

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 = False # set to True if using double DQN agent

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

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

1.2.1 Main Training Loop

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

```

[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_dqn.png") # save graph for
→training visualization

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

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

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

```

```

episode: 0   score: 0.0   memory length: 122   epsilon: 1.0   steps: 122   lr:
0.0001   evaluation reward: 0.0
episode: 1   score: 2.0   memory length: 341   epsilon: 1.0   steps: 219   lr:
0.0001   evaluation reward: 1.0
episode: 2   score: 1.0   memory length: 509   epsilon: 1.0   steps: 168   lr:
0.0001   evaluation reward: 1.0
episode: 3   score: 1.0   memory length: 661   epsilon: 1.0   steps: 152   lr:
0.0001   evaluation reward: 1.0
episode: 4   score: 0.0   memory length: 783   epsilon: 1.0   steps: 122   lr:
0.0001   evaluation reward: 0.8
episode: 5   score: 4.0   memory length: 1076   epsilon: 1.0   steps: 293
lr: 0.0001   evaluation reward: 1.3333333333333333
episode: 6   score: 2.0   memory length: 1292   epsilon: 1.0   steps: 216
lr: 0.0001   evaluation reward: 1.4285714285714286
episode: 7   score: 0.0   memory length: 1415   epsilon: 1.0   steps: 123
lr: 0.0001   evaluation reward: 1.25
episode: 8   score: 0.0   memory length: 1538   epsilon: 1.0   steps: 123
lr: 0.0001   evaluation reward: 1.1111111111111112
episode: 9   score: 3.0   memory length: 1764   epsilon: 1.0   steps: 226
lr: 0.0001   evaluation reward: 1.3
episode: 10  score: 2.0   memory length: 1964   epsilon: 1.0   steps: 200

```

lr: 0.0001	evaluation reward: 1.3636363636363635			
episode: 11	score: 2.0	memory length: 2161	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.4166666666666667			
episode: 12	score: 1.0	memory length: 2331	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.3846153846153846			
episode: 13	score: 0.0	memory length: 2453	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.2857142857142858			
episode: 14	score: 2.0	memory length: 2674	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.3333333333333333			
episode: 15	score: 0.0	memory length: 2797	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.25			
episode: 16	score: 3.0	memory length: 3046	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.3529411764705883			
episode: 17	score: 2.0	memory length: 3228	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.3888888888888888			
episode: 18	score: 0.0	memory length: 3351	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.3157894736842106			
episode: 19	score: 2.0	memory length: 3569	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.35			
episode: 20	score: 0.0	memory length: 3692	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.2857142857142858			
episode: 21	score: 1.0	memory length: 3862	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.2727272727272727			
episode: 22	score: 4.0	memory length: 4156	epsilon: 1.0	steps: 294
lr: 0.0001	evaluation reward: 1.391304347826087			
episode: 23	score: 2.0	memory length: 4354	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4166666666666667			
episode: 24	score: 1.0	memory length: 4504	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4			
episode: 25	score: 3.0	memory length: 4771	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.4615384615384615			
episode: 26	score: 4.0	memory length: 5065	epsilon: 1.0	steps: 294
lr: 0.0001	evaluation reward: 1.5555555555555556			
episode: 27	score: 4.0	memory length: 5364	epsilon: 1.0	steps: 299
lr: 0.0001	evaluation reward: 1.6428571428571428			
episode: 28	score: 1.0	memory length: 5515	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.6206896551724137			
episode: 29	score: 5.0	memory length: 5818	epsilon: 1.0	steps: 303
lr: 0.0001	evaluation reward: 1.7333333333333334			
episode: 30	score: 1.0	memory length: 5988	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.7096774193548387			
episode: 31	score: 1.0	memory length: 6158	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.6875			
episode: 32	score: 4.0	memory length: 6451	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.7575757575757576			
episode: 33	score: 0.0	memory length: 6574	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7058823529411764			
episode: 34	score: 2.0	memory length: 6790	epsilon: 1.0	steps: 216

lr: 0.0001	evaluation reward: 1.7142857142857142		
episode: 35	score: 2.0	memory length: 7007	epsilon: 1.0 steps: 217
lr: 0.0001	evaluation reward: 1.7222222222222223		
episode: 36	score: 2.0	memory length: 7207	epsilon: 1.0 steps: 200
lr: 0.0001	evaluation reward: 1.7297297297297298		
episode: 37	score: 3.0	memory length: 7433	epsilon: 1.0 steps: 226
lr: 0.0001	evaluation reward: 1.763157894736842		
episode: 38	score: 0.0	memory length: 7556	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7179487179487178		
episode: 39	score: 3.0	memory length: 7821	epsilon: 1.0 steps: 265
lr: 0.0001	evaluation reward: 1.75		
episode: 40	score: 1.0	memory length: 7993	epsilon: 1.0 steps: 172
lr: 0.0001	evaluation reward: 1.7317073170731707		
episode: 41	score: 2.0	memory length: 8191	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.7380952380952381		
episode: 42	score: 0.0	memory length: 8314	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.697674418604651		
episode: 43	score: 0.0	memory length: 8437	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.6590909090909092		
episode: 44	score: 1.0	memory length: 8588	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.6444444444444444		
episode: 45	score: 2.0	memory length: 8785	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.6521739130434783		
episode: 46	score: 3.0	memory length: 9015	epsilon: 1.0 steps: 230
lr: 0.0001	evaluation reward: 1.6808510638297873		
episode: 47	score: 6.0	memory length: 9353	epsilon: 1.0 steps: 338
lr: 0.0001	evaluation reward: 1.7708333333333333		
episode: 48	score: 3.0	memory length: 9602	epsilon: 1.0 steps: 249
lr: 0.0001	evaluation reward: 1.7959183673469388		
episode: 49	score: 2.0	memory length: 9803	epsilon: 1.0 steps: 201
lr: 0.0001	evaluation reward: 1.8		
episode: 50	score: 0.0	memory length: 9926	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7647058823529411		
episode: 51	score: 2.0	memory length: 10124	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.7692307692307692		
episode: 52	score: 2.0	memory length: 10342	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.7735849056603774		
episode: 53	score: 3.0	memory length: 10588	epsilon: 1.0 steps: 246
lr: 0.0001	evaluation reward: 1.7962962962962963		
episode: 54	score: 0.0	memory length: 10711	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7636363636363637		
episode: 55	score: 0.0	memory length: 10834	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7321428571428572		
episode: 56	score: 0.0	memory length: 10956	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.7017543859649122		
episode: 57	score: 1.0	memory length: 11127	epsilon: 1.0 steps: 171
lr: 0.0001	evaluation reward: 1.6896551724137931		
episode: 58	score: 1.0	memory length: 11280	epsilon: 1.0 steps: 153

lr: 0.0001	evaluation reward: 1.6779661016949152		
episode: 59	score: 2.0	memory length: 11502	epsilon: 1.0 steps: 222
lr: 0.0001	evaluation reward: 1.6833333333333333		
episode: 60	score: 2.0	memory length: 11722	epsilon: 1.0 steps: 220
lr: 0.0001	evaluation reward: 1.6885245901639345		
episode: 61	score: 1.0	memory length: 11872	epsilon: 1.0 steps: 150
lr: 0.0001	evaluation reward: 1.6774193548387097		
episode: 62	score: 2.0	memory length: 12076	epsilon: 1.0 steps: 204
lr: 0.0001	evaluation reward: 1.6825396825396826		
episode: 63	score: 0.0	memory length: 12199	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.65625		
episode: 64	score: 2.0	memory length: 12396	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.6615384615384616		
episode: 65	score: 1.0	memory length: 12546	epsilon: 1.0 steps: 150
lr: 0.0001	evaluation reward: 1.6515151515151516		
episode: 66	score: 3.0	memory length: 12795	epsilon: 1.0 steps: 249
lr: 0.0001	evaluation reward: 1.671641791044776		
episode: 67	score: 1.0	memory length: 12964	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.661764705882353		
episode: 68	score: 0.0	memory length: 13087	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.6376811594202898		
episode: 69	score: 1.0	memory length: 13238	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.6285714285714286		
episode: 70	score: 3.0	memory length: 13464	epsilon: 1.0 steps: 226
lr: 0.0001	evaluation reward: 1.647887323943662		
episode: 71	score: 0.0	memory length: 13587	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.625		
episode: 72	score: 1.0	memory length: 13756	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.6164383561643836		
episode: 73	score: 0.0	memory length: 13879	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5945945945945945		
episode: 74	score: 0.0	memory length: 14002	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5733333333333333		
episode: 75	score: 0.0	memory length: 14124	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.5526315789473684		
episode: 76	score: 2.0	memory length: 14344	epsilon: 1.0 steps: 220
lr: 0.0001	evaluation reward: 1.5584415584415585		
episode: 77	score: 1.0	memory length: 14512	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.5512820512820513		
episode: 78	score: 0.0	memory length: 14635	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5316455696202531		
episode: 79	score: 3.0	memory length: 14903	epsilon: 1.0 steps: 268
lr: 0.0001	evaluation reward: 1.55		
episode: 80	score: 4.0	memory length: 15181	epsilon: 1.0 steps: 278
lr: 0.0001	evaluation reward: 1.5802469135802468		
episode: 81	score: 0.0	memory length: 15303	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.5609756097560976		
episode: 82	score: 1.0	memory length: 15475	epsilon: 1.0 steps: 172

lr: 0.0001	evaluation reward: 1.5542168674698795		
episode: 83	score: 2.0	memory length: 15673	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.5595238095238095		
episode: 84	score: 0.0	memory length: 15796	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5411764705882354		
episode: 85	score: 0.0	memory length: 15919	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5232558139534884		
episode: 86	score: 2.0	memory length: 16117	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.528735632183908		
episode: 87	score: 1.0	memory length: 16285	epsilon: 1.0 steps: 168
lr: 0.0001	evaluation reward: 1.5227272727272727		
episode: 88	score: 2.0	memory length: 16509	epsilon: 1.0 steps: 224
lr: 0.0001	evaluation reward: 1.5280898876404494		
episode: 89	score: 3.0	memory length: 16776	epsilon: 1.0 steps: 267
lr: 0.0001	evaluation reward: 1.5444444444444445		
episode: 90	score: 2.0	memory length: 16974	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.5494505494505495		
episode: 91	score: 0.0	memory length: 17097	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5326086956521738		
episode: 92	score: 1.0	memory length: 17248	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.5268817204301075		
episode: 93	score: 0.0	memory length: 17371	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5106382978723405		
episode: 94	score: 0.0	memory length: 17493	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.4947368421052631		
episode: 95	score: 2.0	memory length: 17711	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.5		
episode: 96	score: 2.0	memory length: 17909	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.5051546391752577		
episode: 97	score: 0.0	memory length: 18032	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.489795918367347		
episode: 98	score: 0.0	memory length: 18155	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4747474747474747		
episode: 99	score: 0.0	memory length: 18277	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.46		
episode: 100	score: 0.0	memory length: 18399	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.46		
episode: 101	score: 2.0	memory length: 18618	epsilon: 1.0 steps: 219
lr: 0.0001	evaluation reward: 1.46		
episode: 102	score: 2.0	memory length: 18816	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 103	score: 0.0	memory length: 18939	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 104	score: 4.0	memory length: 19217	epsilon: 1.0 steps: 278
lr: 0.0001	evaluation reward: 1.5		
episode: 105	score: 1.0	memory length: 19369	epsilon: 1.0 steps: 152
lr: 0.0001	evaluation reward: 1.47		
episode: 106	score: 4.0	memory length: 19686	epsilon: 1.0 steps: 317

