15	
episode: 1341	score: 7.0	memory length: 279758	epsilon: 0.6440771800077267
steps: 428	lr: 0.0001	evaluation reward: 4.18	
episode: 1342	score: 5.0	memory length: 280044	epsilon: 0.643510900007739
steps: 286	lr: 0.0001	evaluation reward: 4.19	
episode: 1343	score: 6.0	memory length: 280400	epsilon: 0.6428060200077543
steps: 356	lr: 0.0001	evaluation reward: 4.22	
episode: 1344	score: 4.0	memory length: 280673	epsilon: 0.6422654800077661
steps: 273	lr: 0.0001	evaluation reward: 4.21	
episode: 1345	score: 7.0	memory length: 281058	epsilon: 0.6415031800077826
steps: 385	lr: 0.0001	evaluation reward: 4.26	
episode: 1346	score: 3.0	memory length: 281290	epsilon: 0.6410438200077926
steps: 232	lr: 0.0001	evaluation reward: 4.25	
episode: 1347	score: 5.0	memory length: 281596	epsilon: 0.6404379400078057
steps: 306	lr: 0.0001	evaluation reward: 4.26	

episode: 1348	score: 4.0	memory length: 281889	epsilon: 0.6398578000078183
steps: 293	lr: 0.0001	evaluation reward: 4.22	
episode: 1349	score: 7.0	memory length: 282278	epsilon: 0.639087580007835
steps: 389	lr: 0.0001	evaluation reward: 4.24	
episode: 1350	score: 6.0	memory length: 282651	epsilon: 0.6383490400078511
steps: 373	lr: 0.0001	evaluation reward: 4.27	
episode: 1351	score: 3.0	memory length: 282878	epsilon: 0.6378995800078608
steps: 227	lr: 0.0001	evaluation reward: 4.3	
episode: 1352	score: 3.0	memory length: 283106	epsilon: 0.6374481400078706
steps: 228	lr: 0.0001	evaluation reward: 4.27	
episode: 1353	score: 4.0	memory length: 283385	epsilon: 0.6368957200078826
steps: 279	lr: 0.0001	evaluation reward: 4.26	
episode: 1354	score: 6.0	memory length: 283756	epsilon: 0.6361611400078986
steps: 371	lr: 0.0001	evaluation reward: 4.24	
episode: 1355	score: 3.0	memory length: 283986	epsilon: 0.6357057400079085
steps: 230	lr: 0.0001	evaluation reward: 4.25	
episode: 1356	score: 4.0	memory length: 284279	epsilon: 0.6351256000079211
steps: 293	lr: 0.0001	evaluation reward: 4.27	
episode: 1357	score: 2.0	memory length: 284495	epsilon: 0.6346979200079303
steps: 216	lr: 0.0001	evaluation reward: 4.22	
episode: 1358	score: 4.0	memory length: 284788	epsilon: 0.6341177800079429
steps: 293	lr: 0.0001	evaluation reward: 4.21	
episode: 1359	score: 5.0	memory length: 285128	epsilon: 0.6334445800079576
steps: 340	lr: 0.0001	evaluation reward: 4.22	
episode: 1360	score: 4.0	memory length: 285403	epsilon: 0.6329000800079694
steps: 275	lr: 0.0001	evaluation reward: 4.22	
episode: 1361	score: 3.0	memory length: 285617	epsilon: 0.6324763600079786
steps: 214	lr: 0.0001	evaluation reward: 4.15	
episode: 1362	score: 1.0	memory length: 285768	epsilon: 0.6321773800079851
steps: 151	lr: 0.0001	evaluation reward: 4.15	
episode: 1363	score: 4.0	memory length: 286031	epsilon: 0.6316566400079964
steps: 263	lr: 0.0001	evaluation reward: 4.14	
episode: 1364	score: 3.0	memory length: 286257	epsilon: 0.6312091600080061
steps: 226	lr: 0.0001	evaluation reward: 4.15	
episode: 1365	score: 6.0	memory length: 286621	epsilon: 0.6304884400080217
steps: 364	lr: 0.0001	evaluation reward: 4.12	
episode: 1366	score: 4.0	memory length: 286878	epsilon: 0.6299795800080328
steps: 257	lr: 0.0001	evaluation reward: 4.11	
episode: 1367	score: 6.0	memory length: 287222	epsilon: 0.6292984600080476
steps: 344	lr: 0.0001	evaluation reward: 4.16	
episode: 1368	score: 2.0	memory length: 287442	epsilon: 0.628862860008057
steps: 220	lr: 0.0001	evaluation reward: 4.16	
episode: 1369	score: 3.0	memory length: 287654	epsilon: 0.6284431000080661
steps: 212	lr: 0.0001	evaluation reward: 4.17	
episode: 1370	score: 3.0	memory length: 287879	epsilon: 0.6279976000080758
steps: 225	lr: 0.0001	evaluation reward: 4.18	
episode: 1371	score: 9.0	memory length: 288363	epsilon: 0.6270392800080966
steps: 484	lr: 0.0001	evaluation reward: 4.23	

episode: 1372 score: 4.0 memory length: 288627 epsilon: 0.626516560008108
 steps: 264 lr: 0.0001 evaluation reward: 4.22
 episode: 1373 score: 11.0 memory length: 289125 epsilon:
 0.6255305200081294 steps: 498 lr: 0.0001 evaluation reward: 4.29
 episode: 1374 score: 7.0 memory length: 289532 epsilon: 0.6247246600081469
 steps: 407 lr: 0.0001 evaluation reward: 4.34
 episode: 1375 score: 11.0 memory length: 290037 epsilon:
 0.6237247600081686 steps: 505 lr: 0.0001 evaluation reward: 4.43
 episode: 1376 score: 3.0 memory length: 290248 epsilon: 0.6233069800081776
 steps: 211 lr: 0.0001 evaluation reward: 4.43
 episode: 1377 score: 4.0 memory length: 290544 epsilon: 0.6227209000081904
 steps: 296 lr: 0.0001 evaluation reward: 4.43
 episode: 1378 score: 3.0 memory length: 290794 epsilon: 0.6222259000082011
 steps: 250 lr: 0.0001 evaluation reward: 4.45
 episode: 1379 score: 4.0 memory length: 291072 epsilon: 0.621675460008213
 steps: 278 lr: 0.0001 evaluation reward: 4.46
 episode: 1380 score: 7.0 memory length: 291468 epsilon: 0.6208913800082301
 steps: 396 lr: 0.0001 evaluation reward: 4.5
 episode: 1381 score: 7.0 memory length: 291887 epsilon: 0.6200617600082481
 steps: 419 lr: 0.0001 evaluation reward: 4.52
 episode: 1382 score: 4.0 memory length: 292166 epsilon: 0.6195093400082601
 steps: 279 lr: 0.0001 evaluation reward: 4.49
 episode: 1383 score: 5.0 memory length: 292488 epsilon: 0.6188717800082739
 steps: 322 lr: 0.0001 evaluation reward: 4.51
 episode: 1384 score: 5.0 memory length: 292794 epsilon: 0.6182659000082871
 steps: 306 lr: 0.0001 evaluation reward: 4.53
 episode: 1385 score: 2.0 memory length: 292992 epsilon: 0.6178738600082956
 steps: 198 lr: 0.0001 evaluation reward: 4.48
 episode: 1386 score: 4.0 memory length: 293254 epsilon: 0.6173551000083068
 steps: 262 lr: 0.0001 evaluation reward: 4.48
 episode: 1387 score: 5.0 memory length: 293558 epsilon: 0.6167531800083199
 steps: 304 lr: 0.0001 evaluation reward: 4.5
 episode: 1388 score: 7.0 memory length: 293982 epsilon: 0.6159136600083381
 steps: 424 lr: 0.0001 evaluation reward: 4.51
 episode: 1389 score: 5.0 memory length: 294307 epsilon: 0.6152701600083521
 steps: 325 lr: 0.0001 evaluation reward: 4.49
 episode: 1390 score: 7.0 memory length: 294666 epsilon: 0.6145593400083675
 steps: 359 lr: 0.0001 evaluation reward: 4.53
 episode: 1391 score: 6.0 memory length: 294986 epsilon: 0.6139257400083813
 steps: 320 lr: 0.0001 evaluation reward: 4.57
 episode: 1392 score: 4.0 memory length: 295282 epsilon: 0.613339660008394
 steps: 296 lr: 0.0001 evaluation reward: 4.58
 episode: 1393 score: 3.0 memory length: 295493 epsilon: 0.6129218800084031
 steps: 211 lr: 0.0001 evaluation reward: 4.59
 episode: 1394 score: 2.0 memory length: 295675 epsilon: 0.6125615200084109
 steps: 182 lr: 0.0001 evaluation reward: 4.57
 episode: 1395 score: 7.0 memory length: 296070 epsilon: 0.6117794200084279
 steps: 395 lr: 0.0001 evaluation reward: 4.6

episode: 1396 score: 5.0 memory length: 296396 epsilon: 0.6111339400084419
 steps: 326 lr: 0.0001 evaluation reward: 4.55
 episode: 1397 score: 5.0 memory length: 296722 epsilon: 0.6104884600084559
 steps: 326 lr: 0.0001 evaluation reward: 4.57
 episode: 1398 score: 7.0 memory length: 297128 epsilon: 0.6096845800084734
 steps: 406 lr: 0.0001 evaluation reward: 4.59
 episode: 1399 score: 12.0 memory length: 297768 epsilon:
 0.6084173800085009 steps: 640 lr: 0.0001 evaluation reward: 4.67
 episode: 1400 score: 11.0 memory length: 298330 epsilon: 0.607304620008525
 steps: 562 lr: 0.0001 evaluation reward: 4.72
 episode: 1401 score: 6.0 memory length: 298708 epsilon: 0.6065561800085413
 steps: 378 lr: 0.0001 evaluation reward: 4.73
 episode: 1402 score: 4.0 memory length: 298984 epsilon: 0.6060097000085531
 steps: 276 lr: 0.0001 evaluation reward: 4.72
 episode: 1403 score: 8.0 memory length: 299405 epsilon: 0.6051761200085712
 steps: 421 lr: 0.0001 evaluation reward: 4.77
 episode: 1404 score: 3.0 memory length: 299634 epsilon: 0.6047227000085811
 steps: 229 lr: 0.0001 evaluation reward: 4.75
 episode: 1405 score: 5.0 memory length: 299909 epsilon: 0.6041782000085929
 steps: 275 lr: 0.0001 evaluation reward: 4.79
 episode: 1406 score: 4.0 memory length: 300184 epsilon: 0.6036337000086047
 steps: 275 lr: 0.0001 evaluation reward: 4.79
 episode: 1407 score: 5.0 memory length: 300514 epsilon: 0.6029803000086189
 steps: 330 lr: 0.0001 evaluation reward: 4.8
 episode: 1408 score: 4.0 memory length: 300772 epsilon: 0.60246946000863
 steps: 258 lr: 0.0001 evaluation reward: 4.79
 episode: 1409 score: 5.0 memory length: 301080 epsilon: 0.6018596200086432
 steps: 308 lr: 0.0001 evaluation reward: 4.81
 episode: 1410 score: 5.0 memory length: 301384 epsilon: 0.6012577000086563
 steps: 304 lr: 0.0001 evaluation reward: 4.84
 episode: 1411 score: 3.0 memory length: 301654 epsilon: 0.6007231000086679
 steps: 270 lr: 0.0001 evaluation reward: 4.83
 episode: 1412 score: 11.0 memory length: 302078 epsilon:
 0.5998835800086861 steps: 424 lr: 0.0001 evaluation reward: 4.89
 episode: 1413 score: 6.0 memory length: 302452 epsilon: 0.5991430600087022
 steps: 374 lr: 0.0001 evaluation reward: 4.92
 episode: 1414 score: 4.0 memory length: 302744 epsilon: 0.5985649000087148
 steps: 292 lr: 0.0001 evaluation reward: 4.91
 episode: 1415 score: 9.0 memory length: 303224 epsilon: 0.5976145000087354
 steps: 480 lr: 0.0001 evaluation reward: 4.99
 episode: 1416 score: 4.0 memory length: 303466 epsilon: 0.5971353400087458
 steps: 242 lr: 0.0001 evaluation reward: 4.97
 episode: 1417 score: 5.0 memory length: 303792 epsilon: 0.5964898600087598
 steps: 326 lr: 0.0001 evaluation reward: 4.97
 episode: 1418 score: 1.0 memory length: 303943 epsilon: 0.5961908800087663
 steps: 151 lr: 0.0001 evaluation reward: 4.93
 episode: 1419 score: 8.0 memory length: 304348 epsilon: 0.5953889800087837
 steps: 405 lr: 0.0001 evaluation reward: 4.93

episode: 1420	score: 6.0	memory length: 304692	epsilon: 0.5947078600087985
steps: 344	lr: 0.0001	evaluation reward: 4.96	
episode: 1421	score: 7.0	memory length: 305133	epsilon: 0.5938346800088175
steps: 441	lr: 0.0001	evaluation reward: 4.99	
episode: 1422	score: 7.0	memory length: 305513	epsilon: 0.5930822800088338
steps: 380	lr: 0.0001	evaluation reward: 4.99	
episode: 1423	score: 9.0	memory length: 305990	epsilon: 0.5921378200088543
steps: 477	lr: 0.0001	evaluation reward: 5.03	
episode: 1424	score: 6.0	memory length: 306308	epsilon: 0.591508180008868
steps: 318	lr: 0.0001	evaluation reward: 5.05	
episode: 1425	score: 3.0	memory length: 306521	epsilon: 0.5910864400088771
steps: 213	lr: 0.0001	evaluation reward: 5.01	
episode: 1426	score: 4.0	memory length: 306763	epsilon: 0.5906072800088875
steps: 242	lr: 0.0001	evaluation reward: 5.02	
episode: 1427	score: 9.0	memory length: 307110	epsilon: 0.5899202200089024
steps: 347	lr: 0.0001	evaluation reward: 5.09	
episode: 1428	score: 9.0	memory length: 307607	epsilon: 0.5889361600089238
steps: 497	lr: 0.0001	evaluation reward: 5.14	
episode: 1429	score: 6.0	memory length: 307928	epsilon: 0.5883005800089376
steps: 321	lr: 0.0001	evaluation reward: 5.15	
episode: 1430	score: 8.0	memory length: 308405	epsilon: 0.5873561200089581
steps: 477	lr: 0.0001	evaluation reward: 5.19	
episode: 1431	score: 5.0	memory length: 308710	epsilon: 0.5867522200089712
steps: 305	lr: 0.0001	evaluation reward: 5.22	
episode: 1432	score: 10.0	memory length: 309191	epsilon:
0.5857998400089919	steps: 481	lr: 0.0001	evaluation reward: 5.29
episode: 1433	score: 8.0	memory length: 309638	epsilon: 0.5849147800090111
steps: 447	lr: 0.0001	evaluation reward: 5.36	
episode: 1434	score: 5.0	memory length: 309970	epsilon: 0.5842574200090254
steps: 332	lr: 0.0001	evaluation reward: 5.38	
episode: 1435	score: 8.0	memory length: 310414	epsilon: 0.5833783000090444
steps: 444	lr: 0.0001	evaluation reward: 5.39	
episode: 1436	score: 5.0	memory length: 310740	epsilon: 0.5827328200090585
steps: 326	lr: 0.0001	evaluation reward: 5.42	
episode: 1437	score: 5.0	memory length: 311024	epsilon: 0.5821705000090707
steps: 284	lr: 0.0001	evaluation reward: 5.43	
episode: 1438	score: 8.0	memory length: 311497	epsilon: 0.581233960009091
steps: 473	lr: 0.0001	evaluation reward: 5.46	
episode: 1439	score: 7.0	memory length: 311854	epsilon: 0.5805271000091063
steps: 357	lr: 0.0001	evaluation reward: 5.47	
episode: 1440	score: 6.0	memory length: 312212	epsilon: 0.5798182600091217
steps: 358	lr: 0.0001	evaluation reward: 5.44	
episode: 1441	score: 3.0	memory length: 312478	epsilon: 0.5792915800091332
steps: 266	lr: 0.0001	evaluation reward: 5.4	
episode: 1442	score: 8.0	memory length: 312902	epsilon: 0.5784520600091514
steps: 424	lr: 0.0001	evaluation reward: 5.43	
episode: 1443	score: 3.0	memory length: 313133	epsilon: 0.5779946800091613
steps: 231	lr: 0.0001	evaluation reward: 5.4	

episode: 1444	score: 9.0	memory length: 313618	epsilon: 0.5770343800091822
steps: 485	lr: 0.0001	evaluation reward: 5.45	
episode: 1445	score: 8.0	memory length: 314058	epsilon: 0.5761631800092011
steps: 440	lr: 0.0001	evaluation reward: 5.46	
episode: 1446	score: 7.0	memory length: 314432	epsilon: 0.5754226600092172
steps: 374	lr: 0.0001	evaluation reward: 5.5	
episode: 1447	score: 6.0	memory length: 314803	epsilon: 0.5746880800092331
steps: 371	lr: 0.0001	evaluation reward: 5.51	
episode: 1448	score: 4.0	memory length: 315095	epsilon: 0.5741099200092457
steps: 292	lr: 0.0001	evaluation reward: 5.51	
episode: 1449	score: 7.0	memory length: 315537	epsilon: 0.5732347600092647
steps: 442	lr: 0.0001	evaluation reward: 5.51	
episode: 1450	score: 6.0	memory length: 315911	epsilon: 0.5724942400092807
steps: 374	lr: 0.0001	evaluation reward: 5.51	
episode: 1451	score: 2.0	memory length: 316109	epsilon: 0.5721022000092892
steps: 198	lr: 0.0001	evaluation reward: 5.5	
episode: 1452	score: 8.0	memory length: 316512	epsilon: 0.5713042600093066
steps: 403	lr: 0.0001	evaluation reward: 5.55	
episode: 1453	score: 8.0	memory length: 316957	epsilon: 0.5704231600093257
steps: 445	lr: 0.0001	evaluation reward: 5.59	
episode: 1454	score: 7.0	memory length: 317349	epsilon: 0.5696470000093425
steps: 392	lr: 0.0001	evaluation reward: 5.6	
episode: 1455	score: 4.0	memory length: 317593	epsilon: 0.569163880009353
steps: 244	lr: 0.0001	evaluation reward: 5.61	
episode: 1456	score: 5.0	memory length: 317900	epsilon: 0.5685560200093662
steps: 307	lr: 0.0001	evaluation reward: 5.62	
episode: 1457	score: 6.0	memory length: 318272	epsilon: 0.5678194600093822
steps: 372	lr: 0.0001	evaluation reward: 5.66	
episode: 1458	score: 11.0	memory length: 318744	epsilon:
0.5668849000094025	steps: 472	lr: 0.0001	evaluation reward: 5.73
episode: 1459	score: 4.0	memory length: 319016	epsilon: 0.5663463400094142
steps: 272	lr: 0.0001	evaluation reward: 5.72	
episode: 1460	score: 5.0	memory length: 319339	epsilon: 0.5657068000094281
steps: 323	lr: 0.0001	evaluation reward: 5.73	
episode: 1461	score: 2.0	memory length: 319523	epsilon: 0.565342480009436
steps: 184	lr: 0.0001	evaluation reward: 5.72	
episode: 1462	score: 4.0	memory length: 319837	epsilon: 0.5647207600094495
steps: 314	lr: 0.0001	evaluation reward: 5.75	
episode: 1463	score: 6.0	memory length: 320229	epsilon: 0.5639446000094663
steps: 392	lr: 0.0001	evaluation reward: 5.77	
episode: 1464	score: 4.0	memory length: 320490	epsilon: 0.5634278200094776
steps: 261	lr: 0.0001	evaluation reward: 5.78	
episode: 1465	score: 6.0	memory length: 320829	epsilon: 0.5627566000094921
steps: 339	lr: 0.0001	evaluation reward: 5.78	
episode: 1466	score: 7.0	memory length: 321218	epsilon: 0.5619863800095088
steps: 389	lr: 0.0001	evaluation reward: 5.81	
episode: 1467	score: 4.0	memory length: 321459	epsilon: 0.5615092000095192
steps: 241	lr: 0.0001	evaluation reward: 5.79	

episode: 1468 score: 7.0 memory length: 321874 epsilon: 0.560687500009537
 steps: 415 lr: 0.0001 evaluation reward: 5.84
 episode: 1469 score: 7.0 memory length: 322276 epsilon: 0.5598915400095543
 steps: 402 lr: 0.0001 evaluation reward: 5.88
 episode: 1470 score: 7.0 memory length: 322664 epsilon: 0.559123300009571
 steps: 388 lr: 0.0001 evaluation reward: 5.92
 episode: 1471 score: 5.0 memory length: 322949 epsilon: 0.5585590000095833
 steps: 285 lr: 0.0001 evaluation reward: 5.88
 episode: 1472 score: 16.0 memory length: 323564 epsilon:
 0.5573413000096097 steps: 615 lr: 0.0001 evaluation reward: 6.0
 episode: 1473 score: 3.0 memory length: 323814 epsilon: 0.5568463000096204
 steps: 250 lr: 0.0001 evaluation reward: 5.92
 episode: 1474 score: 6.0 memory length: 324154 epsilon: 0.556173100009635
 steps: 340 lr: 0.0001 evaluation reward: 5.91
 episode: 1475 score: 6.0 memory length: 324464 epsilon: 0.5555593000096484
 steps: 310 lr: 0.0001 evaluation reward: 5.86
 episode: 1476 score: 8.0 memory length: 324918 epsilon: 0.5546603800096679
 steps: 454 lr: 0.0001 evaluation reward: 5.91
 episode: 1477 score: 7.0 memory length: 325310 epsilon: 0.5538842200096847
 steps: 392 lr: 0.0001 evaluation reward: 5.94
 episode: 1478 score: 10.0 memory length: 325837 epsilon:
 0.5528407600097074 steps: 527 lr: 0.0001 evaluation reward: 6.01
 episode: 1479 score: 7.0 memory length: 326229 epsilon: 0.5520646000097242
 steps: 392 lr: 0.0001 evaluation reward: 6.04
 episode: 1480 score: 4.0 memory length: 326503 epsilon: 0.551522080009736
 steps: 274 lr: 0.0001 evaluation reward: 6.01
 episode: 1481 score: 6.0 memory length: 326841 epsilon: 0.5508528400097505
 steps: 338 lr: 0.0001 evaluation reward: 6.0
 episode: 1482 score: 11.0 memory length: 327405 epsilon:
 0.5497361200097748 steps: 564 lr: 0.0001 evaluation reward: 6.07
 episode: 1483 score: 5.0 memory length: 327734 epsilon: 0.5490847000097889
 steps: 329 lr: 0.0001 evaluation reward: 6.07
 episode: 1484 score: 6.0 memory length: 328119 epsilon: 0.5483224000098055
 steps: 385 lr: 0.0001 evaluation reward: 6.08
 episode: 1485 score: 6.0 memory length: 328495 epsilon: 0.5475779200098216
 steps: 376 lr: 0.0001 evaluation reward: 6.12
 episode: 1486 score: 6.0 memory length: 328849 epsilon: 0.5468770000098369
 steps: 354 lr: 0.0001 evaluation reward: 6.14
 episode: 1487 score: 7.0 memory length: 329241 epsilon: 0.5461008400098537
 steps: 392 lr: 0.0001 evaluation reward: 6.16
 episode: 1488 score: 3.0 memory length: 329489 epsilon: 0.5456098000098644
 steps: 248 lr: 0.0001 evaluation reward: 6.12
 episode: 1489 score: 5.0 memory length: 329793 epsilon: 0.5450078800098774
 steps: 304 lr: 0.0001 evaluation reward: 6.12
 episode: 1490 score: 10.0 memory length: 330310 epsilon:
 0.5439842200098997 steps: 517 lr: 0.0001 evaluation reward: 6.15
 episode: 1491 score: 7.0 memory length: 330712 epsilon: 0.5431882600099169
 steps: 402 lr: 0.0001 evaluation reward: 6.16