lr: 0.0001	evaluation reward: 1.49			
episode: 107	score: 2.0	memory length: 19904	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.51			
episode: 108	score: 0.0	memory length: 20027	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51			
episode: 109	score: 0.0	memory length: 20150	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 110	score: 3.0	memory length: 20379	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.49			
episode: 111	score: 1.0	memory length: 20551	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.48			
episode: 112	score: 2.0	memory length: 20766	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.49			
episode: 113	score: 0.0	memory length: 20888	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.49			
episode: 114	score: 1.0	memory length: 21056	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.48			
episode: 115	score: 0.0	memory length: 21178	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.48			
episode: 116	score: 3.0	memory length: 21403	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.48			
episode: 117	score: 3.0	memory length: 21672	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.49			
episode: 118	score: 8.0	memory length: 22162	epsilon: 1.0	steps: 490
lr: 0.0001	evaluation reward: 1.57			
episode: 119	score: 2.0	memory length: 22359	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.57			
episode: 120	score: 0.0	memory length: 22482	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 121	score: 1.0	memory length: 22632	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.57			
episode: 122	score: 2.0	memory length: 22832	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.55			
episode: 123	score: 3.0	memory length: 23106	epsilon: 1.0	steps: 274
lr: 0.0001	evaluation reward: 1.56			
episode: 124	score: 1.0	memory length: 23258	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.56			
episode: 125	score: 4.0	memory length: 23521	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.57			
episode: 126	score: 1.0	memory length: 23671	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.54			
episode: 127	score: 1.0	memory length: 23840	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.51			
episode: 128	score: 0.0	memory length: 23962	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.5			
episode: 129	score: 2.0	memory length: 24159	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.47			
episode: 130	score: 2.0	memory length: 24377	epsilon: 1.0	steps: 218

lr: 0.0001	evaluation reward: 1.48			
episode: 131	score: 2.0	memory length: 24574	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.49			
episode: 132	score: 0.0	memory length: 24697	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45			
episode: 133	score: 2.0	memory length: 24918	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.47			
episode: 134	score: 4.0	memory length: 25213	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.49			
episode: 135	score: 0.0	memory length: 25336	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 136	score: 2.0	memory length: 25533	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.47			
episode: 137	score: 0.0	memory length: 25655	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44			
episode: 138	score: 0.0	memory length: 25777	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44			
episode: 139	score: 3.0	memory length: 26024	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.44			
episode: 140	score: 2.0	memory length: 26222	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 141	score: 2.0	memory length: 26438	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.45			
episode: 142	score: 2.0	memory length: 26638	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.47			
episode: 143	score: 0.0	memory length: 26761	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 144	score: 3.0	memory length: 27026	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.49			
episode: 145	score: 2.0	memory length: 27224	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49			
episode: 146	score: 1.0	memory length: 27395	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.47			
episode: 147	score: 0.0	memory length: 27518	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41			
episode: 148	score: 2.0	memory length: 27733	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.4			
episode: 149	score: 2.0	memory length: 27951	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.4			
episode: 150	score: 1.0	memory length: 28122	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.41			
episode: 151	score: 1.0	memory length: 28290	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.4			
episode: 152	score: 0.0	memory length: 28412	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38			
episode: 153	score: 2.0	memory length: 28609	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.37			
episode: 154	score: 3.0	memory length: 28855	epsilon: 1.0	steps: 246

lr: 0.0001	evaluation reward: 1.4			
episode: 155	score: 3.0	memory length: 29084	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.43			
episode: 156	score: 2.0	memory length: 29304	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.45			
episode: 157	score: 3.0	memory length: 29551	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.47			
episode: 158	score: 0.0	memory length: 29673	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.46			
episode: 159	score: 0.0	memory length: 29795	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44			
episode: 160	score: 0.0	memory length: 29917	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.42			
episode: 161	score: 0.0	memory length: 30040	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41			
episode: 162	score: 1.0	memory length: 30212	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.4			
episode: 163	score: 0.0	memory length: 30335	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4			
episode: 164	score: 0.0	memory length: 30458	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38			
episode: 165	score: 3.0	memory length: 30685	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.4			
episode: 166	score: 0.0	memory length: 30808	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 167	score: 3.0	memory length: 31072	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.39			
episode: 168	score: 3.0	memory length: 31322	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.42			
episode: 169	score: 3.0	memory length: 31550	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.44			
episode: 170	score: 0.0	memory length: 31673	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41			
episode: 171	score: 1.0	memory length: 31843	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.42			
episode: 172	score: 3.0	memory length: 32071	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.44			
episode: 173	score: 1.0	memory length: 32241	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.45			
episode: 174	score: 1.0	memory length: 32412	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.46			
episode: 175	score: 2.0	memory length: 32609	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.48			
episode: 176	score: 1.0	memory length: 32780	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.47			
episode: 177	score: 3.0	memory length: 33025	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.49			
episode: 178	score: 2.0	memory length: 33225	epsilon: 1.0	steps: 200

lr: 0.0001	evaluation reward: 1.51			
episode: 179	score: 1.0	memory length: 33375	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.49			
episode: 180	score: 3.0	memory length: 33607	epsilon: 1.0	steps: 232
lr: 0.0001	evaluation reward: 1.48			
episode: 181	score: 0.0	memory length: 33730	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 182	score: 0.0	memory length: 33853	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 183	score: 3.0	memory length: 34121	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.48			
episode: 184	score: 0.0	memory length: 34243	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.48			
episode: 185	score: 1.0	memory length: 34413	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.49			
episode: 186	score: 0.0	memory length: 34536	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 187	score: 2.0	memory length: 34756	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.48			
episode: 188	score: 3.0	memory length: 35018	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.49			
episode: 189	score: 1.0	memory length: 35188	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.47			
episode: 190	score: 1.0	memory length: 35339	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.46			
episode: 191	score: 0.0	memory length: 35462	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46			
episode: 192	score: 7.0	memory length: 35762	epsilon: 1.0	steps: 300
lr: 0.0001	evaluation reward: 1.52			
episode: 193	score: 2.0	memory length: 35960	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.54			
episode: 194	score: 3.0	memory length: 36208	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.57			
episode: 195	score: 2.0	memory length: 36426	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.57			
episode: 196	score: 0.0	memory length: 36549	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55			
episode: 197	score: 3.0	memory length: 36775	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.58			
episode: 198	score: 0.0	memory length: 36898	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58			
episode: 199	score: 2.0	memory length: 37095	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.6			
episode: 200	score: 1.0	memory length: 37265	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.61			
episode: 201	score: 0.0	memory length: 37387	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.59			
episode: 202	score: 1.0	memory length: 37556	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.58			
episode: 203	score: 1.0	memory length: 37707	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.59			
episode: 204	score: 1.0	memory length: 37876	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.56			
episode: 205	score: 0.0	memory length: 37999	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55			
episode: 206	score: 1.0	memory length: 38167	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.52			
episode: 207	score: 2.0	memory length: 38365	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.52			
episode: 208	score: 1.0	memory length: 38516	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.53			
episode: 209	score: 5.0	memory length: 38854	epsilon: 1.0	steps: 338
lr: 0.0001	evaluation reward: 1.58			
episode: 210	score: 1.0	memory length: 39005	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.56			
episode: 211	score: 1.0	memory length: 39157	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.56			
episode: 212	score: 1.0	memory length: 39308	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.55			
episode: 213	score: 3.0	memory length: 39571	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.58			
episode: 214	score: 0.0	memory length: 39694	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 215	score: 2.0	memory length: 39912	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.59			
episode: 216	score: 3.0	memory length: 40138	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.59			
episode: 217	score: 3.0	memory length: 40387	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.59			
episode: 218	score: 0.0	memory length: 40510	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51			
episode: 219	score: 3.0	memory length: 40741	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.52			
episode: 220	score: 0.0	memory length: 40864	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52			
episode: 221	score: 0.0	memory length: 40987	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51			
episode: 222	score: 0.0	memory length: 41110	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 223	score: 1.0	memory length: 41282	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.47			
episode: 224	score: 2.0	memory length: 41480	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.48			
episode: 225	score: 0.0	memory length: 41602	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44			
episode: 226	score: 0.0	memory length: 41724	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.43			
episode: 227	score: 1.0	memory length: 41895	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.43			
episode: 228	score: 4.0	memory length: 42171	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.47			
episode: 229	score: 1.0	memory length: 42343	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.46			
episode: 230	score: 3.0	memory length: 42568	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.47			
episode: 231	score: 3.0	memory length: 42799	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.48			
episode: 232	score: 2.0	memory length: 43016	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.5			
episode: 233	score: 4.0	memory length: 43292	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.52			
episode: 234	score: 1.0	memory length: 43461	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.49			
episode: 235	score: 2.0	memory length: 43659	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.51			
episode: 236	score: 1.0	memory length: 43828	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5			
episode: 237	score: 2.0	memory length: 44025	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.52			
episode: 238	score: 3.0	memory length: 44236	epsilon: 1.0	steps: 211
lr: 0.0001	evaluation reward: 1.55			
episode: 239	score: 1.0	memory length: 44405	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.53			
episode: 240	score: 0.0	memory length: 44528	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51			
episode: 241	score: 0.0	memory length: 44651	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 242	score: 1.0	memory length: 44801	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.48			
episode: 243	score: 1.0	memory length: 44952	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.49			
episode: 244	score: 3.0	memory length: 45220	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.49			
episode: 245	score: 0.0	memory length: 45343	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 246	score: 0.0	memory length: 45465	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.46			
episode: 247	score: 1.0	memory length: 45616	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.47			
episode: 248	score: 1.0	memory length: 45766	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.46			
episode: 249	score: 1.0	memory length: 45935	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.45			
episode: 250	score: 4.0	memory length: 46212	epsilon: 1.0	steps: 277

lr: 0.0001	evaluation reward: 1.48			
episode: 251	score: 0.0	memory length: 46335	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 252	score: 2.0	memory length: 46532	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.49			
episode: 253	score: 2.0	memory length: 46732	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.49			
episode: 254	score: 0.0	memory length: 46855	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46			
episode: 255	score: 2.0	memory length: 47052	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.45			
episode: 256	score: 3.0	memory length: 47278	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.46			
episode: 257	score: 2.0	memory length: 47475	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.45			
episode: 258	score: 5.0	memory length: 47802	epsilon: 1.0	steps: 327
lr: 0.0001	evaluation reward: 1.5			
episode: 259	score: 2.0	memory length: 48021	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.52			
episode: 260	score: 0.0	memory length: 48143	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.52			
episode: 261	score: 2.0	memory length: 48362	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.54			
episode: 262	score: 3.0	memory length: 48609	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.56			
episode: 263	score: 0.0	memory length: 48732	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56			
episode: 264	score: 1.0	memory length: 48882	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.57			
episode: 265	score: 1.0	memory length: 49032	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.55			
episode: 266	score: 2.0	memory length: 49230	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.57			
episode: 267	score: 1.0	memory length: 49381	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.55			
episode: 268	score: 2.0	memory length: 49598	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.54			
episode: 269	score: 1.0	memory length: 49767	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.52			
episode: 270	score: 1.0	memory length: 49938	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.53			
episode: 271	score: 0.0	memory length: 50061	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52			
episode: 272	score: 2.0	memory length: 50260	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.51			
episode: 273	score: 3.0	memory length: 50505	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.53			
episode: 274	score: 1.0	memory length: 50656	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.53			
episode: 275	score: 6.0	memory length: 51005	epsilon: 1.0	steps: 349
lr: 0.0001	evaluation reward: 1.57			
episode: 276	score: 0.0	memory length: 51128	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56			
episode: 277	score: 1.0	memory length: 51279	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.54			
episode: 278	score: 2.0	memory length: 51477	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.54			
episode: 279	score: 3.0	memory length: 51723	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.56			
episode: 280	score: 0.0	memory length: 51845	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.53			
episode: 281	score: 2.0	memory length: 52042	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.55			
episode: 282	score: 3.0	memory length: 52304	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.58			
episode: 283	score: 2.0	memory length: 52521	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.57			
episode: 284	score: 0.0	memory length: 52644	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 285	score: 2.0	memory length: 52843	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.58			
episode: 286	score: 0.0	memory length: 52965	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.58			
episode: 287	score: 2.0	memory length: 53183	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.58			
episode: 288	score: 3.0	memory length: 53426	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.58			
episode: 289	score: 2.0	memory length: 53623	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.59			
episode: 290	score: 3.0	memory length: 53848	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.61			
episode: 291	score: 0.0	memory length: 53971	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61			
episode: 292	score: 1.0	memory length: 54142	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.55			
episode: 293	score: 1.0	memory length: 54293	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.54			
episode: 294	score: 3.0	memory length: 54544	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.54			
episode: 295	score: 3.0	memory length: 54810	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.55			
episode: 296	score: 3.0	memory length: 55054	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.58			
episode: 297	score: 3.0	memory length: 55318	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.58			
episode: 298	score: 4.0	memory length: 55591	epsilon: 1.0	steps: 273

lr: 0.0001	evaluation reward: 1.62			
episode: 299	score: 0.0	memory length: 55714	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6			
episode: 300	score: 2.0	memory length: 55932	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.61			
episode: 301	score: 2.0	memory length: 56131	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.63			
episode: 302	score: 3.0	memory length: 56362	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.65			
episode: 303	score: 1.0	memory length: 56530	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.65			
episode: 304	score: 3.0	memory length: 56793	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.67			
episode: 305	score: 3.0	memory length: 57020	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.7			
episode: 306	score: 2.0	memory length: 57218	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.71			
episode: 307	score: 0.0	memory length: 57341	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.69			
episode: 308	score: 2.0	memory length: 57539	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.7			
episode: 309	score: 0.0	memory length: 57662	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 310	score: 2.0	memory length: 57880	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.66			
episode: 311	score: 0.0	memory length: 58003	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 312	score: 5.0	memory length: 58368	epsilon: 1.0	steps: 365
lr: 0.0001	evaluation reward: 1.69			
episode: 313	score: 2.0	memory length: 58566	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.68			
episode: 314	score: 1.0	memory length: 58736	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.69			
episode: 315	score: 2.0	memory length: 58915	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.69			
episode: 316	score: 0.0	memory length: 59037	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.66			
episode: 317	score: 0.0	memory length: 59159	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.63			
episode: 318	score: 3.0	memory length: 59403	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.66			
episode: 319	score: 2.0	memory length: 59603	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.65			
episode: 320	score: 2.0	memory length: 59820	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.67			
episode: 321	score: 0.0	memory length: 59942	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.67			
episode: 322	score: 2.0	memory length: 60161	epsilon: 1.0	steps: 219

lr: 0.0001	evaluation reward: 1.69			
episode: 323	score: 1.0	memory length: 60312	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.69			
episode: 324	score: 2.0	memory length: 60510	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.69			
episode: 325	score: 0.0	memory length: 60632	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.69			
episode: 326	score: 0.0	memory length: 60755	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.69			
episode: 327	score: 3.0	memory length: 61024	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.71			
episode: 328	score: 2.0	memory length: 61222	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.69			
episode: 329	score: 0.0	memory length: 61344	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.68			
episode: 330	score: 0.0	memory length: 61467	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 331	score: 2.0	memory length: 61666	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.64			
episode: 332	score: 2.0	memory length: 61865	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.64			
episode: 333	score: 3.0	memory length: 62132	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.63			
episode: 334	score: 3.0	memory length: 62362	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.65			
episode: 335	score: 0.0	memory length: 62484	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.63			
episode: 336	score: 3.0	memory length: 62731	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.65			
episode: 337	score: 2.0	memory length: 62929	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65			
episode: 338	score: 3.0	memory length: 63175	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.65			
episode: 339	score: 0.0	memory length: 63297	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.64			
episode: 340	score: 0.0	memory length: 63419	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.64			
episode: 341	score: 0.0	memory length: 63542	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 342	score: 1.0	memory length: 63711	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.64			
episode: 343	score: 1.0	memory length: 63880	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.64			
episode: 344	score: 2.0	memory length: 64077	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.63			
episode: 345	score: 2.0	memory length: 64274	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.65			
episode: 346	score: 3.0	memory length: 64537	epsilon: 1.0	steps: 263

lr: 0.0001	evaluation reward: 1.68			
episode: 347	score: 0.0	memory length: 64660	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67			
episode: 348	score: 0.0	memory length: 64782	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.66			
episode: 349	score: 5.0	memory length: 65125	epsilon: 1.0	steps: 343
lr: 0.0001	evaluation reward: 1.7			
episode: 350	score: 2.0	memory length: 65328	epsilon: 1.0	steps: 203
lr: 0.0001	evaluation reward: 1.68			
episode: 351	score: 1.0	memory length: 65497	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.69			
episode: 352	score: 1.0	memory length: 65666	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.68			
episode: 353	score: 2.0	memory length: 65884	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.68			
episode: 354	score: 5.0	memory length: 66188	epsilon: 1.0	steps: 304
lr: 0.0001	evaluation reward: 1.73			
episode: 355	score: 1.0	memory length: 66357	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.72			
episode: 356	score: 3.0	memory length: 66582	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.72			
episode: 357	score: 0.0	memory length: 66705	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7			
episode: 358	score: 3.0	memory length: 66951	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.68			
episode: 359	score: 0.0	memory length: 67074	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66			
episode: 360	score: 2.0	memory length: 67272	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.68			
episode: 361	score: 4.0	memory length: 67550	epsilon: 1.0	steps: 278
lr: 0.0001	evaluation reward: 1.7			
episode: 362	score: 2.0	memory length: 67767	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.69			
episode: 363	score: 3.0	memory length: 67994	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.72			
episode: 364	score: 0.0	memory length: 68117	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.71			
episode: 365	score: 2.0	memory length: 68315	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.72			
episode: 366	score: 1.0	memory length: 68484	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.71			
episode: 367	score: 3.0	memory length: 68710	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.73			
episode: 368	score: 2.0	memory length: 68908	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.73			
episode: 369	score: 1.0	memory length: 69059	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.73			
episode: 370	score: 2.0	memory length: 69259	epsilon: 1.0	steps: 200

lr: 0.0001	evaluation reward: 1.74			
episode: 371	score: 1.0	memory length: 69429	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.75			
episode: 372	score: 0.0	memory length: 69552	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.73			
episode: 373	score: 0.0	memory length: 69675	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7			
episode: 374	score: 3.0	memory length: 69900	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.72			
episode: 375	score: 0.0	memory length: 70023	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66			
episode: 376	score: 0.0	memory length: 70146	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.66			
episode: 377	score: 2.0	memory length: 70344	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.67			
episode: 378	score: 0.0	memory length: 70467	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 379	score: 1.0	memory length: 70617	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.63			
episode: 380	score: 2.0	memory length: 70814	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.65			
episode: 381	score: 1.0	memory length: 70986	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.64			
episode: 382	score: 2.0	memory length: 71184	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.63			
episode: 383	score: 0.0	memory length: 71307	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61			
episode: 384	score: 1.0	memory length: 71477	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.62			
episode: 385	score: 0.0	memory length: 71599	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.6			
episode: 386	score: 0.0	memory length: 71722	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6			
episode: 387	score: 3.0	memory length: 71968	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.61			
episode: 388	score: 2.0	memory length: 72165	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.6			
episode: 389	score: 3.0	memory length: 72413	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.61			
episode: 390	score: 0.0	memory length: 72536	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58			
episode: 391	score: 2.0	memory length: 72755	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.6			
episode: 392	score: 0.0	memory length: 72877	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.59			
episode: 393	score: 0.0	memory length: 73000	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58			
episode: 394	score: 2.0	memory length: 73198	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.57			
episode: 395	score: 2.0	memory length: 73396	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.56			
episode: 396	score: 3.0	memory length: 73647	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.56			
episode: 397	score: 1.0	memory length: 73798	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.54			
episode: 398	score: 2.0	memory length: 73996	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.52			
episode: 399	score: 0.0	memory length: 74119	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52			
episode: 400	score: 2.0	memory length: 74317	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.52			
episode: 401	score: 6.0	memory length: 74657	epsilon: 1.0	steps: 340
lr: 0.0001	evaluation reward: 1.56			
episode: 402	score: 3.0	memory length: 74924	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.56			
episode: 403	score: 1.0	memory length: 75074	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.56			
episode: 404	score: 0.0	memory length: 75197	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53			
episode: 405	score: 2.0	memory length: 75415	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.52			
episode: 406	score: 0.0	memory length: 75537	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.5			
episode: 407	score: 2.0	memory length: 75734	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.52			
episode: 408	score: 2.0	memory length: 75931	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.52			
episode: 409	score: 3.0	memory length: 76164	epsilon: 1.0	steps: 233
lr: 0.0001	evaluation reward: 1.55			
episode: 410	score: 0.0	memory length: 76286	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.53			
episode: 411	score: 3.0	memory length: 76533	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.56			
episode: 412	score: 1.0	memory length: 76684	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.52			
episode: 413	score: 1.0	memory length: 76853	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.51			
episode: 414	score: 1.0	memory length: 77003	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51			
episode: 415	score: 2.0	memory length: 77218	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.51			
episode: 416	score: 2.0	memory length: 77415	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.53			
episode: 417	score: 2.0	memory length: 77633	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.55			
episode: 418	score: 1.0	memory length: 77803	epsilon: 1.0	steps: 170

lr: 0.0001	evaluation reward: 1.53			
episode: 419	score: 1.0	memory length: 77973	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.52			
episode: 420	score: 2.0	memory length: 78192	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.52			
episode: 421	score: 3.0	memory length: 78418	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.55			
episode: 422	score: 2.0	memory length: 78638	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.55			
episode: 423	score: 0.0	memory length: 78761	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.54			
episode: 424	score: 3.0	memory length: 79028	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.55			
episode: 425	score: 0.0	memory length: 79151	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55			
episode: 426	score: 1.0	memory length: 79319	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.56			
episode: 427	score: 1.0	memory length: 79490	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.54			
episode: 428	score: 1.0	memory length: 79640	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.53			
episode: 429	score: 0.0	memory length: 79763	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53			
episode: 430	score: 0.0	memory length: 79885	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.53			
episode: 431	score: 1.0	memory length: 80036	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.52			
episode: 432	score: 0.0	memory length: 80159	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5			
episode: 433	score: 1.0	memory length: 80329	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.48			
episode: 434	score: 1.0	memory length: 80500	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.46			
episode: 435	score: 1.0	memory length: 80671	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.47			
episode: 436	score: 0.0	memory length: 80794	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 437	score: 0.0	memory length: 80917	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42			
episode: 438	score: 1.0	memory length: 81088	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.4			
episode: 439	score: 3.0	memory length: 81318	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.43			
episode: 440	score: 2.0	memory length: 81516	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45			
episode: 441	score: 1.0	memory length: 81666	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.46			
episode: 442	score: 4.0	memory length: 81959	epsilon: 1.0	steps: 293