episode: 1492	score: 7.0	memory length: 331119	epsilon: 0.5423824000099344
steps: 407	lr: 0.0001	evaluation reward: 6.19	
episode: 1493	score: 6.0	memory length: 331494	epsilon: 0.5416399000099505
steps: 375	lr: 0.0001	evaluation reward: 6.22	
episode: 1494	score: 3.0	memory length: 331722	epsilon: 0.5411884600099603
steps: 228	lr: 0.0001	evaluation reward: 6.23	
episode: 1495	score: 5.0	memory length: 332013	epsilon: 0.5406122800099729
steps: 291	lr: 0.0001	evaluation reward: 6.21	
episode: 1496	score: 2.0	memory length: 332211	epsilon: 0.5402202400099814
steps: 198	lr: 0.0001	evaluation reward: 6.18	
episode: 1497	score: 9.0	memory length: 332637	epsilon: 0.5393767600099997
steps: 426	lr: 0.0001	evaluation reward: 6.22	
episode: 1498	score: 9.0	memory length: 333109	epsilon: 0.53844220001002
steps: 472	lr: 0.0001	evaluation reward: 6.24	
episode: 1499	score: 5.0	memory length: 333433	epsilon: 0.5378006800100339
steps: 324	lr: 0.0001	evaluation reward: 6.17	
episode: 1500	score: 8.0	memory length: 333886	epsilon: 0.5369037400100534
steps: 453	lr: 0.0001	evaluation reward: 6.14	
episode: 1501	score: 8.0	memory length: 334296	epsilon: 0.536091940010071
steps: 410	lr: 0.0001	evaluation reward: 6.16	
episode: 1502	score: 9.0	memory length: 334747	epsilon: 0.5351989600100904
steps: 451	lr: 0.0001	evaluation reward: 6.21	
episode: 1503	score: 9.0	memory length: 335207	epsilon: 0.5342881600101101
steps: 460	lr: 0.0001	evaluation reward: 6.22	
episode: 1504	score: 3.0	memory length: 335433	epsilon: 0.5338406800101199
steps: 226	lr: 0.0001	evaluation reward: 6.22	
episode: 1505	score: 8.0	memory length: 335880	epsilon: 0.5329556200101391
steps: 447	lr: 0.0001	evaluation reward: 6.25	
episode: 1506	score: 8.0	memory length: 336374	epsilon: 0.5319775000101603
steps: 494	lr: 0.0001	evaluation reward: 6.29	
episode: 1507	score: 5.0	memory length: 336706	epsilon: 0.5313201400101746
steps: 332	lr: 0.0001	evaluation reward: 6.29	
episode: 1508	score: 6.0	memory length: 337085	epsilon: 0.5305697200101909
steps: 379	lr: 0.0001	evaluation reward: 6.31	
episode: 1509	score: 9.0	memory length: 337555	epsilon: 0.5296391200102111
steps: 470	lr: 0.0001	evaluation reward: 6.35	
episode: 1510	score: 3.0	memory length: 337784	epsilon: 0.5291857000102209
steps: 229	lr: 0.0001	evaluation reward: 6.33	
episode: 1511	score: 5.0	memory length: 338108	epsilon: 0.5285441800102348
steps: 324	lr: 0.0001	evaluation reward: 6.35	
episode: 1512	score: 2.0	memory length: 338307	epsilon: 0.5281501600102434
steps: 199	lr: 0.0001	evaluation reward: 6.26	
episode: 1513	score: 11.0	memory length: 338892	epsilon:
0.5269918600102685	steps: 585	lr: 0.0001	evaluation reward: 6.31
episode: 1514	score: 6.0	memory length: 339263	epsilon: 0.5262572800102845
steps: 371	lr: 0.0001	evaluation reward: 6.33	
episode: 1515	score: 5.0	memory length: 339601	epsilon: 0.525588040010299
steps: 338	lr: 0.0001	evaluation reward: 6.29	

episode: 1516 score: 4.0 memory length: 339863 epsilon: 0.5250692800103103
 steps: 262 lr: 0.0001 evaluation reward: 6.29
 episode: 1517 score: 5.0 memory length: 340206 epsilon: 0.524390140010325
 steps: 343 lr: 0.0001 evaluation reward: 6.29
 episode: 1518 score: 7.0 memory length: 340667 epsilon: 0.5234773600103448
 steps: 461 lr: 0.0001 evaluation reward: 6.35
 episode: 1519 score: 7.0 memory length: 341094 epsilon: 0.5226319000103632
 steps: 427 lr: 0.0001 evaluation reward: 6.34
 episode: 1520 score: 14.0 memory length: 341802 epsilon:
 0.5212300600103936 steps: 708 lr: 0.0001 evaluation reward: 6.42
 episode: 1521 score: 8.0 memory length: 342272 epsilon: 0.5202994600104138
 steps: 470 lr: 0.0001 evaluation reward: 6.43
 episode: 1522 score: 1.0 memory length: 342423 epsilon: 0.5200004800104203
 steps: 151 lr: 0.0001 evaluation reward: 6.37
 episode: 1523 score: 6.0 memory length: 342781 epsilon: 0.5192916400104357
 steps: 358 lr: 0.0001 evaluation reward: 6.34
 episode: 1524 score: 9.0 memory length: 343255 epsilon: 0.5183531200104561
 steps: 474 lr: 0.0001 evaluation reward: 6.37
 episode: 1525 score: 8.0 memory length: 343681 epsilon: 0.5175096400104744
 steps: 426 lr: 0.0001 evaluation reward: 6.42
 episode: 1526 score: 3.0 memory length: 343894 epsilon: 0.5170879000104835
 steps: 213 lr: 0.0001 evaluation reward: 6.41
 episode: 1527 score: 7.0 memory length: 344266 epsilon: 0.5163513400104995
 steps: 372 lr: 0.0001 evaluation reward: 6.39
 episode: 1528 score: 4.0 memory length: 344564 epsilon: 0.5157613000105123
 steps: 298 lr: 0.0001 evaluation reward: 6.34
 episode: 1529 score: 9.0 memory length: 345013 epsilon: 0.5148722800105316
 steps: 449 lr: 0.0001 evaluation reward: 6.37
 episode: 1530 score: 6.0 memory length: 345384 epsilon: 0.5141377000105476
 steps: 371 lr: 0.0001 evaluation reward: 6.35
 episode: 1531 score: 13.0 memory length: 345965 epsilon:
 0.5129873200105726 steps: 581 lr: 0.0001 evaluation reward: 6.43
 episode: 1532 score: 5.0 memory length: 346275 epsilon: 0.5123735200105859
 steps: 310 lr: 0.0001 evaluation reward: 6.38
 episode: 1533 score: 10.0 memory length: 346788 epsilon:
 0.5113577800106079 steps: 513 lr: 0.0001 evaluation reward: 6.4
 episode: 1534 score: 9.0 memory length: 347274 epsilon: 0.5103955000106288
 steps: 486 lr: 0.0001 evaluation reward: 6.44
 episode: 1535 score: 8.0 memory length: 347689 epsilon: 0.5095738000106467
 steps: 415 lr: 0.0001 evaluation reward: 6.44
 episode: 1536 score: 11.0 memory length: 348238 epsilon:
 0.5084867800106703 steps: 549 lr: 0.0001 evaluation reward: 6.5
 episode: 1537 score: 5.0 memory length: 348565 epsilon: 0.5078393200106843
 steps: 327 lr: 0.0001 evaluation reward: 6.5
 episode: 1538 score: 8.0 memory length: 349019 epsilon: 0.5069404000107038
 steps: 454 lr: 0.0001 evaluation reward: 6.5
 episode: 1539 score: 6.0 memory length: 349394 epsilon: 0.50619790001072
 steps: 375 lr: 0.0001 evaluation reward: 6.49

episode: 1540 score: 9.0 memory length: 349886 epsilon: 0.5052237400107411
 steps: 492 lr: 0.0001 evaluation reward: 6.52
 episode: 1541 score: 3.0 memory length: 350133 epsilon: 0.5047346800107517
 steps: 247 lr: 0.0001 evaluation reward: 6.52
 episode: 1542 score: 7.0 memory length: 350519 epsilon: 0.5039704000107683
 steps: 386 lr: 0.0001 evaluation reward: 6.51
 episode: 1543 score: 8.0 memory length: 350953 epsilon: 0.503111080010787
 steps: 434 lr: 0.0001 evaluation reward: 6.56
 episode: 1544 score: 3.0 memory length: 351219 epsilon: 0.5025844000107984
 steps: 266 lr: 0.0001 evaluation reward: 6.5
 episode: 1545 score: 7.0 memory length: 351625 epsilon: 0.5017805200108159
 steps: 406 lr: 0.0001 evaluation reward: 6.49
 episode: 1546 score: 13.0 memory length: 352132 epsilon:
 0.5007766600108376 steps: 507 lr: 0.0001 evaluation reward: 6.55
 episode: 1547 score: 7.0 memory length: 352536 epsilon:
 0.49997674001085435 steps: 404 lr: 0.0001 evaluation reward: 6.56
 episode: 1548 score: 4.0 memory length: 352795 epsilon: 0.4994639200108511
 steps: 259 lr: 0.0001 evaluation reward: 6.56
 episode: 1549 score: 8.0 memory length: 353235 epsilon: 0.4985927200108456
 steps: 440 lr: 0.0001 evaluation reward: 6.57
 episode: 1550 score: 4.0 memory length: 353474 epsilon: 0.4981195000108426
 steps: 239 lr: 0.0001 evaluation reward: 6.55
 episode: 1551 score: 10.0 memory length: 353989 epsilon:
 0.49709980001083615 steps: 515 lr: 0.0001 evaluation reward: 6.63
 episode: 1552 score: 15.0 memory length: 354586 epsilon:
 0.49591774001082867 steps: 597 lr: 0.0001 evaluation reward: 6.7
 episode: 1553 score: 3.0 memory length: 354833 epsilon: 0.4954286800108256
 steps: 247 lr: 0.0001 evaluation reward: 6.65
 episode: 1554 score: 4.0 memory length: 355109 epsilon: 0.4948822000108221
 steps: 276 lr: 0.0001 evaluation reward: 6.62
 episode: 1555 score: 10.0 memory length: 355642 epsilon:
 0.49382686001081544 steps: 533 lr: 0.0001 evaluation reward: 6.68
 episode: 1556 score: 8.0 memory length: 356086 epsilon: 0.4929477400108099
 steps: 444 lr: 0.0001 evaluation reward: 6.71
 episode: 1557 score: 5.0 memory length: 356410 epsilon: 0.4923062200108058
 steps: 324 lr: 0.0001 evaluation reward: 6.7
 episode: 1558 score: 8.0 memory length: 356836 epsilon: 0.4914627400108005
 steps: 426 lr: 0.0001 evaluation reward: 6.67
 episode: 1559 score: 6.0 memory length: 357192 epsilon: 0.490757860010796
 steps: 356 lr: 0.0001 evaluation reward: 6.69
 episode: 1560 score: 7.0 memory length: 357597 epsilon:
 0.48995596001079095 steps: 405 lr: 0.0001 evaluation reward: 6.71
 episode: 1561 score: 6.0 memory length: 357931 epsilon:
 0.48929464001078676 steps: 334 lr: 0.0001 evaluation reward: 6.75
 episode: 1562 score: 6.0 memory length: 358286 epsilon: 0.4885917400107823
 steps: 355 lr: 0.0001 evaluation reward: 6.77
 episode: 1563 score: 6.0 memory length: 358621 epsilon: 0.4879284400107781
 steps: 335 lr: 0.0001 evaluation reward: 6.77