lr: 0.0001	evaluation reward: 1.49			
episode: 443	score: 3.0	memory length: 82190	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.51			
episode: 444	score: 2.0	memory length: 82408	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.51			
episode: 445	score: 0.0	memory length: 82531	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 446	score: 1.0	memory length: 82681	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.47			
episode: 447	score: 0.0	memory length: 82804	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47			
episode: 448	score: 2.0	memory length: 83001	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.49			
episode: 449	score: 5.0	memory length: 83333	epsilon: 1.0	steps: 332
lr: 0.0001	evaluation reward: 1.49			
episode: 450	score: 1.0	memory length: 83484	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48			
episode: 451	score: 0.0	memory length: 83606	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.47			
episode: 452	score: 0.0	memory length: 83729	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46			
episode: 453	score: 0.0	memory length: 83852	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44			
episode: 454	score: 0.0	memory length: 83974	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39			
episode: 455	score: 1.0	memory length: 84125	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.39			
episode: 456	score: 3.0	memory length: 84392	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.39			
episode: 457	score: 2.0	memory length: 84612	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.41			
episode: 458	score: 0.0	memory length: 84735	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38			
episode: 459	score: 3.0	memory length: 84969	epsilon: 1.0	steps: 234
lr: 0.0001	evaluation reward: 1.41			
episode: 460	score: 1.0	memory length: 85120	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.4			
episode: 461	score: 1.0	memory length: 85291	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.37			
episode: 462	score: 1.0	memory length: 85442	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.36			
episode: 463	score: 3.0	memory length: 85707	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.36			
episode: 464	score: 0.0	memory length: 85829	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36			
episode: 465	score: 1.0	memory length: 85999	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.35			
episode: 466	score: 0.0	memory length: 86121	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.34			
episode: 467	score: 2.0	memory length: 86319	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.33			
episode: 468	score: 0.0	memory length: 86442	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31			
episode: 469	score: 0.0	memory length: 86564	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3			
episode: 470	score: 1.0	memory length: 86735	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.29			
episode: 471	score: 1.0	memory length: 86886	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.29			
episode: 472	score: 0.0	memory length: 87008	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.29			
episode: 473	score: 0.0	memory length: 87131	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.29			
episode: 474	score: 0.0	memory length: 87253	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.26			
episode: 475	score: 2.0	memory length: 87450	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.28			
episode: 476	score: 2.0	memory length: 87647	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.3			
episode: 477	score: 2.0	memory length: 87844	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.3			
episode: 478	score: 4.0	memory length: 88137	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.34			
episode: 479	score: 0.0	memory length: 88259	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.33			
episode: 480	score: 0.0	memory length: 88382	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31			
episode: 481	score: 3.0	memory length: 88628	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.33			
episode: 482	score: 3.0	memory length: 88876	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.34			
episode: 483	score: 1.0	memory length: 89027	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.35			
episode: 484	score: 4.0	memory length: 89313	epsilon: 1.0	steps: 286
lr: 0.0001	evaluation reward: 1.38			
episode: 485	score: 2.0	memory length: 89511	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4			
episode: 486	score: 0.0	memory length: 89633	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4			
episode: 487	score: 1.0	memory length: 89804	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.38			
episode: 488	score: 3.0	memory length: 90050	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.39			
episode: 489	score: 0.0	memory length: 90173	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36			
episode: 490	score: 1.0	memory length: 90324	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.37			
episode: 491	score: 2.0	memory length: 90526	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.37			
episode: 492	score: 0.0	memory length: 90648	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37			
episode: 493	score: 2.0	memory length: 90846	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.39			
episode: 494	score: 2.0	memory length: 91044	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.39			
episode: 495	score: 2.0	memory length: 91243	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.39			
episode: 496	score: 4.0	memory length: 91521	epsilon: 1.0	steps: 278
lr: 0.0001	evaluation reward: 1.4			
episode: 497	score: 3.0	memory length: 91749	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.42			
episode: 498	score: 1.0	memory length: 91918	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.41			
episode: 499	score: 2.0	memory length: 92136	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.43			
episode: 500	score: 2.0	memory length: 92353	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.43			
episode: 501	score: 2.0	memory length: 92551	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.39			
episode: 502	score: 3.0	memory length: 92777	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.39			
episode: 503	score: 2.0	memory length: 92995	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.4			
episode: 504	score: 0.0	memory length: 93117	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4			
episode: 505	score: 0.0	memory length: 93239	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38			
episode: 506	score: 0.0	memory length: 93361	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38			
episode: 507	score: 3.0	memory length: 93631	epsilon: 1.0	steps: 270
lr: 0.0001	evaluation reward: 1.39			
episode: 508	score: 4.0	memory length: 93931	epsilon: 1.0	steps: 300
lr: 0.0001	evaluation reward: 1.41			
episode: 509	score: 2.0	memory length: 94148	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.4			
episode: 510	score: 2.0	memory length: 94346	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.42			
episode: 511	score: 2.0	memory length: 94544	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41			
episode: 512	score: 1.0	memory length: 94715	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.41			
episode: 513	score: 2.0	memory length: 94933	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.42			
episode: 514	score: 0.0	memory length: 95056	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.41			
episode: 515	score: 2.0	memory length: 95255	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.41			
episode: 516	score: 0.0	memory length: 95377	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39			
episode: 517	score: 3.0	memory length: 95602	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.4			
episode: 518	score: 0.0	memory length: 95725	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39			
episode: 519	score: 3.0	memory length: 95951	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.41			
episode: 520	score: 1.0	memory length: 96122	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.4			
episode: 521	score: 1.0	memory length: 96291	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.38			
episode: 522	score: 0.0	memory length: 96413	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36			
episode: 523	score: 4.0	memory length: 96693	epsilon: 1.0	steps: 280
lr: 0.0001	evaluation reward: 1.4			
episode: 524	score: 2.0	memory length: 96910	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.39			
episode: 525	score: 3.0	memory length: 97180	epsilon: 1.0	steps: 270
lr: 0.0001	evaluation reward: 1.42			
episode: 526	score: 1.0	memory length: 97350	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.42			
episode: 527	score: 1.0	memory length: 97519	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.42			
episode: 528	score: 0.0	memory length: 97641	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41			
episode: 529	score: 4.0	memory length: 97919	epsilon: 1.0	steps: 278
lr: 0.0001	evaluation reward: 1.45			
episode: 530	score: 1.0	memory length: 98070	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.46			
episode: 531	score: 3.0	memory length: 98316	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.48			
episode: 532	score: 2.0	memory length: 98496	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.5			
episode: 533	score: 0.0	memory length: 98618	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.49			
episode: 534	score: 1.0	memory length: 98786	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.49			
episode: 535	score: 0.0	memory length: 98909	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 536	score: 1.0	memory length: 99077	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.49			
episode: 537	score: 1.0	memory length: 99245	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.5			
episode: 538	score: 1.0	memory length: 99414	epsilon: 1.0	steps: 169

```

lr: 0.0001      evaluation reward: 1.5
episode: 539    score: 2.0    memory length: 99612    epsilon: 1.0    steps: 198
lr: 0.0001      evaluation reward: 1.49
episode: 540    score: 3.0    memory length: 99856    epsilon: 1.0    steps: 244
lr: 0.0001      evaluation reward: 1.5

```

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.py:67:

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: 541    score: 2.0    memory length: 100074    epsilon: 0.9998515000000032
steps: 218      lr: 0.0001      evaluation reward: 1.51
episode: 542    score: 0.0    memory length: 100196    epsilon: 0.9996099400000085
steps: 122      lr: 0.0001      evaluation reward: 1.47
episode: 543    score: 2.0    memory length: 100394    epsilon: 0.9992179000000017
steps: 198      lr: 0.0001      evaluation reward: 1.46
episode: 544    score: 0.0    memory length: 100517    epsilon: 0.99897436000000223
steps: 123      lr: 0.0001      evaluation reward: 1.44
episode: 545    score: 0.0    memory length: 100640    epsilon: 0.99873082000000276
steps: 123      lr: 0.0001      evaluation reward: 1.44
episode: 546    score: 1.0    memory length: 100812    epsilon: 0.9983902600000035
steps: 172      lr: 0.0001      evaluation reward: 1.44
episode: 547    score: 1.0    memory length: 100963    epsilon: 0.99809128000000414
steps: 151      lr: 0.0001      evaluation reward: 1.45
episode: 548    score: 0.0    memory length: 101085    epsilon: 0.99784972000000467
steps: 122      lr: 0.0001      evaluation reward: 1.43
episode: 549    score: 3.0    memory length: 101333    epsilon: 0.99735868000000573
steps: 248      lr: 0.0001      evaluation reward: 1.41
episode: 550    score: 2.0    memory length: 101551    epsilon: 0.99692704000000667
steps: 218      lr: 0.0001      evaluation reward: 1.42
episode: 551    score: 0.0    memory length: 101674    epsilon: 0.9966835000000072
steps: 123      lr: 0.0001      evaluation reward: 1.42
episode: 552    score: 0.0    memory length: 101797    epsilon: 0.99643996000000773
steps: 123      lr: 0.0001      evaluation reward: 1.42
episode: 553    score: 0.0    memory length: 101919    epsilon: 0.99619840000000825
steps: 122      lr: 0.0001      evaluation reward: 1.42
episode: 554    score: 2.0    memory length: 102117    epsilon: 0.9958063600000091
steps: 198      lr: 0.0001      evaluation reward: 1.44
episode: 555    score: 2.0    memory length: 102336    epsilon: 0.99537274000001005
steps: 219      lr: 0.0001      evaluation reward: 1.45

```

episode: 556	score: 3.0	memory length: 102600	epsilon: 0.9948500200001118
steps: 264	lr: 0.0001	evaluation reward: 1.45	
episode: 557	score: 0.0	memory length: 102723	epsilon: 0.9946064800001171
steps: 123	lr: 0.0001	evaluation reward: 1.43	
episode: 558	score: 0.0	memory length: 102845	epsilon: 0.9943649200001223
steps: 122	lr: 0.0001	evaluation reward: 1.43	
episode: 559	score: 2.0	memory length: 103024	epsilon: 0.994010500000013
steps: 179	lr: 0.0001	evaluation reward: 1.42	
episode: 560	score: 4.0	memory length: 103320	epsilon: 0.9934244200001427
steps: 296	lr: 0.0001	evaluation reward: 1.45	
episode: 561	score: 2.0	memory length: 103518	epsilon: 0.9930323800001513
steps: 198	lr: 0.0001	evaluation reward: 1.46	
episode: 562	score: 1.0	memory length: 103669	epsilon: 0.9927334000001578
steps: 151	lr: 0.0001	evaluation reward: 1.46	
episode: 563	score: 1.0	memory length: 103838	epsilon: 0.992398780000165
steps: 169	lr: 0.0001	evaluation reward: 1.44	
episode: 564	score: 1.0	memory length: 104008	epsilon: 0.9920621800001723
steps: 170	lr: 0.0001	evaluation reward: 1.45	
episode: 565	score: 0.0	memory length: 104130	epsilon: 0.9918206200001776
steps: 122	lr: 0.0001	evaluation reward: 1.44	
episode: 566	score: 3.0	memory length: 104377	epsilon: 0.9913315600001882
steps: 247	lr: 0.0001	evaluation reward: 1.47	
episode: 567	score: 0.0	memory length: 104499	epsilon: 0.9910900000001934
steps: 122	lr: 0.0001	evaluation reward: 1.45	
episode: 568	score: 2.0	memory length: 104696	epsilon: 0.9906999400002019
steps: 197	lr: 0.0001	evaluation reward: 1.47	
episode: 569	score: 0.0	memory length: 104819	epsilon: 0.9904564000002072
steps: 123	lr: 0.0001	evaluation reward: 1.47	
episode: 570	score: 0.0	memory length: 104941	epsilon: 0.9902148400002124
steps: 122	lr: 0.0001	evaluation reward: 1.46	
episode: 571	score: 2.0	memory length: 105158	epsilon: 0.9897851800002218
steps: 217	lr: 0.0001	evaluation reward: 1.47	
episode: 572	score: 1.0	memory length: 105326	epsilon: 0.989452540000229
steps: 168	lr: 0.0001	evaluation reward: 1.48	
episode: 573	score: 0.0	memory length: 105449	epsilon: 0.9892090000002343
steps: 123	lr: 0.0001	evaluation reward: 1.48	
episode: 574	score: 3.0	memory length: 105676	epsilon: 0.988759540000244
steps: 227	lr: 0.0001	evaluation reward: 1.51	
episode: 575	score: 2.0	memory length: 105873	epsilon: 0.9883694800002525
steps: 197	lr: 0.0001	evaluation reward: 1.51	
episode: 576	score: 4.0	memory length: 106147	epsilon: 0.9878269600002643
steps: 274	lr: 0.0001	evaluation reward: 1.53	
episode: 577	score: 0.0	memory length: 106270	epsilon: 0.9875834200002696
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 578	score: 1.0	memory length: 106439	epsilon: 0.9872488000002768
steps: 169	lr: 0.0001	evaluation reward: 1.48	
episode: 579	score: 0.0	memory length: 106562	epsilon: 0.9870052600002821
steps: 123	lr: 0.0001	evaluation reward: 1.48	

episode: 580	score: 1.0	memory length: 106712	epsilon: 0.9867082600002886
steps: 150	lr: 0.0001	evaluation reward: 1.49	
episode: 581	score: 0.0	memory length: 106835	epsilon: 0.9864647200002938
steps: 123	lr: 0.0001	evaluation reward: 1.46	
episode: 582	score: 0.0	memory length: 106957	epsilon: 0.9862231600002991
steps: 122	lr: 0.0001	evaluation reward: 1.43	
episode: 583	score: 1.0	memory length: 107108	epsilon: 0.9859241800003056
steps: 151	lr: 0.0001	evaluation reward: 1.43	
episode: 584	score: 1.0	memory length: 107279	epsilon: 0.9855856000003129
steps: 171	lr: 0.0001	evaluation reward: 1.4	
episode: 585	score: 0.0	memory length: 107402	epsilon: 0.9853420600003182
steps: 123	lr: 0.0001	evaluation reward: 1.38	
episode: 586	score: 2.0	memory length: 107620	epsilon: 0.9849104200003276
steps: 218	lr: 0.0001	evaluation reward: 1.4	
episode: 587	score: 2.0	memory length: 107839	epsilon: 0.984476800000337
steps: 219	lr: 0.0001	evaluation reward: 1.41	
episode: 588	score: 3.0	memory length: 108085	epsilon: 0.9839897200003476
steps: 246	lr: 0.0001	evaluation reward: 1.41	
episode: 589	score: 4.0	memory length: 108381	epsilon: 0.9834036400003603
steps: 296	lr: 0.0001	evaluation reward: 1.45	
episode: 590	score: 1.0	memory length: 108553	epsilon: 0.9830630800003677
steps: 172	lr: 0.0001	evaluation reward: 1.45	
episode: 591	score: 6.0	memory length: 108918	epsilon: 0.9823403800003834
steps: 365	lr: 0.0001	evaluation reward: 1.49	
episode: 592	score: 3.0	memory length: 109166	epsilon: 0.981849340000394
steps: 248	lr: 0.0001	evaluation reward: 1.52	
episode: 593	score: 3.0	memory length: 109412	epsilon: 0.9813622600004046
steps: 246	lr: 0.0001	evaluation reward: 1.53	
episode: 594	score: 1.0	memory length: 109582	epsilon: 0.9810256600004119
steps: 170	lr: 0.0001	evaluation reward: 1.52	
episode: 595	score: 1.0	memory length: 109733	epsilon: 0.9807266800004184
steps: 151	lr: 0.0001	evaluation reward: 1.51	
episode: 596	score: 1.0	memory length: 109883	epsilon: 0.9804296800004249
steps: 150	lr: 0.0001	evaluation reward: 1.48	
episode: 597	score: 2.0	memory length: 110106	epsilon: 0.9799881400004344
steps: 223	lr: 0.0001	evaluation reward: 1.47	
episode: 598	score: 0.0	memory length: 110229	epsilon: 0.9797446000004397
steps: 123	lr: 0.0001	evaluation reward: 1.46	
episode: 599	score: 3.0	memory length: 110477	epsilon: 0.9792535600004504
steps: 248	lr: 0.0001	evaluation reward: 1.47	
episode: 600	score: 2.0	memory length: 110675	epsilon: 0.9788615200004589
steps: 198	lr: 0.0001	evaluation reward: 1.47	
episode: 601	score: 4.0	memory length: 110953	epsilon: 0.9783110800004708
steps: 278	lr: 0.0001	evaluation reward: 1.49	
episode: 602	score: 2.0	memory length: 111168	epsilon: 0.9778853800004801
steps: 215	lr: 0.0001	evaluation reward: 1.48	
episode: 603	score: 2.0	memory length: 111387	epsilon: 0.9774517600004895
steps: 219	lr: 0.0001	evaluation reward: 1.48	

episode: 604	score: 2.0	memory length: 111585	epsilon: 0.977059720000498
steps: 198	lr: 0.0001	evaluation reward: 1.5	
episode: 605	score: 1.0	memory length: 111754	epsilon: 0.9767251000005053
steps: 169	lr: 0.0001	evaluation reward: 1.51	
episode: 606	score: 3.0	memory length: 111999	epsilon: 0.9762400000005158
steps: 245	lr: 0.0001	evaluation reward: 1.54	
episode: 607	score: 1.0	memory length: 112170	epsilon: 0.9759014200005232
steps: 171	lr: 0.0001	evaluation reward: 1.52	
episode: 608	score: 1.0	memory length: 112338	epsilon: 0.9755687800005304
steps: 168	lr: 0.0001	evaluation reward: 1.49	
episode: 609	score: 1.0	memory length: 112489	epsilon: 0.9752698000005369
steps: 151	lr: 0.0001	evaluation reward: 1.48	
episode: 610	score: 5.0	memory length: 112793	epsilon: 0.9746678800005499
steps: 304	lr: 0.0001	evaluation reward: 1.51	
episode: 611	score: 0.0	memory length: 112916	epsilon: 0.9744243400005552
steps: 123	lr: 0.0001	evaluation reward: 1.49	
episode: 612	score: 1.0	memory length: 113066	epsilon: 0.9741273400005617
steps: 150	lr: 0.0001	evaluation reward: 1.49	
episode: 613	score: 1.0	memory length: 113216	epsilon: 0.9738303400005681
steps: 150	lr: 0.0001	evaluation reward: 1.48	
episode: 614	score: 0.0	memory length: 113339	epsilon: 0.9735868000005734
steps: 123	lr: 0.0001	evaluation reward: 1.48	
episode: 615	score: 0.0	memory length: 113462	epsilon: 0.9733432600005787
steps: 123	lr: 0.0001	evaluation reward: 1.46	
episode: 616	score: 1.0	memory length: 113631	epsilon: 0.973008640000586
steps: 169	lr: 0.0001	evaluation reward: 1.47	
episode: 617	score: 1.0	memory length: 113800	epsilon: 0.9726740200005932
steps: 169	lr: 0.0001	evaluation reward: 1.45	
episode: 618	score: 0.0	memory length: 113923	epsilon: 0.9724304800005985
steps: 123	lr: 0.0001	evaluation reward: 1.45	
episode: 619	score: 5.0	memory length: 114232	epsilon: 0.9718186600006118
steps: 309	lr: 0.0001	evaluation reward: 1.47	
episode: 620	score: 3.0	memory length: 114458	epsilon: 0.9713711800006215
steps: 226	lr: 0.0001	evaluation reward: 1.49	
episode: 621	score: 1.0	memory length: 114609	epsilon: 0.971072200000628
steps: 151	lr: 0.0001	evaluation reward: 1.49	
episode: 622	score: 0.0	memory length: 114732	epsilon: 0.9708286600006333
steps: 123	lr: 0.0001	evaluation reward: 1.49	
episode: 623	score: 1.0	memory length: 114902	epsilon: 0.9704920600006406
steps: 170	lr: 0.0001	evaluation reward: 1.46	
episode: 624	score: 1.0	memory length: 115070	epsilon: 0.9701594200006478
steps: 168	lr: 0.0001	evaluation reward: 1.45	
episode: 625	score: 2.0	memory length: 115291	epsilon: 0.9697218400006573
steps: 221	lr: 0.0001	evaluation reward: 1.44	
episode: 626	score: 0.0	memory length: 115414	epsilon: 0.9694783000006626
steps: 123	lr: 0.0001	evaluation reward: 1.43	
episode: 627	score: 3.0	memory length: 115661	epsilon: 0.9689892400006732
steps: 247	lr: 0.0001	evaluation reward: 1.45	

episode: 628	score: 1.0	memory length: 115833	epsilon: 0.9686486800006806
steps: 172	lr: 0.0001	evaluation reward: 1.46	
episode: 629	score: 0.0	memory length: 115956	epsilon: 0.9684051400006859
steps: 123	lr: 0.0001	evaluation reward: 1.42	