episode: 1564 score: 7.0 memory length: 359005 epsilon: 0.4871681200107733
 steps: 384 lr: 0.0001 evaluation reward: 6.8
 episode: 1565 score: 12.0 memory length: 359435 epsilon:
 0.4863167200107679 steps: 430 lr: 0.0001 evaluation reward: 6.86
 episode: 1566 score: 8.0 memory length: 359842 epsilon: 0.4855108600107628
 steps: 407 lr: 0.0001 evaluation reward: 6.87
 episode: 1567 score: 6.0 memory length: 360194 epsilon: 0.4848139000107584
 steps: 352 lr: 0.0001 evaluation reward: 6.89
 episode: 1568 score: 12.0 memory length: 360673 epsilon:
 0.4838654800107524 steps: 479 lr: 0.0001 evaluation reward: 6.94
 episode: 1569 score: 4.0 memory length: 360932 epsilon:
 0.48335266001074917 steps: 259 lr: 0.0001 evaluation reward: 6.91
 episode: 1570 score: 8.0 memory length: 361388 epsilon:
 0.48244978001074346 steps: 456 lr: 0.0001 evaluation reward: 6.92
 episode: 1571 score: 11.0 memory length: 361922 epsilon:
 0.48139246001073677 steps: 534 lr: 0.0001 evaluation reward: 6.98
 episode: 1572 score: 14.0 memory length: 362477 epsilon:
 0.4802935600107298 steps: 555 lr: 0.0001 evaluation reward: 6.96
 episode: 1573 score: 3.0 memory length: 362685 epsilon: 0.4798817200107272
 steps: 208 lr: 0.0001 evaluation reward: 6.96
 episode: 1574 score: 2.0 memory length: 362883 epsilon:
 0.47948968001072473 steps: 198 lr: 0.0001 evaluation reward: 6.92
 episode: 1575 score: 5.0 memory length: 363229 epsilon: 0.4788046000107204
 steps: 346 lr: 0.0001 evaluation reward: 6.91
 episode: 1576 score: 3.0 memory length: 363474 epsilon: 0.4783195000107173
 steps: 245 lr: 0.0001 evaluation reward: 6.86
 episode: 1577 score: 11.0 memory length: 363987 epsilon:
 0.4773037600107109 steps: 513 lr: 0.0001 evaluation reward: 6.9
 episode: 1578 score: 6.0 memory length: 364330 epsilon: 0.4766246200107066
 steps: 343 lr: 0.0001 evaluation reward: 6.86
 episode: 1579 score: 4.0 memory length: 364594 epsilon: 0.4761019000107033
 steps: 264 lr: 0.0001 evaluation reward: 6.83
 episode: 1580 score: 8.0 memory length: 365032 epsilon: 0.4752346600106978
 steps: 438 lr: 0.0001 evaluation reward: 6.87
 episode: 1581 score: 4.0 memory length: 365291 epsilon:
 0.47472184001069456 steps: 259 lr: 0.0001 evaluation reward: 6.85
 episode: 1582 score: 13.0 memory length: 365829 epsilon:
 0.4736566000106878 steps: 538 lr: 0.0001 evaluation reward: 6.87
 episode: 1583 score: 7.0 memory length: 366208 epsilon: 0.4729061800106831
 steps: 379 lr: 0.0001 evaluation reward: 6.89
 episode: 1584 score: 3.0 memory length: 366475 epsilon:
 0.47237752001067973 steps: 267 lr: 0.0001 evaluation reward: 6.86
 episode: 1585 score: 7.0 memory length: 366899 epsilon: 0.4715380000106744
 steps: 424 lr: 0.0001 evaluation reward: 6.87
 episode: 1586 score: 4.0 memory length: 367197 epsilon: 0.4709479600106707
 steps: 298 lr: 0.0001 evaluation reward: 6.85
 episode: 1587 score: 6.0 memory length: 367569 epsilon: 0.470211400010666
 steps: 372 lr: 0.0001 evaluation reward: 6.84

episode: 1588 score: 4.0 memory length: 367809 epsilon: 0.469736200010663
 steps: 240 lr: 0.0001 evaluation reward: 6.85
 episode: 1589 score: 9.0 memory length: 368224 epsilon: 0.4689145000106578
 steps: 415 lr: 0.0001 evaluation reward: 6.89
 episode: 1590 score: 7.0 memory length: 368649 epsilon: 0.4680730000106525
 steps: 425 lr: 0.0001 evaluation reward: 6.86
 episode: 1591 score: 5.0 memory length: 368961 epsilon: 0.4674552400106486
 steps: 312 lr: 0.0001 evaluation reward: 6.84
 episode: 1592 score: 3.0 memory length: 369193 epsilon: 0.4669958800106457
 steps: 232 lr: 0.0001 evaluation reward: 6.8
 episode: 1593 score: 6.0 memory length: 369567 epsilon: 0.466255360010641
 steps: 374 lr: 0.0001 evaluation reward: 6.8
 episode: 1594 score: 10.0 memory length: 370104 epsilon:
 0.46519210001063427 steps: 537 lr: 0.0001 evaluation reward: 6.87
 episode: 1595 score: 6.0 memory length: 370425 epsilon:
 0.46455652001063025 steps: 321 lr: 0.0001 evaluation reward: 6.88
 episode: 1596 score: 5.0 memory length: 370756 epsilon: 0.4639011400106261
 steps: 331 lr: 0.0001 evaluation reward: 6.91
 episode: 1597 score: 12.0 memory length: 371237 epsilon:
 0.4629487600106201 steps: 481 lr: 0.0001 evaluation reward: 6.94
 episode: 1598 score: 8.0 memory length: 371672 epsilon: 0.4620874600106146
 steps: 435 lr: 0.0001 evaluation reward: 6.93
 episode: 1599 score: 2.0 memory length: 371853 epsilon:
 0.46172908001061236 steps: 181 lr: 0.0001 evaluation reward: 6.9
 episode: 1600 score: 12.0 memory length: 372455 epsilon:
 0.4605371200106048 steps: 602 lr: 0.0001 evaluation reward: 6.94
 episode: 1601 score: 6.0 memory length: 372781 epsilon:
 0.45989164001060073 steps: 326 lr: 0.0001 evaluation reward: 6.92
 episode: 1602 score: 6.0 memory length: 373138 epsilon:
 0.45918478001059626 steps: 357 lr: 0.0001 evaluation reward: 6.89
 episode: 1603 score: 3.0 memory length: 373371 epsilon:
 0.45872344001059334 steps: 233 lr: 0.0001 evaluation reward: 6.83
 episode: 1604 score: 11.0 memory length: 373918 epsilon:
 0.4576403800105865 steps: 547 lr: 0.0001 evaluation reward: 6.91
 episode: 1605 score: 5.0 memory length: 374226 epsilon:
 0.45703054001058263 steps: 308 lr: 0.0001 evaluation reward: 6.88
 episode: 1606 score: 12.0 memory length: 374699 epsilon:
 0.4560940000105767 steps: 473 lr: 0.0001 evaluation reward: 6.92
 episode: 1607 score: 20.0 memory length: 375413 epsilon:
 0.45468028001056776 steps: 714 lr: 0.0001 evaluation reward: 7.07
 episode: 1608 score: 12.0 memory length: 375814 epsilon:
 0.45388630001056274 steps: 401 lr: 0.0001 evaluation reward: 7.13
 episode: 1609 score: 7.0 memory length: 376190 epsilon: 0.453141820010558
 steps: 376 lr: 0.0001 evaluation reward: 7.11
 episode: 1610 score: 4.0 memory length: 376432 epsilon: 0.452662660010555
 steps: 242 lr: 0.0001 evaluation reward: 7.12
 episode: 1611 score: 4.0 memory length: 376727 epsilon: 0.4520785600105513
 steps: 295 lr: 0.0001 evaluation reward: 7.11

episode: 1612 score: 7.0 memory length: 377135 epsilon: 0.4512707200105462
 steps: 408 lr: 0.0001 evaluation reward: 7.16
 episode: 1613 score: 8.0 memory length: 377570 epsilon:
 0.45040942001054074 steps: 435 lr: 0.0001 evaluation reward: 7.13
 episode: 1614 score: 9.0 memory length: 378042 epsilon: 0.4494748600105348
 steps: 472 lr: 0.0001 evaluation reward: 7.16
 episode: 1615 score: 4.0 memory length: 378316 epsilon: 0.4489323400105314
 steps: 274 lr: 0.0001 evaluation reward: 7.15
 episode: 1616 score: 6.0 memory length: 378675 epsilon: 0.4482215200105269
 steps: 359 lr: 0.0001 evaluation reward: 7.17
 episode: 1617 score: 7.0 memory length: 379035 epsilon: 0.4475087200105224
 steps: 360 lr: 0.0001 evaluation reward: 7.19
 episode: 1618 score: 6.0 memory length: 379375 epsilon:
 0.44683552001051813 steps: 340 lr: 0.0001 evaluation reward: 7.18
 episode: 1619 score: 11.0 memory length: 379868 epsilon:
 0.44585938001051195 steps: 493 lr: 0.0001 evaluation reward: 7.22
 episode: 1620 score: 6.0 memory length: 380208 epsilon: 0.4451861800105077
 steps: 340 lr: 0.0001 evaluation reward: 7.14
 episode: 1621 score: 6.0 memory length: 380570 epsilon:
 0.44446942001050316 steps: 362 lr: 0.0001 evaluation reward: 7.12
 episode: 1622 score: 7.0 memory length: 380944 epsilon: 0.4437289000104985
 steps: 374 lr: 0.0001 evaluation reward: 7.18
 episode: 1623 score: 10.0 memory length: 381445 epsilon:
 0.4427369200104922 steps: 501 lr: 0.0001 evaluation reward: 7.22
 episode: 1624 score: 8.0 memory length: 381890 epsilon: 0.4418558200104866
 steps: 445 lr: 0.0001 evaluation reward: 7.21
 episode: 1625 score: 12.0 memory length: 382536 epsilon:
 0.44057674001047853 steps: 646 lr: 0.0001 evaluation reward: 7.25
 episode: 1626 score: 4.0 memory length: 382812 epsilon:
 0.44003026001047507 steps: 276 lr: 0.0001 evaluation reward: 7.26
 episode: 1627 score: 9.0 memory length: 383328 epsilon: 0.4390085800104686
 steps: 516 lr: 0.0001 evaluation reward: 7.28
 episode: 1628 score: 10.0 memory length: 383834 epsilon:
 0.43800670001046227 steps: 506 lr: 0.0001 evaluation reward: 7.34
 episode: 1629 score: 9.0 memory length: 384310 epsilon: 0.4370642200104563
 steps: 476 lr: 0.0001 evaluation reward: 7.34
 episode: 1630 score: 7.0 memory length: 384701 epsilon: 0.4362900400104514
 steps: 391 lr: 0.0001 evaluation reward: 7.35
 episode: 1631 score: 13.0 memory length: 385162 epsilon:
 0.43537726001044563 steps: 461 lr: 0.0001 evaluation reward: 7.35
 episode: 1632 score: 5.0 memory length: 385457 epsilon:
 0.43479316001044194 steps: 295 lr: 0.0001 evaluation reward: 7.35
 episode: 1633 score: 5.0 memory length: 385803 epsilon: 0.4341080800104376
 steps: 346 lr: 0.0001 evaluation reward: 7.3
 episode: 1634 score: 5.0 memory length: 386148 epsilon: 0.4334249800104333
 steps: 345 lr: 0.0001 evaluation reward: 7.26
 episode: 1635 score: 6.0 memory length: 386520 epsilon: 0.4326884200104286
 steps: 372 lr: 0.0001 evaluation reward: 7.24