episode: 630	score: 1.0	memory length: 116124	epsilon: 0.9680725000006931
steps: 168	lr: 0.0001	evaluation reward: 1.42	
episode: 631	score: 5.0	memory length: 116485	epsilon: 0.9673577200007086
steps: 361	lr: 0.0001	evaluation reward: 1.44	
episode: 632	score: 1.0	memory length: 116655	epsilon: 0.9670211200007159
steps: 170	lr: 0.0001	evaluation reward: 1.43	
episode: 633	score: 3.0	memory length: 116900	epsilon: 0.9665360200007265
steps: 245	lr: 0.0001	evaluation reward: 1.46	
episode: 634	score: 1.0	memory length: 117069	epsilon: 0.9662014000007337
steps: 169	lr: 0.0001	evaluation reward: 1.46	
episode: 635	score: 1.0	memory length: 117240	epsilon: 0.9658628200007411
steps: 171	lr: 0.0001	evaluation reward: 1.47	
episode: 636	score: 0.0	memory length: 117363	epsilon: 0.9656192800007464
steps: 123	lr: 0.0001	evaluation reward: 1.46	
episode: 637	score: 0.0	memory length: 117486	epsilon: 0.9653757400007517
steps: 123	lr: 0.0001	evaluation reward: 1.45	
episode: 638	score: 0.0	memory length: 117609	epsilon: 0.965132200000757
steps: 123	lr: 0.0001	evaluation reward: 1.44	
episode: 639	score: 0.0	memory length: 117731	epsilon: 0.9648906400007622
steps: 122	lr: 0.0001	evaluation reward: 1.42	
episode: 640	score: 2.0	memory length: 117930	epsilon: 0.9644966200007707
steps: 199	lr: 0.0001	evaluation reward: 1.41	
episode: 641	score: 2.0	memory length: 118131	epsilon: 0.9640986400007794
steps: 201	lr: 0.0001	evaluation reward: 1.41	
episode: 642	score: 2.0	memory length: 118349	epsilon: 0.9636670000007888
steps: 218	lr: 0.0001	evaluation reward: 1.43	
episode: 643	score: 1.0	memory length: 118501	epsilon: 0.9633660400007953
steps: 152	lr: 0.0001	evaluation reward: 1.42	
episode: 644	score: 2.0	memory length: 118699	epsilon: 0.9629740000008038
steps: 198	lr: 0.0001	evaluation reward: 1.44	
episode: 645	score: 4.0	memory length: 118994	epsilon: 0.9623899000008165
steps: 295	lr: 0.0001	evaluation reward: 1.48	
episode: 646	score: 2.0	memory length: 119192	epsilon: 0.961997860000825
steps: 198	lr: 0.0001	evaluation reward: 1.49	
episode: 647	score: 3.0	memory length: 119459	epsilon: 0.9614692000008365
steps: 267	lr: 0.0001	evaluation reward: 1.51	
episode: 648	score: 0.0	memory length: 119582	epsilon: 0.9612256600008418
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 649	score: 0.0	memory length: 119704	epsilon: 0.960984100000847
steps: 122	lr: 0.0001	evaluation reward: 1.48	
episode: 650	score: 3.0	memory length: 119932	epsilon: 0.9605326600008568
steps: 228	lr: 0.0001	evaluation reward: 1.49	
episode: 651	score: 1.0	memory length: 120085	epsilon: 0.9602297200008634
steps: 153	lr: 0.0001	evaluation reward: 1.5	

episode: 652	score: 1.0	memory length: 120257	epsilon: 0.9598891600008708
steps: 172	lr: 0.0001	evaluation reward: 1.51	
episode: 653	score: 0.0	memory length: 120380	epsilon: 0.959645620000876
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 654	score: 3.0	memory length: 120611	epsilon: 0.959188240000886
steps: 231	lr: 0.0001	evaluation reward: 1.52	
episode: 655	score: 3.0	memory length: 120858	epsilon: 0.9586991800008966
steps: 247	lr: 0.0001	evaluation reward: 1.53	
episode: 656	score: 3.0	memory length: 121103	epsilon: 0.9582140800009071
steps: 245	lr: 0.0001	evaluation reward: 1.53	
episode: 657	score: 3.0	memory length: 121369	epsilon: 0.9576874000009186
steps: 266	lr: 0.0001	evaluation reward: 1.56	
episode: 658	score: 1.0	memory length: 121520	epsilon: 0.9573884200009251
steps: 151	lr: 0.0001	evaluation reward: 1.57	
episode: 659	score: 3.0	memory length: 121766	epsilon: 0.9569013400009356
steps: 246	lr: 0.0001	evaluation reward: 1.58	
episode: 660	score: 5.0	memory length: 122073	epsilon: 0.9562934800009488
steps: 307	lr: 0.0001	evaluation reward: 1.59	
episode: 661	score: 1.0	memory length: 122223	epsilon: 0.9559964800009553
steps: 150	lr: 0.0001	evaluation reward: 1.58	
episode: 662	score: 0.0	memory length: 122346	epsilon: 0.9557529400009606
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 663	score: 3.0	memory length: 122615	epsilon: 0.9552203200009721
steps: 269	lr: 0.0001	evaluation reward: 1.59	
episode: 664	score: 1.0	memory length: 122784	epsilon: 0.9548857000009794
steps: 169	lr: 0.0001	evaluation reward: 1.59	
episode: 665	score: 2.0	memory length: 122982	epsilon: 0.9544936600009879
steps: 198	lr: 0.0001	evaluation reward: 1.61	
episode: 666	score: 0.0	memory length: 123105	epsilon: 0.9542501200009932
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 667	score: 0.0	memory length: 123227	epsilon: 0.9540085600009984
steps: 122	lr: 0.0001	evaluation reward: 1.58	
episode: 668	score: 3.0	memory length: 123473	epsilon: 0.953521480001009
steps: 246	lr: 0.0001	evaluation reward: 1.59	
episode: 669	score: 2.0	memory length: 123693	epsilon: 0.9530858800010185
steps: 220	lr: 0.0001	evaluation reward: 1.61	
episode: 670	score: 1.0	memory length: 123844	epsilon: 0.952786900001025
steps: 151	lr: 0.0001	evaluation reward: 1.62	
episode: 671	score: 0.0	memory length: 123967	epsilon: 0.9525433600010302
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 672	score: 0.0	memory length: 124090	epsilon: 0.9522998200010355
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 673	score: 2.0	memory length: 124287	epsilon: 0.951909760001044
steps: 197	lr: 0.0001	evaluation reward: 1.61	
episode: 674	score: 2.0	memory length: 124485	epsilon: 0.9515177200010525
steps: 198	lr: 0.0001	evaluation reward: 1.6	
episode: 675	score: 3.0	memory length: 124751	epsilon: 0.9509910400010639
steps: 266	lr: 0.0001	evaluation reward: 1.61	

episode: 676	score: 2.0	memory length: 124949	epsilon: 0.9505990000010724
steps: 198	lr: 0.0001	evaluation reward: 1.59	
episode: 677	score: 2.0	memory length: 125146	epsilon: 0.9502089400010809
steps: 197	lr: 0.0001	evaluation reward: 1.61	
episode: 678	score: 0.0	memory length: 125268	epsilon: 0.9499673800010862
steps: 122	lr: 0.0001	evaluation reward: 1.6	
episode: 679	score: 1.0	memory length: 125418	epsilon: 0.9496703800010926
steps: 150	lr: 0.0001	evaluation reward: 1.61	
episode: 680	score: 2.0	memory length: 125615	epsilon: 0.9492803200011011
steps: 197	lr: 0.0001	evaluation reward: 1.62	
episode: 681	score: 0.0	memory length: 125738	epsilon: 0.9490367800011064
steps: 123	lr: 0.0001	evaluation reward: 1.62	
episode: 682	score: 2.0	memory length: 125935	epsilon: 0.9486467200011148
steps: 197	lr: 0.0001	evaluation reward: 1.64	
episode: 683	score: 2.0	memory length: 126133	epsilon: 0.9482546800011233
steps: 198	lr: 0.0001	evaluation reward: 1.65	
episode: 684	score: 2.0	memory length: 126331	epsilon: 0.9478626400011319
steps: 198	lr: 0.0001	evaluation reward: 1.66	
episode: 685	score: 2.0	memory length: 126528	epsilon: 0.9474725800011403
steps: 197	lr: 0.0001	evaluation reward: 1.68	
episode: 686	score: 3.0	memory length: 126772	epsilon: 0.9469894600011508
steps: 244	lr: 0.0001	evaluation reward: 1.69	
episode: 687	score: 2.0	memory length: 126990	epsilon: 0.9465578200011602
steps: 218	lr: 0.0001	evaluation reward: 1.69	
episode: 688	score: 10.0	memory length: 127429	epsilon: 0.945688600001179
steps: 439	lr: 0.0001	evaluation reward: 1.76	
episode: 689	score: 2.0	memory length: 127629	epsilon: 0.9452926000011876
steps: 200	lr: 0.0001	evaluation reward: 1.74	
episode: 690	score: 2.0	memory length: 127827	epsilon: 0.9449005600011962
steps: 198	lr: 0.0001	evaluation reward: 1.75	
episode: 691	score: 1.0	memory length: 127998	epsilon: 0.9445619800012035
steps: 171	lr: 0.0001	evaluation reward: 1.7	
episode: 692	score: 2.0	memory length: 128196	epsilon: 0.944169940001212
steps: 198	lr: 0.0001	evaluation reward: 1.69	
episode: 693	score: 0.0	memory length: 128318	epsilon: 0.9439283800012173
steps: 122	lr: 0.0001	evaluation reward: 1.66	
episode: 694	score: 0.0	memory length: 128440	epsilon: 0.9436868200012225
steps: 122	lr: 0.0001	evaluation reward: 1.65	
episode: 695	score: 2.0	memory length: 128659	epsilon: 0.9432532000012319
steps: 219	lr: 0.0001	evaluation reward: 1.66	
episode: 696	score: 2.0	memory length: 128875	epsilon: 0.9428255200012412
steps: 216	lr: 0.0001	evaluation reward: 1.67	
episode: 697	score: 3.0	memory length: 129100	epsilon: 0.9423800200012509
steps: 225	lr: 0.0001	evaluation reward: 1.68	
episode: 698	score: 0.0	memory length: 129222	epsilon: 0.9421384600012561
steps: 122	lr: 0.0001	evaluation reward: 1.68	
episode: 699	score: 0.0	memory length: 129344	epsilon: 0.9418969000012614
steps: 122	lr: 0.0001	evaluation reward: 1.65	

episode: 700	score: 0.0	memory length: 129466	epsilon: 0.9416553400012666
steps: 122	lr: 0.0001	evaluation reward: 1.63	
episode: 701	score: 3.0	memory length: 129728	epsilon: 0.9411365800012779
steps: 262	lr: 0.0001	evaluation reward: 1.62	
episode: 702	score: 1.0	memory length: 129897	epsilon: 0.9408019600012851
steps: 169	lr: 0.0001	evaluation reward: 1.61	
episode: 703	score: 1.0	memory length: 130048	epsilon: 0.9405029800012916
steps: 151	lr: 0.0001	evaluation reward: 1.6	
episode: 704	score: 2.0	memory length: 130230	epsilon: 0.9401426200012994
steps: 182	lr: 0.0001	evaluation reward: 1.6	
episode: 705	score: 5.0	memory length: 130522	epsilon: 0.939564460001312
steps: 292	lr: 0.0001	evaluation reward: 1.64	
episode: 706	score: 1.0	memory length: 130691	epsilon: 0.9392298400013193
steps: 169	lr: 0.0001	evaluation reward: 1.62	
episode: 707	score: 0.0	memory length: 130814	epsilon: 0.9389863000013245
steps: 123	lr: 0.0001	evaluation reward: 1.61	
episode: 708	score: 1.0	memory length: 130983	epsilon: 0.9386516800013318
steps: 169	lr: 0.0001	evaluation reward: 1.61	
episode: 709	score: 2.0	memory length: 131201	epsilon: 0.9382200400013412
steps: 218	lr: 0.0001	evaluation reward: 1.62	
episode: 710	score: 4.0	memory length: 131476	epsilon: 0.937675540001353
steps: 275	lr: 0.0001	evaluation reward: 1.61	
episode: 711	score: 0.0	memory length: 131598	epsilon: 0.9374339800013582
steps: 122	lr: 0.0001	evaluation reward: 1.61	
episode: 712	score: 2.0	memory length: 131820	epsilon: 0.9369944200013678
steps: 222	lr: 0.0001	evaluation reward: 1.62	
episode: 713	score: 1.0	memory length: 131971	epsilon: 0.9366954400013743
steps: 151	lr: 0.0001	evaluation reward: 1.62	
episode: 714	score: 5.0	memory length: 132287	epsilon: 0.9360697600013879
steps: 316	lr: 0.0001	evaluation reward: 1.67	
episode: 715	score: 0.0	memory length: 132410	epsilon: 0.9358262200013931
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 716	score: 2.0	memory length: 132610	epsilon: 0.9354302200014017
steps: 200	lr: 0.0001	evaluation reward: 1.68	
episode: 717	score: 0.0	memory length: 132733	epsilon: 0.935186680001407
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 718	score: 3.0	memory length: 132960	epsilon: 0.9347372200014168
steps: 227	lr: 0.0001	evaluation reward: 1.7	
episode: 719	score: 3.0	memory length: 133226	epsilon: 0.9342105400014282
steps: 266	lr: 0.0001	evaluation reward: 1.68	
episode: 720	score: 3.0	memory length: 133452	epsilon: 0.9337630600014379
steps: 226	lr: 0.0001	evaluation reward: 1.68	
episode: 721	score: 1.0	memory length: 133622	epsilon: 0.9334264600014452
steps: 170	lr: 0.0001	evaluation reward: 1.68	
episode: 722	score: 1.0	memory length: 133791	epsilon: 0.9330918400014525
steps: 169	lr: 0.0001	evaluation reward: 1.69	
episode: 723	score: 0.0	memory length: 133914	epsilon: 0.9328483000014578
steps: 123	lr: 0.0001	evaluation reward: 1.68	

episode: 724	score: 2.0	memory length: 134130	epsilon: 0.9324206200014671
steps: 216	lr: 0.0001	evaluation reward: 1.69	
episode: 725	score: 2.0	memory length: 134351	epsilon: 0.9319830400014766
steps: 221	lr: 0.0001	evaluation reward: 1.69	
episode: 726	score: 1.0	memory length: 134522	epsilon: 0.9316444600014839
steps: 171	lr: 0.0001	evaluation reward: 1.7	
episode: 727	score: 2.0	memory length: 134740	epsilon: 0.9312128200014933
steps: 218	lr: 0.0001	evaluation reward: 1.69	
episode: 728	score: 0.0	memory length: 134863	epsilon: 0.9309692800014986
steps: 123	lr: 0.0001	evaluation reward: 1.68	
episode: 729	score: 1.0	memory length: 135014	epsilon: 0.9306703000015051
steps: 151	lr: 0.0001	evaluation reward: 1.69	
episode: 730	score: 1.0	memory length: 135183	epsilon: 0.9303356800015123
steps: 169	lr: 0.0001	evaluation reward: 1.69	
episode: 731	score: 2.0	memory length: 135401	epsilon: 0.9299040400015217
steps: 218	lr: 0.0001	evaluation reward: 1.66	
episode: 732	score: 2.0	memory length: 135601	epsilon: 0.9295080400015303
steps: 200	lr: 0.0001	evaluation reward: 1.67	
episode: 733	score: 3.0	memory length: 135846	epsilon: 0.9290229400015408
steps: 245	lr: 0.0001	evaluation reward: 1.67	
episode: 734	score: 1.0	memory length: 136015	epsilon: 0.9286883200015481
steps: 169	lr: 0.0001	evaluation reward: 1.67	
episode: 735	score: 3.0	memory length: 136265	epsilon: 0.9281933200015589
steps: 250	lr: 0.0001	evaluation reward: 1.69	
episode: 736	score: 2.0	memory length: 136462	epsilon: 0.9278032600015673
steps: 197	lr: 0.0001	evaluation reward: 1.71	
episode: 737	score: 2.0	memory length: 136681	epsilon: 0.9273696400015767
steps: 219	lr: 0.0001	evaluation reward: 1.73	
episode: 738	score: 4.0	memory length: 136977	epsilon: 0.9267835600015895
steps: 296	lr: 0.0001	evaluation reward: 1.77	
episode: 739	score: 0.0	memory length: 137099	epsilon: 0.9265420000015947
steps: 122	lr: 0.0001	evaluation reward: 1.77	
episode: 740	score: 1.0	memory length: 137270	epsilon: 0.926203420001602
steps: 171	lr: 0.0001	evaluation reward: 1.76	
episode: 741	score: 2.0	memory length: 137489	epsilon: 0.9257698000016115
steps: 219	lr: 0.0001	evaluation reward: 1.76	
episode: 742	score: 2.0	memory length: 137686	epsilon: 0.9253797400016199
steps: 197	lr: 0.0001	evaluation reward: 1.76	
episode: 743	score: 0.0	memory length: 137808	epsilon: 0.9251381800016252
steps: 122	lr: 0.0001	evaluation reward: 1.75	
episode: 744	score: 1.0	memory length: 137959	epsilon: 0.9248392000016317
steps: 151	lr: 0.0001	evaluation reward: 1.74	
episode: 745	score: 1.0	memory length: 138128	epsilon: 0.9245045800016389
steps: 169	lr: 0.0001	evaluation reward: 1.71	
episode: 746	score: 0.0	memory length: 138251	epsilon: 0.9242610400016442
steps: 123	lr: 0.0001	evaluation reward: 1.69	
episode: 747	score: 2.0	memory length: 138449	epsilon: 0.9238690000016527
steps: 198	lr: 0.0001	evaluation reward: 1.68	

episode: 748	score: 1.0	memory length: 138619	epsilon: 0.92353240000166
steps: 170	lr: 0.0001	evaluation reward: 1.69	
episode: 749	score: 1.0	memory length: 138788	epsilon: 0.9231977800016673
steps: 169	lr: 0.0001	evaluation reward: 1.7	
episode: 750	score: 0.0	memory length: 138911	epsilon: 0.9229542400016726
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 751	score: 0.0	memory length: 139033	epsilon: 0.9227126800016778
steps: 122	lr: 0.0001	evaluation reward: 1.66	
episode: 752	score: 0.0	memory length: 139156	epsilon: 0.9224691400016831
steps: 123	lr: 0.0001	evaluation reward: 1.65	
episode: 753	score: 1.0	memory length: 139325	epsilon: 0.9221345200016904
steps: 169	lr: 0.0001	evaluation reward: 1.66	
episode: 754	score: 1.0	memory length: 139494	epsilon: 0.9217999000016976
steps: 169	lr: 0.0001	evaluation reward: 1.64	
episode: 755	score: 2.0	memory length: 139694	epsilon: 0.9214039000017062
steps: 200	lr: 0.0001	evaluation reward: 1.63	
episode: 756	score: 2.0	memory length: 139891	epsilon: 0.9210138400017147
steps: 197	lr: 0.0001	evaluation reward: 1.62	
episode: 757	score: 0.0	memory length: 140014	epsilon: 0.92077030000172
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 758	score: 0.0	memory length: 140137	epsilon: 0.9205267600017253
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 759	score: 0.0	memory length: 140260	epsilon: 0.9202832200017306
steps: 123	lr: 0.0001	evaluation reward: 1.55	
episode: 760	score: 1.0	memory length: 140431	epsilon: 0.9199446400017379
steps: 171	lr: 0.0001	evaluation reward: 1.51	
episode: 761	score: 0.0	memory length: 140554	epsilon: 0.9197011000017432
steps: 123	lr: 0.0001	evaluation reward: 1.5	
episode: 762	score: 1.0	memory length: 140724	epsilon: 0.9193645000017505
steps: 170	lr: 0.0001	evaluation reward: 1.51	
episode: 763	score: 0.0	memory length: 140847	epsilon: 0.9191209600017558
steps: 123	lr: 0.0001	evaluation reward: 1.48	
episode: 764	score: 0.0	memory length: 140970	epsilon: 0.9188774200017611
steps: 123	lr: 0.0001	evaluation reward: 1.47	
episode: 765	score: 2.0	memory length: 141170	epsilon: 0.9184814200017697
steps: 200	lr: 0.0001	evaluation reward: 1.47	
episode: 766	score: 2.0	memory length: 141351	epsilon: 0.9181230400017775
steps: 181	lr: 0.0001	evaluation reward: 1.49	
episode: 767	score: 2.0	memory length: 141551	epsilon: 0.9177270400017861
steps: 200	lr: 0.0001	evaluation reward: 1.51	
episode: 768	score: 0.0	memory length: 141674	epsilon: 0.9174835000017914
steps: 123	lr: 0.0001	evaluation reward: 1.48	
episode: 769	score: 2.0	memory length: 141892	epsilon: 0.9170518600018007
steps: 218	lr: 0.0001	evaluation reward: 1.48	
episode: 770	score: 5.0	memory length: 142215	epsilon: 0.9164123200018146
steps: 323	lr: 0.0001	evaluation reward: 1.52	
episode: 771	score: 0.0	memory length: 142338	epsilon: 0.9161687800018199
steps: 123	lr: 0.0001	evaluation reward: 1.52	

episode: 772	score: 2.0	memory length: 142556	epsilon: 0.9157371400018293
steps: 218	lr: 0.0001	evaluation reward: 1.54	
episode: 773	score: 1.0	memory length: 142727	epsilon: 0.9153985600018366
steps: 171	lr: 0.0001	evaluation reward: 1.53	
episode: 774	score: 0.0	memory length: 142850	epsilon: 0.9151550200018419
steps: 123	lr: 0.0001	evaluation reward: 1.51	
episode: 775	score: 3.0	memory length: 143076	epsilon: 0.9147075400018516
steps: 226	lr: 0.0001	evaluation reward: 1.51	
episode: 776	score: 1.0	memory length: 143227	epsilon: 0.9144085600018581
steps: 151	lr: 0.0001	evaluation reward: 1.5	
episode: 777	score: 4.0	memory length: 143504	epsilon: 0.91386010000187
steps: 277	lr: 0.0001	evaluation reward: 1.52	
episode: 778	score: 1.0	memory length: 143675	epsilon: 0.9135215200018774
steps: 171	lr: 0.0001	evaluation reward: 1.53	
episode: 779	score: 2.0	memory length: 143873	epsilon: 0.9131294800018859
steps: 198	lr: 0.0001	evaluation reward: 1.54	
episode: 780	score: 2.0	memory length: 144092	epsilon: 0.9126958600018953
steps: 219	lr: 0.0001	evaluation reward: 1.54	
episode: 781	score: 1.0	memory length: 144242	epsilon: 0.9123988600019017
steps: 150	lr: 0.0001	evaluation reward: 1.55	
episode: 782	score: 9.0	memory length: 144655	epsilon: 0.9115811200019195
steps: 413	lr: 0.0001	evaluation reward: 1.62	
episode: 783	score: 0.0	memory length: 144778	epsilon: 0.9113375800019248
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 784	score: 1.0	memory length: 144947	epsilon: 0.911002960001932
steps: 169	lr: 0.0001	evaluation reward: 1.59	
episode: 785	score: 4.0	memory length: 145242	epsilon: 0.9104188600019447
steps: 295	lr: 0.0001	evaluation reward: 1.61	
episode: 786	score: 2.0	memory length: 145425	epsilon: 0.9100565200019526
steps: 183	lr: 0.0001	evaluation reward: 1.6	
episode: 787	score: 3.0	memory length: 145671	epsilon: 0.9095694400019632
steps: 246	lr: 0.0001	evaluation reward: 1.61	
episode: 788	score: 1.0	memory length: 145842	epsilon: 0.9092308600019705
steps: 171	lr: 0.0001	evaluation reward: 1.52	
episode: 789	score: 0.0	memory length: 145965	epsilon: 0.9089873200019758
steps: 123	lr: 0.0001	evaluation reward: 1.5	
episode: 790	score: 0.0	memory length: 146088	epsilon: 0.9087437800019811
steps: 123	lr: 0.0001	evaluation reward: 1.48	
episode: 791	score: 1.0	memory length: 146256	epsilon: 0.9084111400019883
steps: 168	lr: 0.0001	evaluation reward: 1.48	
episode: 792	score: 2.0	memory length: 146456	epsilon: 0.9080151400019969
steps: 200	lr: 0.0001	evaluation reward: 1.48	
episode: 793	score: 3.0	memory length: 146721	epsilon: 0.9074904400020083
steps: 265	lr: 0.0001	evaluation reward: 1.51	
episode: 794	score: 3.0	memory length: 146968	epsilon: 0.9070013800020189
steps: 247	lr: 0.0001	evaluation reward: 1.54	
episode: 795	score: 0.0	memory length: 147090	epsilon: 0.9067598200020242
steps: 122	lr: 0.0001	evaluation reward: 1.52	

episode: 796	score: 1.0	memory length: 147240	epsilon: 0.9064628200020306
steps: 150	lr: 0.0001	evaluation reward: 1.51	
episode: 797	score: 1.0	memory length: 147408	epsilon: 0.9061301800020378
steps: 168	lr: 0.0001	evaluation reward: 1.49	
episode: 798	score: 2.0	memory length: 147607	epsilon: 0.9057361600020464
steps: 199	lr: 0.0001	evaluation reward: 1.51	
episode: 799	score: 1.0	memory length: 147775	epsilon: 0.9054035200020536
steps: 168	lr: 0.0001	evaluation reward: 1.52	
episode: 800	score: 1.0	memory length: 147946	epsilon: 0.905064940002061
steps: 171	lr: 0.0001	evaluation reward: 1.53	
episode: 801	score: 3.0	memory length: 148192	epsilon: 0.9045778600020715
steps: 246	lr: 0.0001	evaluation reward: 1.53	
episode: 802	score: 0.0	memory length: 148315	epsilon: 0.9043343200020768
steps: 123	lr: 0.0001	evaluation reward: 1.52	
episode: 803	score: 1.0	memory length: 148484	epsilon: 0.9039997000020841
steps: 169	lr: 0.0001	evaluation reward: 1.52	
episode: 804	score: 2.0	memory length: 148668	epsilon: 0.903635380002092
steps: 184	lr: 0.0001	evaluation reward: 1.52	
episode: 805	score: 1.0	memory length: 148819	epsilon: 0.9033364000020985
steps: 151	lr: 0.0001	evaluation reward: 1.48	
episode: 806	score: 4.0	memory length: 149105	epsilon: 0.9027701200021108
steps: 286	lr: 0.0001	evaluation reward: 1.51	
episode: 807	score: 2.0	memory length: 149323	epsilon: 0.9023384800021201
steps: 218	lr: 0.0001	evaluation reward: 1.53	
episode: 808	score: 2.0	memory length: 149521	epsilon: 0.9019464400021286
steps: 198	lr: 0.0001	evaluation reward: 1.54	
episode: 809	score: 0.0	memory length: 149643	epsilon: 0.9017048800021339
steps: 122	lr: 0.0001	evaluation reward: 1.52	
episode: 810	score: 4.0	memory length: 149922	epsilon: 0.9011524600021459
steps: 279	lr: 0.0001	evaluation reward: 1.52	
episode: 811	score: 0.0	memory length: 150044	epsilon: 0.9009109000021511
steps: 122	lr: 0.0001	evaluation reward: 1.52	
episode: 812	score: 1.0	memory length: 150195	epsilon: 0.9006119200021576
steps: 151	lr: 0.0001	evaluation reward: 1.51	
episode: 813	score: 0.0	memory length: 150318	epsilon: 0.9003683800021629
steps: 123	lr: 0.0001	evaluation reward: 1.5	