episode: 1636 score: 7.0 memory length: 386946 epsilon: 0.4318449400104233
steps: 426 lr: 0.0001 evaluation reward: 7.2
episode: 1637 score: 5.0 memory length: 387238 epsilon: 0.4312667800104196
steps: 292 lr: 0.0001 evaluation reward: 7.2
episode: 1638 score: 10.0 memory length: 387761 epsilon:
0.4302312400104131 steps: 523 lr: 0.0001 evaluation reward: 7.22
episode: 1639 score: 4.0 memory length: 388039 epsilon: 0.4296808000104096
steps: 278 lr: 0.0001 evaluation reward: 7.2
episode: 1640 score: 6.0 memory length: 388390 epsilon: 0.4289858200104052
steps: 351 lr: 0.0001 evaluation reward: 7.17
episode: 1641 score: 21.0 memory length: 388926 epsilon:
0.4279245400103985 steps: 536 lr: 0.0001 evaluation reward: 7.35
episode: 1642 score: 5.0 memory length: 389198 epsilon: 0.4273859800103951
steps: 272 lr: 0.0001 evaluation reward: 7.33
episode: 1643 score: 11.0 memory length: 389741 epsilon:
0.42631084001038827 steps: 543 lr: 0.0001 evaluation reward: 7.36
episode: 1644 score: 14.0 memory length: 390255 epsilon:
0.42529312001038183 steps: 514 lr: 0.0001 evaluation reward: 7.47
episode: 1645 score: 2.0 memory length: 390473 epsilon: 0.4248614800103791
steps: 218 lr: 0.0001 evaluation reward: 7.42
episode: 1646 score: 7.0 memory length: 390842 epsilon: 0.4241308600103745
steps: 369 lr: 0.0001 evaluation reward: 7.36
episode: 1647 score: 6.0 memory length: 391144 epsilon: 0.4235329000103707
steps: 302 lr: 0.0001 evaluation reward: 7.35
episode: 1648 score: 7.0 memory length: 391528 epsilon: 0.4227725800103659
steps: 384 lr: 0.0001 evaluation reward: 7.38
episode: 1649 score: 6.0 memory length: 391898 epsilon:
0.42203998001036125 steps: 370 lr: 0.0001 evaluation reward: 7.36
episode: 1650 score: 7.0 memory length: 392305 epsilon:
0.42123412001035615 steps: 407 lr: 0.0001 evaluation reward: 7.39
episode: 1651 score: 8.0 memory length: 392729 epsilon:
0.42039460001035084 steps: 424 lr: 0.0001 evaluation reward: 7.37
episode: 1652 score: 6.0 memory length: 393070 epsilon:
0.41971942001034657 steps: 341 lr: 0.0001 evaluation reward: 7.28
episode: 1653 score: 9.0 memory length: 393525 epsilon:
0.41881852001034087 steps: 455 lr: 0.0001 evaluation reward: 7.34
episode: 1654 score: 9.0 memory length: 393998 epsilon:
0.41788198001033494 steps: 473 lr: 0.0001 evaluation reward: 7.39
episode: 1655 score: 11.0 memory length: 394555 epsilon:
0.41677912001032796 steps: 557 lr: 0.0001 evaluation reward: 7.4
episode: 1656 score: 3.0 memory length: 394768 epsilon: 0.4163573800103253
steps: 213 lr: 0.0001 evaluation reward: 7.35
episode: 1657 score: 9.0 memory length: 395215 epsilon: 0.4154723200103197
steps: 447 lr: 0.0001 evaluation reward: 7.39
episode: 1658 score: 7.0 memory length: 395621 epsilon: 0.4146684400103146
steps: 406 lr: 0.0001 evaluation reward: 7.38
episode: 1659 score: 5.0 memory length: 395930 epsilon:
0.41405662001031074 steps: 309 lr: 0.0001 evaluation reward: 7.37

episode: 1660 score: 6.0 memory length: 396264 epsilon:
 0.41339530001030655 steps: 334 lr: 0.0001 evaluation reward: 7.36
 episode: 1661 score: 10.0 memory length: 396734 epsilon:
 0.41246470001030067 steps: 470 lr: 0.0001 evaluation reward: 7.4
 episode: 1662 score: 6.0 memory length: 397095 epsilon:
 0.41174992001029614 steps: 361 lr: 0.0001 evaluation reward: 7.4
 episode: 1663 score: 4.0 memory length: 397372 epsilon: 0.4112014600102927
 steps: 277 lr: 0.0001 evaluation reward: 7.38
 episode: 1664 score: 10.0 memory length: 397915 epsilon:
 0.41012632001028587 steps: 543 lr: 0.0001 evaluation reward: 7.41
 episode: 1665 score: 8.0 memory length: 398331 epsilon:
 0.40930264001028066 steps: 416 lr: 0.0001 evaluation reward: 7.37
 episode: 1666 score: 7.0 memory length: 398724 epsilon:
 0.40852450001027574 steps: 393 lr: 0.0001 evaluation reward: 7.36
 episode: 1667 score: 9.0 memory length: 399179 epsilon:
 0.40762360001027004 steps: 455 lr: 0.0001 evaluation reward: 7.39
 episode: 1668 score: 6.0 memory length: 399518 epsilon: 0.4069523800102658
 steps: 339 lr: 0.0001 evaluation reward: 7.33
 episode: 1669 score: 11.0 memory length: 400038 epsilon:
 0.4059227800102593 steps: 520 lr: 0.0001 evaluation reward: 7.4
 episode: 1670 score: 16.0 memory length: 400638 epsilon:
 0.40473478001025176 steps: 600 lr: 0.0001 evaluation reward: 7.48
 episode: 1671 score: 7.0 memory length: 401018 epsilon: 0.403982380010247
 steps: 380 lr: 0.0001 evaluation reward: 7.44
 episode: 1672 score: 5.0 memory length: 401304 epsilon: 0.4034161000102434
 steps: 286 lr: 0.0001 evaluation reward: 7.35
 episode: 1673 score: 3.0 memory length: 401534 epsilon:
 0.40296070001024054 steps: 230 lr: 0.0001 evaluation reward: 7.35
 episode: 1674 score: 6.0 memory length: 401869 epsilon:
 0.40229740001023634 steps: 335 lr: 0.0001 evaluation reward: 7.39
 episode: 1675 score: 11.0 memory length: 402491 epsilon:
 0.40106584001022855 steps: 622 lr: 0.0001 evaluation reward: 7.45
 episode: 1676 score: 3.0 memory length: 402721 epsilon:
 0.40061044001022567 steps: 230 lr: 0.0001 evaluation reward: 7.45
 episode: 1677 score: 9.0 memory length: 403194 epsilon:
 0.39967390001021974 steps: 473 lr: 0.0001 evaluation reward: 7.43
 episode: 1678 score: 15.0 memory length: 403795 epsilon:
 0.3984839200102122 steps: 601 lr: 0.0001 evaluation reward: 7.52
 episode: 1679 score: 10.0 memory length: 404281 epsilon:
 0.3975216400102061 steps: 486 lr: 0.0001 evaluation reward: 7.58
 episode: 1680 score: 7.0 memory length: 404692 epsilon: 0.396707860010201
 steps: 411 lr: 0.0001 evaluation reward: 7.57
 episode: 1681 score: 12.0 memory length: 405194 epsilon:
 0.3957139000101947 steps: 502 lr: 0.0001 evaluation reward: 7.65
 episode: 1682 score: 11.0 memory length: 405564 epsilon:
 0.39498130001019005 steps: 370 lr: 0.0001 evaluation reward: 7.63
 episode: 1683 score: 5.0 memory length: 405876 epsilon:
 0.39436354001018614 steps: 312 lr: 0.0001 evaluation reward: 7.61

episode: 1684 score: 11.0 memory length: 406428 epsilon:
 0.3932705800101792 steps: 552 lr: 0.0001 evaluation reward: 7.69
 episode: 1685 score: 11.0 memory length: 406937 epsilon:
 0.39226276001017285 steps: 509 lr: 0.0001 evaluation reward: 7.73
 episode: 1686 score: 18.0 memory length: 407578 epsilon:
 0.3909935800101648 steps: 641 lr: 0.0001 evaluation reward: 7.87
 episode: 1687 score: 14.0 memory length: 408116 epsilon:
 0.3899283400101581 steps: 538 lr: 0.0001 evaluation reward: 7.95
 episode: 1688 score: 4.0 memory length: 408369 epsilon: 0.3894274000101549
 steps: 253 lr: 0.0001 evaluation reward: 7.95
 episode: 1689 score: 3.0 memory length: 408619 epsilon: 0.3889324000101518
 steps: 250 lr: 0.0001 evaluation reward: 7.89
 episode: 1690 score: 9.0 memory length: 408981 epsilon:
 0.38821564001014724 steps: 362 lr: 0.0001 evaluation reward: 7.91
 episode: 1691 score: 8.0 memory length: 409415 epsilon: 0.3873563200101418
 steps: 434 lr: 0.0001 evaluation reward: 7.94
 episode: 1692 score: 11.0 memory length: 409949 epsilon:
 0.3862990000101351 steps: 534 lr: 0.0001 evaluation reward: 8.02
 episode: 1693 score: 12.0 memory length: 410413 epsilon:
 0.3853802800101293 steps: 464 lr: 0.0001 evaluation reward: 8.08
 episode: 1694 score: 11.0 memory length: 410940 epsilon:
 0.3843368200101227 steps: 527 lr: 0.0001 evaluation reward: 8.09
 episode: 1695 score: 11.0 memory length: 411461 epsilon:
 0.3833052400101162 steps: 521 lr: 0.0001 evaluation reward: 8.14
 episode: 1696 score: 11.0 memory length: 411987 epsilon:
 0.3822637600101096 steps: 526 lr: 0.0001 evaluation reward: 8.2
 episode: 1697 score: 11.0 memory length: 412576 epsilon:
 0.3810975400101022 steps: 589 lr: 0.0001 evaluation reward: 8.19
 episode: 1698 score: 6.0 memory length: 412950 epsilon: 0.3803570200100975
 steps: 374 lr: 0.0001 evaluation reward: 8.17
 episode: 1699 score: 7.0 memory length: 413340 epsilon:
 0.37958482001009264 steps: 390 lr: 0.0001 evaluation reward: 8.22
 episode: 1700 score: 10.0 memory length: 413747 epsilon:
 0.37877896001008754 steps: 407 lr: 0.0001 evaluation reward: 8.2
 episode: 1701 score: 5.0 memory length: 414077 epsilon: 0.3781255600100834
 steps: 330 lr: 0.0001 evaluation reward: 8.19
 episode: 1702 score: 12.0 memory length: 414666 epsilon: 0.376959340010076
 steps: 589 lr: 0.0001 evaluation reward: 8.25
 episode: 1703 score: 7.0 memory length: 415071 epsilon:
 0.37615744001007095 steps: 405 lr: 0.0001 evaluation reward: 8.29
 episode: 1704 score: 10.0 memory length: 415621 epsilon:
 0.37506844001006406 steps: 550 lr: 0.0001 evaluation reward: 8.28
 episode: 1705 score: 16.0 memory length: 416188 epsilon:
 0.37394578001005696 steps: 567 lr: 0.0001 evaluation reward: 8.39
 episode: 1706 score: 3.0 memory length: 416436 epsilon:
 0.37345474001005385 steps: 248 lr: 0.0001 evaluation reward: 8.3
 episode: 1707 score: 8.0 memory length: 416890 epsilon:
 0.37255582001004817 steps: 454 lr: 0.0001 evaluation reward: 8.18

episode: 1708 score: 9.0 memory length: 417368 epsilon: 0.3716093800100422
 steps: 478 lr: 0.0001 evaluation reward: 8.15
 episode: 1709 score: 5.0 memory length: 417659 epsilon:
 0.37103320001003853 steps: 291 lr: 0.0001 evaluation reward: 8.13
 episode: 1710 score: 8.0 memory length: 418114 epsilon:
 0.37013230001003283 steps: 455 lr: 0.0001 evaluation reward: 8.17
 episode: 1711 score: 7.0 memory length: 418491 epsilon: 0.3693858400100281
 steps: 377 lr: 0.0001 evaluation reward: 8.2
 episode: 1712 score: 10.0 memory length: 418988 epsilon:
 0.3684017800100219 steps: 497 lr: 0.0001 evaluation reward: 8.23
 episode: 1713 score: 8.0 memory length: 419423 epsilon:
 0.36754048001001643 steps: 435 lr: 0.0001 evaluation reward: 8.23
 episode: 1714 score: 8.0 memory length: 419898 epsilon: 0.3665999800100105
 steps: 475 lr: 0.0001 evaluation reward: 8.22
 episode: 1715 score: 4.0 memory length: 420157 epsilon:
 0.36608716001000724 steps: 259 lr: 0.0001 evaluation reward: 8.22
 episode: 1716 score: 8.0 memory length: 420591 epsilon: 0.3652278400100018
 steps: 434 lr: 0.0001 evaluation reward: 8.24
 episode: 1717 score: 5.0 memory length: 420866 epsilon:
 0.36468334000999836 steps: 275 lr: 0.0001 evaluation reward: 8.22
 episode: 1718 score: 9.0 memory length: 421318 epsilon: 0.3637883800099927
 steps: 452 lr: 0.0001 evaluation reward: 8.25
 episode: 1719 score: 7.0 memory length: 421659 epsilon: 0.3631132000099884
 steps: 341 lr: 0.0001 evaluation reward: 8.21
 episode: 1720 score: 6.0 memory length: 421960 epsilon:
 0.36251722000998465 steps: 301 lr: 0.0001 evaluation reward: 8.21
 episode: 1721 score: 7.0 memory length: 422350 epsilon:
 0.36174502000997977 steps: 390 lr: 0.0001 evaluation reward: 8.22
 episode: 1722 score: 11.0 memory length: 422924 epsilon:
 0.3606085000099726 steps: 574 lr: 0.0001 evaluation reward: 8.26
 episode: 1723 score: 8.0 memory length: 423397 epsilon:
 0.35967196000996665 steps: 473 lr: 0.0001 evaluation reward: 8.24
 episode: 1724 score: 5.0 memory length: 423709 epsilon:
 0.35905420000996274 steps: 312 lr: 0.0001 evaluation reward: 8.21
 episode: 1725 score: 7.0 memory length: 424079 epsilon: 0.3583216000099581
 steps: 370 lr: 0.0001 evaluation reward: 8.16
 episode: 1726 score: 8.0 memory length: 424497 epsilon:
 0.35749396000995287 steps: 418 lr: 0.0001 evaluation reward: 8.2
 episode: 1727 score: 5.0 memory length: 424823 epsilon: 0.3568484800099488
 steps: 326 lr: 0.0001 evaluation reward: 8.16
 episode: 1728 score: 10.0 memory length: 425313 epsilon:
 0.35587828000994265 steps: 490 lr: 0.0001 evaluation reward: 8.16
 episode: 1729 score: 7.0 memory length: 425700 epsilon: 0.3551120200099378
 steps: 387 lr: 0.0001 evaluation reward: 8.14
 episode: 1730 score: 9.0 memory length: 426152 epsilon:
 0.35421706000993214 steps: 452 lr: 0.0001 evaluation reward: 8.16
 episode: 1731 score: 6.0 memory length: 426490 epsilon: 0.3535478200099279
 steps: 338 lr: 0.0001 evaluation reward: 8.09

episode: 1732 score: 4.0 memory length: 426747 epsilon: 0.3530389600099247
 steps: 257 lr: 0.0001 evaluation reward: 8.08
 episode: 1733 score: 8.0 memory length: 427186 epsilon: 0.3521697400099192
 steps: 439 lr: 0.0001 evaluation reward: 8.11
 episode: 1734 score: 10.0 memory length: 427735 epsilon:
 0.3510827200099123 steps: 549 lr: 0.0001 evaluation reward: 8.16
 episode: 1735 score: 6.0 memory length: 428080 epsilon: 0.350399620009908
 steps: 345 lr: 0.0001 evaluation reward: 8.16
 episode: 1736 score: 5.0 memory length: 428366 epsilon: 0.3498333400099044
 steps: 286 lr: 0.0001 evaluation reward: 8.14
 episode: 1737 score: 9.0 memory length: 428889 epsilon:
 0.34879780000989785 steps: 523 lr: 0.0001 evaluation reward: 8.18
 episode: 1738 score: 10.0 memory length: 429343 epsilon:
 0.34789888000989216 steps: 454 lr: 0.0001 evaluation reward: 8.18
 episode: 1739 score: 15.0 memory length: 429944 epsilon:
 0.34670890000988464 steps: 601 lr: 0.0001 evaluation reward: 8.29
 episode: 1740 score: 8.0 memory length: 430391 epsilon:
 0.34582384000987904 steps: 447 lr: 0.0001 evaluation reward: 8.31
 episode: 1741 score: 7.0 memory length: 430778 epsilon: 0.3450575800098742
 steps: 387 lr: 0.0001 evaluation reward: 8.17
 episode: 1742 score: 7.0 memory length: 431174 epsilon: 0.3442735000098692
 steps: 396 lr: 0.0001 evaluation reward: 8.19
 episode: 1743 score: 8.0 memory length: 431594 epsilon:
 0.34344190000986397 steps: 420 lr: 0.0001 evaluation reward: 8.16
 episode: 1744 score: 8.0 memory length: 432018 epsilon:
 0.34260238000985865 steps: 424 lr: 0.0001 evaluation reward: 8.1
 episode: 1745 score: 10.0 memory length: 432397 epsilon:
 0.3418519600098539 steps: 379 lr: 0.0001 evaluation reward: 8.18
 episode: 1746 score: 9.0 memory length: 432844 epsilon: 0.3409669000098483
 steps: 447 lr: 0.0001 evaluation reward: 8.2
 episode: 1747 score: 15.0 memory length: 433398 epsilon:
 0.33986998000984137 steps: 554 lr: 0.0001 evaluation reward: 8.29
 episode: 1748 score: 10.0 memory length: 433877 epsilon:
 0.33892156000983537 steps: 479 lr: 0.0001 evaluation reward: 8.32
 episode: 1749 score: 19.0 memory length: 434568 epsilon:
 0.3375533800098267 steps: 691 lr: 0.0001 evaluation reward: 8.45
 episode: 1750 score: 7.0 memory length: 434994 epsilon: 0.3367099000098214
 steps: 426 lr: 0.0001 evaluation reward: 8.45
 episode: 1751 score: 5.0 memory length: 435320 epsilon: 0.3360644200098173
 steps: 326 lr: 0.0001 evaluation reward: 8.42
 episode: 1752 score: 11.0 memory length: 435850 epsilon:
 0.33501502000981065 steps: 530 lr: 0.0001 evaluation reward: 8.47
 episode: 1753 score: 11.0 memory length: 436412 epsilon:
 0.3339022600098036 steps: 562 lr: 0.0001 evaluation reward: 8.49
 episode: 1754 score: 10.0 memory length: 436932 epsilon:
 0.3328726600097971 steps: 520 lr: 0.0001 evaluation reward: 8.5
 episode: 1755 score: 9.0 memory length: 437363 epsilon: 0.3320192800097917
 steps: 431 lr: 0.0001 evaluation reward: 8.48

episode: 1756 score: 11.0 memory length: 437863 epsilon:
 0.33102928000978543 steps: 500 lr: 0.0001 evaluation reward: 8.56
 episode: 1757 score: 8.0 memory length: 438299 epsilon:
 0.33016600000977997 steps: 436 lr: 0.0001 evaluation reward: 8.55
 episode: 1758 score: 11.0 memory length: 438800 epsilon:
 0.3291740200097737 steps: 501 lr: 0.0001 evaluation reward: 8.59
 episode: 1759 score: 15.0 memory length: 439361 epsilon:
 0.32806324000976667 steps: 561 lr: 0.0001 evaluation reward: 8.69
 episode: 1760 score: 14.0 memory length: 439993 epsilon:
 0.32681188000975875 steps: 632 lr: 0.0001 evaluation reward: 8.77
 episode: 1761 score: 16.0 memory length: 440605 epsilon:
 0.3256001200097511 steps: 612 lr: 0.0001 evaluation reward: 8.83
 episode: 1762 score: 7.0 memory length: 440992 epsilon:
 0.32483386000974623 steps: 387 lr: 0.0001 evaluation reward: 8.84
 episode: 1763 score: 13.0 memory length: 441616 epsilon:
 0.3235983400097384 steps: 624 lr: 0.0001 evaluation reward: 8.93
 episode: 1764 score: 9.0 memory length: 442115 epsilon:
 0.32261032000973217 steps: 499 lr: 0.0001 evaluation reward: 8.92
 episode: 1765 score: 5.0 memory length: 442423 epsilon: 0.3220004800097283
 steps: 308 lr: 0.0001 evaluation reward: 8.89
 episode: 1766 score: 9.0 memory length: 442840 epsilon: 0.3211748200097231
 steps: 417 lr: 0.0001 evaluation reward: 8.91
 episode: 1767 score: 19.0 memory length: 443618 epsilon:
 0.31963438000971334 steps: 778 lr: 0.0001 evaluation reward: 9.01
 episode: 1768 score: 15.0 memory length: 444243 epsilon:
 0.3183968800097055 steps: 625 lr: 0.0001 evaluation reward: 9.1
 episode: 1769 score: 8.0 memory length: 444677 epsilon:
 0.31753756000970007 steps: 434 lr: 0.0001 evaluation reward: 9.07
 episode: 1770 score: 7.0 memory length: 445072 epsilon: 0.3167554600096951
 steps: 395 lr: 0.0001 evaluation reward: 8.98
 episode: 1771 score: 9.0 memory length: 445541 epsilon:
 0.31582684000968925 steps: 469 lr: 0.0001 evaluation reward: 9.0
 episode: 1772 score: 8.0 memory length: 445941 epsilon:
 0.31503484000968424 steps: 400 lr: 0.0001 evaluation reward: 9.03
 episode: 1773 score: 11.0 memory length: 446330 epsilon:
 0.31426462000967936 steps: 389 lr: 0.0001 evaluation reward: 9.11
 episode: 1774 score: 10.0 memory length: 446830 epsilon:
 0.3132746200096731 steps: 500 lr: 0.0001 evaluation reward: 9.15
 episode: 1775 score: 11.0 memory length: 447246 epsilon:
 0.3124509400096679 steps: 416 lr: 0.0001 evaluation reward: 9.15
 episode: 1776 score: 12.0 memory length: 447828 epsilon:
 0.3112985800096606 steps: 582 lr: 0.0001 evaluation reward: 9.24
 episode: 1777 score: 19.0 memory length: 448377 epsilon:
 0.3102115600096537 steps: 549 lr: 0.0001 evaluation reward: 9.34
 episode: 1778 score: 9.0 memory length: 448887 epsilon:
 0.30920176000964733 steps: 510 lr: 0.0001 evaluation reward: 9.28
 episode: 1779 score: 19.0 memory length: 449624 epsilon:
 0.3077425000096381 steps: 737 lr: 0.0001 evaluation reward: 9.37

episode: 1780 score: 9.0 memory length: 449931 epsilon:
 0.30713464000963425 steps: 307 lr: 0.0001 evaluation reward: 9.39
 episode: 1781 score: 9.0 memory length: 450403 epsilon:
 0.30620008000962834 steps: 472 lr: 0.0001 evaluation reward: 9.36
 episode: 1782 score: 6.0 memory length: 450754 epsilon:
 0.30550510000962394 steps: 351 lr: 0.0001 evaluation reward: 9.31
 episode: 1783 score: 12.0 memory length: 451281 epsilon:
 0.30446164000961734 steps: 527 lr: 0.0001 evaluation reward: 9.38
 episode: 1784 score: 19.0 memory length: 451763 epsilon:
 0.3035072800096113 steps: 482 lr: 0.0001 evaluation reward: 9.46
 episode: 1785 score: 7.0 memory length: 452116 epsilon: 0.3028083400096069
 steps: 353 lr: 0.0001 evaluation reward: 9.42
 episode: 1786 score: 14.0 memory length: 452762 epsilon:
 0.3015292600095988 steps: 646 lr: 0.0001 evaluation reward: 9.38
 episode: 1787 score: 11.0 memory length: 453320 epsilon:
 0.3004244200095918 steps: 558 lr: 0.0001 evaluation reward: 9.35
 episode: 1788 score: 14.0 memory length: 453987 epsilon:
 0.29910376000958344 steps: 667 lr: 0.0001 evaluation reward: 9.45
 episode: 1789 score: 5.0 memory length: 454315 epsilon:
 0.29845432000957933 steps: 328 lr: 0.0001 evaluation reward: 9.47
 episode: 1790 score: 12.0 memory length: 454881 epsilon:
 0.29733364000957224 steps: 566 lr: 0.0001 evaluation reward: 9.5
 episode: 1791 score: 7.0 memory length: 455304 epsilon:
 0.29649610000956694 steps: 423 lr: 0.0001 evaluation reward: 9.49
 episode: 1792 score: 13.0 memory length: 455787 epsilon:
 0.2955397600095609 steps: 483 lr: 0.0001 evaluation reward: 9.51
 episode: 1793 score: 11.0 memory length: 456295 epsilon:
 0.29453392000955453 steps: 508 lr: 0.0001 evaluation reward: 9.5
 episode: 1794 score: 15.0 memory length: 456920 epsilon:
 0.2932964200095467 steps: 625 lr: 0.0001 evaluation reward: 9.54
 episode: 1795 score: 10.0 memory length: 457454 epsilon: 0.29223910000954
 steps: 534 lr: 0.0001 evaluation reward: 9.53
 episode: 1796 score: 7.0 memory length: 457879 epsilon: 0.2913976000095347
 steps: 425 lr: 0.0001 evaluation reward: 9.49
 episode: 1797 score: 9.0 memory length: 458320 epsilon:
 0.29052442000952916 steps: 441 lr: 0.0001 evaluation reward: 9.47
 episode: 1798 score: 9.0 memory length: 458777 epsilon:
 0.28961956000952344 steps: 457 lr: 0.0001 evaluation reward: 9.5
 episode: 1799 score: 9.0 memory length: 459214 epsilon:
 0.28875430000951796 steps: 437 lr: 0.0001 evaluation reward: 9.52
 episode: 1800 score: 9.0 memory length: 459647 epsilon:
 0.28789696000951254 steps: 433 lr: 0.0001 evaluation reward: 9.51
 episode: 1801 score: 11.0 memory length: 460186 epsilon:
 0.2868297400095058 steps: 539 lr: 0.0001 evaluation reward: 9.57
 episode: 1802 score: 13.0 memory length: 460806 epsilon: 0.285602140009498
 steps: 620 lr: 0.0001 evaluation reward: 9.58
 episode: 1803 score: 8.0 memory length: 461213 epsilon: 0.2847962800094929
 steps: 407 lr: 0.0001 evaluation reward: 9.59

episode: 1804 score: 12.0 memory length: 461817 epsilon:
 0.28360036000948535 steps: 604 lr: 0.0001 evaluation reward: 9.61
 episode: 1805 score: 15.0 memory length: 462395 epsilon:
 0.2824559200094781 steps: 578 lr: 0.0001 evaluation reward: 9.6
 episode: 1806 score: 12.0 memory length: 462966 epsilon:
 0.28132534000947096 steps: 571 lr: 0.0001 evaluation reward: 9.69
 episode: 1807 score: 16.0 memory length: 463613 epsilon:
 0.28004428000946285 steps: 647 lr: 0.0001 evaluation reward: 9.77
 episode: 1808 score: 18.0 memory length: 464319 epsilon: 0.278646400009454
 steps: 706 lr: 0.0001 evaluation reward: 9.86
 episode: 1809 score: 7.0 memory length: 464689 epsilon: 0.2779138000094494
 steps: 370 lr: 0.0001 evaluation reward: 9.88
 episode: 1810 score: 8.0 memory length: 465156 epsilon: 0.2769891400094435
 steps: 467 lr: 0.0001 evaluation reward: 9.88
 episode: 1811 score: 9.0 memory length: 465650 epsilon:
 0.27601102000943734 steps: 494 lr: 0.0001 evaluation reward: 9.9
 episode: 1812 score: 9.0 memory length: 465976 epsilon:
 0.27536554000943325 steps: 326 lr: 0.0001 evaluation reward: 9.89
 episode: 1813 score: 9.0 memory length: 466427 epsilon: 0.2744725600094276
 steps: 451 lr: 0.0001 evaluation reward: 9.9
 episode: 1814 score: 16.0 memory length: 467052 epsilon:
 0.2732350600094198 steps: 625 lr: 0.0001 evaluation reward: 9.98
 episode: 1815 score: 12.0 memory length: 467597 epsilon:
 0.27215596000941294 steps: 545 lr: 0.0001 evaluation reward: 10.06
 episode: 1816 score: 6.0 memory length: 467934 epsilon: 0.2714887000094087
 steps: 337 lr: 0.0001 evaluation reward: 10.04
 episode: 1817 score: 14.0 memory length: 468629 epsilon: 0.2701126000094
 steps: 695 lr: 0.0001 evaluation reward: 10.13
 episode: 1818 score: 15.0 memory length: 469163 epsilon:
 0.2690552800093933 steps: 534 lr: 0.0001 evaluation reward: 10.19
 episode: 1819 score: 9.0 memory length: 469647 epsilon:
 0.26809696000938726 steps: 484 lr: 0.0001 evaluation reward: 10.21
 episode: 1820 score: 9.0 memory length: 470159 epsilon:
 0.26708320000938085 steps: 512 lr: 0.0001 evaluation reward: 10.24
 episode: 1821 score: 5.0 memory length: 470484 epsilon: 0.2664397000093768
 steps: 325 lr: 0.0001 evaluation reward: 10.22
 episode: 1822 score: 9.0 memory length: 470990 epsilon:
 0.26543782000937044 steps: 506 lr: 0.0001 evaluation reward: 10.2
 episode: 1823 score: 7.0 memory length: 471377 epsilon: 0.2646715600093656
 steps: 387 lr: 0.0001 evaluation reward: 10.19
 episode: 1824 score: 17.0 memory length: 471999 epsilon:
 0.2634400000093578 steps: 622 lr: 0.0001 evaluation reward: 10.31
 episode: 1825 score: 13.0 memory length: 472608 epsilon:
 0.26223418000935017 steps: 609 lr: 0.0001 evaluation reward: 10.37
 episode: 1826 score: 10.0 memory length: 473092 epsilon:
 0.2612758600093441 steps: 484 lr: 0.0001 evaluation reward: 10.39
 episode: 1827 score: 7.0 memory length: 473542 epsilon:
 0.26038486000933847 steps: 450 lr: 0.0001 evaluation reward: 10.41

episode: 1828 score: 15.0 memory length: 474093 epsilon:
 0.25929388000933157 steps: 551 lr: 0.0001 evaluation reward: 10.46
 episode: 1829 score: 11.0 memory length: 474549 epsilon:
 0.25839100000932586 steps: 456 lr: 0.0001 evaluation reward: 10.5
 episode: 1830 score: 5.0 memory length: 474878 epsilon:
 0.25773958000932173 steps: 329 lr: 0.0001 evaluation reward: 10.46
 episode: 1831 score: 7.0 memory length: 475229 epsilon:
 0.25704460000931734 steps: 351 lr: 0.0001 evaluation reward: 10.47
 episode: 1832 score: 9.0 memory length: 475693 epsilon: 0.2561258800093115
 steps: 464 lr: 0.0001 evaluation reward: 10.52
 episode: 1833 score: 16.0 memory length: 476280 epsilon:
 0.25496362000930417 steps: 587 lr: 0.0001 evaluation reward: 10.6
 episode: 1834 score: 8.0 memory length: 476695 epsilon:
 0.25414192000929897 steps: 415 lr: 0.0001 evaluation reward: 10.58
 episode: 1835 score: 21.0 memory length: 477467 epsilon:
 0.2526133600092893 steps: 772 lr: 0.0001 evaluation reward: 10.73
 episode: 1836 score: 14.0 memory length: 478117 epsilon:
 0.25132636000928116 steps: 650 lr: 0.0001 evaluation reward: 10.82
 episode: 1837 score: 11.0 memory length: 478669 epsilon:
 0.25023340000927424 steps: 552 lr: 0.0001 evaluation reward: 10.84
 episode: 1838 score: 8.0 memory length: 479094 epsilon:
 0.24939190000926892 steps: 425 lr: 0.0001 evaluation reward: 10.82
 episode: 1839 score: 15.0 memory length: 479686 epsilon:
 0.2482197400092615 steps: 592 lr: 0.0001 evaluation reward: 10.82
 episode: 1840 score: 13.0 memory length: 480299 epsilon:
 0.24700600000925382 steps: 613 lr: 0.0001 evaluation reward: 10.87
 episode: 1841 score: 8.0 memory length: 480778 epsilon:
 0.24605758000924782 steps: 479 lr: 0.0001 evaluation reward: 10.88
 episode: 1842 score: 9.0 memory length: 481251 epsilon: 0.2451210400092419
 steps: 473 lr: 0.0001 evaluation reward: 10.9
 episode: 1843 score: 8.0 memory length: 481673 epsilon: 0.2442854800092366
 steps: 422 lr: 0.0001 evaluation reward: 10.9
 episode: 1844 score: 10.0 memory length: 482083 epsilon:
 0.24347368000923147 steps: 410 lr: 0.0001 evaluation reward: 10.92
 episode: 1845 score: 8.0 memory length: 482476 epsilon:
 0.24269554000922655 steps: 393 lr: 0.0001 evaluation reward: 10.9
 episode: 1846 score: 6.0 memory length: 482848 epsilon: 0.2419589800092219
 steps: 372 lr: 0.0001 evaluation reward: 10.87
 episode: 1847 score: 12.0 memory length: 483407 epsilon:
 0.2408521600092149 steps: 559 lr: 0.0001 evaluation reward: 10.84
 episode: 1848 score: 14.0 memory length: 483926 epsilon:
 0.2398245400092084 steps: 519 lr: 0.0001 evaluation reward: 10.88
 episode: 1849 score: 14.0 memory length: 484483 epsilon:
 0.2387216800092014 steps: 557 lr: 0.0001 evaluation reward: 10.83
 episode: 1850 score: 7.0 memory length: 484874 epsilon: 0.2379475000091965
 steps: 391 lr: 0.0001 evaluation reward: 10.83
 episode: 1851 score: 7.0 memory length: 485276 epsilon:
 0.23715154000919147 steps: 402 lr: 0.0001 evaluation reward: 10.85

```

episode: 1852  score: 12.0  memory length: 485810  epsilon:
0.23609422000918479  steps: 534  lr: 0.0001  evaluation reward: 10.86
episode: 1853  score: 8.0  memory length: 486263  epsilon: 0.2351972800091791
steps: 453  lr: 0.0001  evaluation reward: 10.83
episode: 1854  score: 8.0  memory length: 486699  epsilon:
0.23433400000917365  steps: 436  lr: 0.0001  evaluation reward: 10.81
episode: 1855  score: 9.0  memory length: 487198  epsilon: 0.2333459800091674
steps: 499  lr: 0.0001  evaluation reward: 10.81
episode: 1856  score: 6.0  memory length: 487556  epsilon: 0.2326371400091629
steps: 358  lr: 0.0001  evaluation reward: 10.76
episode: 1857  score: 21.0  memory length: 488229  epsilon:
0.23130460000915448  steps: 673  lr: 0.0001  evaluation reward: 10.89
episode: 1858  score: 6.0  memory length: 488568  epsilon:
0.23063338000915024  steps: 339  lr: 0.0001  evaluation reward: 10.84
episode: 1859  score: 9.0  memory length: 489061  epsilon:
0.22965724000914406  steps: 493  lr: 0.0001  evaluation reward: 10.78
episode: 1860  score: 8.0  memory length: 489499  epsilon:
0.22879000000913857  steps: 438  lr: 0.0001  evaluation reward: 10.72
episode: 1861  score: 12.0  memory length: 490115  epsilon:
0.22757032000913086  steps: 616  lr: 0.0001  evaluation reward: 10.68
episode: 1862  score: 16.0  memory length: 490741  epsilon: 0.226330840009123
steps: 626  lr: 0.0001  evaluation reward: 10.77
episode: 1863  score: 11.0  memory length: 491285  epsilon:
0.2252537200091162  steps: 544  lr: 0.0001  evaluation reward: 10.75
episode: 1864  score: 10.0  memory length: 491706  epsilon:
0.22442014000911092  steps: 421  lr: 0.0001  evaluation reward: 10.76
episode: 1865  score: 12.0  memory length: 492161  epsilon:
0.22351924000910522  steps: 455  lr: 0.0001  evaluation reward: 10.83
episode: 1866  score: 13.0  memory length: 492512  epsilon:
0.22282426000910083  steps: 351  lr: 0.0001  evaluation reward: 10.87
episode: 1867  score: 10.0  memory length: 493085  epsilon:
0.22168972000909365  steps: 573  lr: 0.0001  evaluation reward: 10.78
episode: 1868  score: 17.0  memory length: 493778  epsilon:
0.22031758000908497  steps: 693  lr: 0.0001  evaluation reward: 10.8
episode: 1869  score: 19.0  memory length: 494340  epsilon:
0.21920482000907793  steps: 562  lr: 0.0001  evaluation reward: 10.91
episode: 1870  score: 10.0  memory length: 494840  epsilon:
0.21821482000907166  steps: 500  lr: 0.0001  evaluation reward: 10.94
episode: 1871  score: 7.0  memory length: 495211  epsilon:
0.21748024000906702  steps: 371  lr: 0.0001  evaluation reward: 10.92

```

KeyboardInterrupt

Traceback (most recent call last)

```

<ipython-input-5-f85d5c09a4b5> in <module>
    53     pylab.ylabel('Rewards')
    54     pylab.title('Episodes vs Reward')

```

```

--> 55         pylab.savefig("./save_graph/breakout_ddqn.png") # save graph
    ↳for training visualization
    56
    57         # every episode, plot the play time

c:\users\computer\miniconda3\lib\site-packages\matplotlib\pyplot.py in
    ↳savefig(*args, **kwargs)
    858         fig = gcf()
    859         res = fig.savefig(*args, **kwargs)
--> 860         fig.canvas.draw_idle() # need this if 'transparent=True' to reset
    ↳colors
    861         return res
    862

c:\users\computer\miniconda3\lib\site-packages\matplotlib\backend_bases.py in
    ↳draw_idle(self, *args, **kwargs)
    2010         if not self._is_idle_drawing:
    2011             with self._idle_draw_cntx():
--> 2012                 self.draw(*args, **kwargs)
    2013
    2014         @cbook.deprecated("3.2")

c:\users\computer\miniconda3\lib\site-packages\matplotlib\backends\backend_agg.py
    ↳in draw(self)
    405         (self.toolbar._wait_cursor_for_draw_cm() if self.toolbar
    406         else nullcontext()):
--> 407         self.figure.draw(self.renderer)
    408         # A GUI class may be need to update a window using this draw,
    ↳so
    409         # don't forget to call the superclass.

c:\users\computer\miniconda3\lib\site-packages\matplotlib\artist.py in
    ↳draw_wrapper(artist, renderer, *args, **kwargs)
    39         renderer.start_filter()
    40
--> 41         return draw(artist, renderer, *args, **kwargs)
    42         finally:
    43             if artist.get_agg_filter() is not None:

c:\users\computer\miniconda3\lib\site-packages\matplotlib\figure.py in draw(self,
    ↳renderer)
    1861
    1862         self.patch.draw(renderer)
--> 1863         mimage._draw_list_compositing_images(
    1864             renderer, self, artists, self.suppressComposite)
    1865

```

```

c:\users\computer\miniconda3\lib\site-packages\matplotlib\image.py in
↳ _draw_list_compositing_images(renderer, parent, artists, suppress_composite)
    129     if not_composite or not has_images:
    130         for a in artists:
--> 131             a.draw(renderer)
    132     else:
    133         # Composite any adjacent images together

c:\users\computer\miniconda3\lib\site-packages\matplotlib\artist.py in
↳ draw_wrapper(artist, renderer, *args, **kwargs)
    39         renderer.start_filter()
    40
---> 41         return draw(artist, renderer, *args, **kwargs)
    42     finally:
    43         if artist.get_agg_filter() is not None:

c:\users\computer\miniconda3\lib\site-packages\matplotlib\cbook\deprecation.py in
↳ wrapper(*inner_args, **inner_kwargs)
    409         else deprecation_addendum,
    410         **kwargs)
--> 411     return func(*inner_args, **inner_kwargs)
    412
    413     return wrapper

c:\users\computer\miniconda3\lib\site-packages\matplotlib\axes\_base.py in
↳ draw(self, renderer, inframe)
    2745         renderer.stop_rasterizing()
    2746
-> 2747         mimage._draw_list_compositing_images(renderer, self, artists)
    2748
    2749         renderer.close_group('axes')

c:\users\computer\miniconda3\lib\site-packages\matplotlib\image.py in
↳ _draw_list_compositing_images(renderer, parent, artists, suppress_composite)
    129     if not_composite or not has_images:
    130         for a in artists:
--> 131             a.draw(renderer)
    132     else:
    133         # Composite any adjacent images together

c:\users\computer\miniconda3\lib\site-packages\matplotlib\artist.py in
↳ draw_wrapper(artist, renderer, *args, **kwargs)
    39         renderer.start_filter()
    40
---> 41         return draw(artist, renderer, *args, **kwargs)
    42     finally:
    43         if artist.get_agg_filter() is not None:

```

```

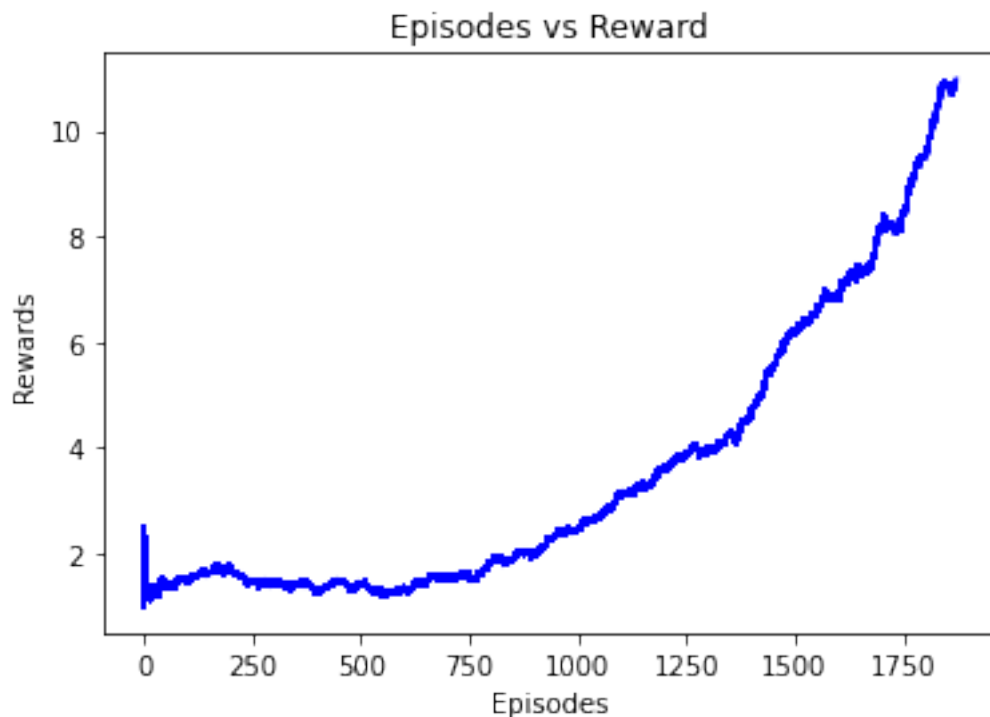
c:\users\computer\miniconda3\lib\site-packages\matplotlib\lines.py in draw(self,
-> renderer)
    784
    785         gc.set_dashes(self._dashOffset, self._dashSeq)
-> 786         renderer.draw_path(gc, tpath, affine.frozen())
    787         gc.restore()
    788

c:\users\computer\miniconda3\lib\site-packages\matplotlib\transforms.py in
-> frozen(self)
    1765     def frozen(self):
    1766         # docstring inherited
-> 1767         return Affine2D(self.get_matrix().copy())
    1768
    1769     @property

c:\users\computer\miniconda3\lib\site-packages\matplotlib\transforms.py in
-> __init__(self, matrix, **kwargs)
    1844         # A bit faster than np.identity(3).
    1845         matrix = IdentityTransform._mtx.copy()
-> 1846         self._mtx = matrix.copy()
    1847         self._invalid = 0
    1848

```

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.

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

[7]: 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
```

```
-----
ModuleNotFoundError                                Traceback (most recent call last)
<ipython-input-7-283b4431c844> in <module>
      7 from IPython import display as ipythondisplay
      8
----> 9 from pyvirtualdisplay import Display
     10
     11 # Displaying the game live

ModuleNotFoundError: No module named 'pyvirtualdisplay'
```

```
[ ]: 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:
    action = agent.get_action(np.float32(history[: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
score += reward

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

```