episode: 814	score: 4.0	memory length: 150592	epsilon: 0.8998258600021747
steps: 274	lr: 0.0001	evaluation reward: 1.49	
episode: 815	score: 4.0	memory length: 150851	epsilon: 0.8993130400021858
steps: 259	lr: 0.0001	evaluation reward: 1.53	
episode: 816	score: 1.0	memory length: 151002	epsilon: 0.8990140600021923
steps: 151	lr: 0.0001	evaluation reward: 1.52	
episode: 817	score: 1.0	memory length: 151173	epsilon: 0.8986754800021997
steps: 171	lr: 0.0001	evaluation reward: 1.53	
episode: 818	score: 1.0	memory length: 151341	epsilon: 0.8983428400022069
steps: 168	lr: 0.0001	evaluation reward: 1.51	
episode: 819	score: 3.0	memory length: 151588	epsilon: 0.8978537800022175
steps: 247	lr: 0.0001	evaluation reward: 1.51	

episode: 820	score: 3.0	memory length: 151833	epsilon: 0.897368680002228
steps: 245	lr: 0.0001	evaluation reward: 1.51	
episode: 821	score: 0.0	memory length: 151956	epsilon: 0.8971251400022333
steps: 123	lr: 0.0001	evaluation reward: 1.5	
episode: 822	score: 1.0	memory length: 152107	epsilon: 0.8968261600022398
steps: 151	lr: 0.0001	evaluation reward: 1.5	
episode: 823	score: 0.0	memory length: 152230	epsilon: 0.8965826200022451
steps: 123	lr: 0.0001	evaluation reward: 1.5	
episode: 824	score: 2.0	memory length: 152449	epsilon: 0.8961490000022545
steps: 219	lr: 0.0001	evaluation reward: 1.5	
episode: 825	score: 2.0	memory length: 152648	epsilon: 0.8957549800022631
steps: 199	lr: 0.0001	evaluation reward: 1.5	
episode: 826	score: 1.0	memory length: 152817	epsilon: 0.8954203600022703
steps: 169	lr: 0.0001	evaluation reward: 1.5	
episode: 827	score: 1.0	memory length: 152987	epsilon: 0.8950837600022776
steps: 170	lr: 0.0001	evaluation reward: 1.49	
episode: 828	score: 0.0	memory length: 153110	epsilon: 0.8948402200022829
steps: 123	lr: 0.0001	evaluation reward: 1.49	
episode: 829	score: 3.0	memory length: 153354	epsilon: 0.8943571000022934
steps: 244	lr: 0.0001	evaluation reward: 1.51	
episode: 830	score: 4.0	memory length: 153629	epsilon: 0.8938126000023052
steps: 275	lr: 0.0001	evaluation reward: 1.54	
episode: 831	score: 3.0	memory length: 153873	epsilon: 0.8933294800023157
steps: 244	lr: 0.0001	evaluation reward: 1.55	
episode: 832	score: 2.0	memory length: 154092	epsilon: 0.8928958600023251
steps: 219	lr: 0.0001	evaluation reward: 1.55	
episode: 833	score: 1.0	memory length: 154261	epsilon: 0.8925612400023324
steps: 169	lr: 0.0001	evaluation reward: 1.53	
episode: 834	score: 3.0	memory length: 154509	epsilon: 0.892070200002343
steps: 248	lr: 0.0001	evaluation reward: 1.55	
episode: 835	score: 1.0	memory length: 154678	epsilon: 0.8917355800023503
steps: 169	lr: 0.0001	evaluation reward: 1.53	
episode: 836	score: 5.0	memory length: 154982	epsilon: 0.8911336600023634
steps: 304	lr: 0.0001	evaluation reward: 1.56	
episode: 837	score: 1.0	memory length: 155151	epsilon: 0.8907990400023706
steps: 169	lr: 0.0001	evaluation reward: 1.55	
episode: 838	score: 1.0	memory length: 155302	epsilon: 0.8905000600023771
steps: 151	lr: 0.0001	evaluation reward: 1.52	
episode: 839	score: 0.0	memory length: 155424	epsilon: 0.8902585000023824
steps: 122	lr: 0.0001	evaluation reward: 1.52	
episode: 840	score: 3.0	memory length: 155650	epsilon: 0.8898110200023921
steps: 226	lr: 0.0001	evaluation reward: 1.54	
episode: 841	score: 1.0	memory length: 155820	epsilon: 0.8894744200023994
steps: 170	lr: 0.0001	evaluation reward: 1.53	
episode: 842	score: 1.0	memory length: 155971	epsilon: 0.8891754400024059
steps: 151	lr: 0.0001	evaluation reward: 1.52	
episode: 843	score: 1.0	memory length: 156140	epsilon: 0.8888408200024132
steps: 169	lr: 0.0001	evaluation reward: 1.53	

episode: 844	score: 2.0	memory length: 156357	epsilon: 0.8884111600024225
steps: 217	lr: 0.0001	evaluation reward: 1.54	
episode: 845	score: 1.0	memory length: 156526	epsilon: 0.8880765400024297
steps: 169	lr: 0.0001	evaluation reward: 1.54	
episode: 846	score: 0.0	memory length: 156649	epsilon: 0.887833000002435
steps: 123	lr: 0.0001	evaluation reward: 1.54	
episode: 847	score: 4.0	memory length: 156929	epsilon: 0.8872786000024471
steps: 280	lr: 0.0001	evaluation reward: 1.56	
episode: 848	score: 2.0	memory length: 157148	epsilon: 0.8868449800024565
steps: 219	lr: 0.0001	evaluation reward: 1.57	
episode: 849	score: 2.0	memory length: 157345	epsilon: 0.886454920002465
steps: 197	lr: 0.0001	evaluation reward: 1.58	
episode: 850	score: 2.0	memory length: 157543	epsilon: 0.8860628800024735
steps: 198	lr: 0.0001	evaluation reward: 1.6	
episode: 851	score: 1.0	memory length: 157714	epsilon: 0.8857243000024808
steps: 171	lr: 0.0001	evaluation reward: 1.61	
episode: 852	score: 1.0	memory length: 157882	epsilon: 0.885391660002488
steps: 168	lr: 0.0001	evaluation reward: 1.62	
episode: 853	score: 4.0	memory length: 158157	epsilon: 0.8848471600024999
steps: 275	lr: 0.0001	evaluation reward: 1.65	
episode: 854	score: 4.0	memory length: 158452	epsilon: 0.8842630600025125
steps: 295	lr: 0.0001	evaluation reward: 1.68	
episode: 855	score: 0.0	memory length: 158574	epsilon: 0.8840215000025178
steps: 122	lr: 0.0001	evaluation reward: 1.66	
episode: 856	score: 0.0	memory length: 158697	epsilon: 0.8837779600025231
steps: 123	lr: 0.0001	evaluation reward: 1.64	
episode: 857	score: 0.0	memory length: 158819	epsilon: 0.8835364000025283
steps: 122	lr: 0.0001	evaluation reward: 1.64	
episode: 858	score: 1.0	memory length: 158970	epsilon: 0.8832374200025348
steps: 151	lr: 0.0001	evaluation reward: 1.65	
episode: 859	score: 0.0	memory length: 159092	epsilon: 0.88299586000254
steps: 122	lr: 0.0001	evaluation reward: 1.65	
episode: 860	score: 0.0	memory length: 159214	epsilon: 0.8827543000025453
steps: 122	lr: 0.0001	evaluation reward: 1.64	
episode: 861	score: 1.0	memory length: 159384	epsilon: 0.8824177000025526
steps: 170	lr: 0.0001	evaluation reward: 1.65	
episode: 862	score: 2.0	memory length: 159602	epsilon: 0.881986060002562
steps: 218	lr: 0.0001	evaluation reward: 1.66	
episode: 863	score: 2.0	memory length: 159800	epsilon: 0.8815940200025705
steps: 198	lr: 0.0001	evaluation reward: 1.68	
episode: 864	score: 1.0	memory length: 159968	epsilon: 0.8812613800025777
steps: 168	lr: 0.0001	evaluation reward: 1.69	
episode: 865	score: 2.0	memory length: 160168	epsilon: 0.8808653800025863
steps: 200	lr: 0.0001	evaluation reward: 1.69	
episode: 866	score: 6.0	memory length: 160422	epsilon: 0.8803624600025972
steps: 254	lr: 0.0001	evaluation reward: 1.73	
episode: 867	score: 1.0	memory length: 160591	epsilon: 0.8800278400026045
steps: 169	lr: 0.0001	evaluation reward: 1.72	

episode: 868	score: 2.0	memory length: 160806	epsilon: 0.8796021400026137
steps: 215	lr: 0.0001	evaluation reward: 1.74	
episode: 869	score: 3.0	memory length: 161052	epsilon: 0.8791150600026243
steps: 246	lr: 0.0001	evaluation reward: 1.75	
episode: 870	score: 3.0	memory length: 161299	epsilon: 0.8786260000026349
steps: 247	lr: 0.0001	evaluation reward: 1.73	
episode: 871	score: 0.0	memory length: 161422	epsilon: 0.8783824600026402
steps: 123	lr: 0.0001	evaluation reward: 1.73	
episode: 872	score: 0.0	memory length: 161544	epsilon: 0.8781409000026454
steps: 122	lr: 0.0001	evaluation reward: 1.71	
episode: 873	score: 1.0	memory length: 161713	epsilon: 0.8778062800026527
steps: 169	lr: 0.0001	evaluation reward: 1.71	
episode: 874	score: 0.0	memory length: 161836	epsilon: 0.877562740002658
steps: 123	lr: 0.0001	evaluation reward: 1.71	
episode: 875	score: 3.0	memory length: 162082	epsilon: 0.8770756600026686
steps: 246	lr: 0.0001	evaluation reward: 1.71	
episode: 876	score: 0.0	memory length: 162205	epsilon: 0.8768321200026739
steps: 123	lr: 0.0001	evaluation reward: 1.7	
episode: 877	score: 7.0	memory length: 162506	epsilon: 0.8762361400026868
steps: 301	lr: 0.0001	evaluation reward: 1.73	
episode: 878	score: 2.0	memory length: 162705	epsilon: 0.8758421200026953
steps: 199	lr: 0.0001	evaluation reward: 1.74	
episode: 879	score: 3.0	memory length: 162953	epsilon: 0.875351080002706
steps: 248	lr: 0.0001	evaluation reward: 1.75	
episode: 880	score: 1.0	memory length: 163104	epsilon: 0.8750521000027125
steps: 151	lr: 0.0001	evaluation reward: 1.74	
episode: 881	score: 7.0	memory length: 163423	epsilon: 0.8744204800027262
steps: 319	lr: 0.0001	evaluation reward: 1.8	
episode: 882	score: 2.0	memory length: 163621	epsilon: 0.8740284400027347
steps: 198	lr: 0.0001	evaluation reward: 1.73	
episode: 883	score: 4.0	memory length: 163899	epsilon: 0.8734780000027467
steps: 278	lr: 0.0001	evaluation reward: 1.77	
episode: 884	score: 2.0	memory length: 164096	epsilon: 0.8730879400027551
steps: 197	lr: 0.0001	evaluation reward: 1.78	
episode: 885	score: 0.0	memory length: 164219	epsilon: 0.8728444000027604
steps: 123	lr: 0.0001	evaluation reward: 1.74	
episode: 886	score: 0.0	memory length: 164342	epsilon: 0.8726008600027657
steps: 123	lr: 0.0001	evaluation reward: 1.72	
episode: 887	score: 0.0	memory length: 164465	epsilon: 0.872357320002771
steps: 123	lr: 0.0001	evaluation reward: 1.69	
episode: 888	score: 1.0	memory length: 164616	epsilon: 0.8720583400027775
steps: 151	lr: 0.0001	evaluation reward: 1.69	
episode: 889	score: 4.0	memory length: 164894	epsilon: 0.8715079000027894
steps: 278	lr: 0.0001	evaluation reward: 1.73	
episode: 890	score: 1.0	memory length: 165045	epsilon: 0.8712089200027959
steps: 151	lr: 0.0001	evaluation reward: 1.74	
episode: 891	score: 1.0	memory length: 165214	epsilon: 0.8708743000028032
steps: 169	lr: 0.0001	evaluation reward: 1.74	

episode: 892	score: 1.0	memory length: 165383	epsilon: 0.8705396800028105
steps: 169	lr: 0.0001	evaluation reward: 1.73	
episode: 893	score: 2.0	memory length: 165580	epsilon: 0.8701496200028189
steps: 197	lr: 0.0001	evaluation reward: 1.72	
episode: 894	score: 3.0	memory length: 165806	epsilon: 0.8697021400028286
steps: 226	lr: 0.0001	evaluation reward: 1.72	
episode: 895	score: 1.0	memory length: 165958	epsilon: 0.8694011800028352
steps: 152	lr: 0.0001	evaluation reward: 1.73	
episode: 896	score: 2.0	memory length: 166175	epsilon: 0.8689715200028445
steps: 217	lr: 0.0001	evaluation reward: 1.74	
episode: 897	score: 1.0	memory length: 166344	epsilon: 0.8686369000028518
steps: 169	lr: 0.0001	evaluation reward: 1.74	
episode: 898	score: 0.0	memory length: 166466	epsilon: 0.868395340002857
steps: 122	lr: 0.0001	evaluation reward: 1.72	
episode: 899	score: 1.0	memory length: 166617	epsilon: 0.8680963600028635
steps: 151	lr: 0.0001	evaluation reward: 1.72	
episode: 900	score: 2.0	memory length: 166814	epsilon: 0.867706300002872
steps: 197	lr: 0.0001	evaluation reward: 1.73	
episode: 901	score: 4.0	memory length: 167090	epsilon: 0.8671598200028838
steps: 276	lr: 0.0001	evaluation reward: 1.74	
episode: 902	score: 1.0	memory length: 167241	epsilon: 0.8668608400028903
steps: 151	lr: 0.0001	evaluation reward: 1.75	
episode: 903	score: 1.0	memory length: 167391	epsilon: 0.8665638400028968
steps: 150	lr: 0.0001	evaluation reward: 1.75	
episode: 904	score: 0.0	memory length: 167514	epsilon: 0.866320300002902
steps: 123	lr: 0.0001	evaluation reward: 1.73	
episode: 905	score: 0.0	memory length: 167637	epsilon: 0.8660767600029073
steps: 123	lr: 0.0001	evaluation reward: 1.72	
episode: 906	score: 0.0	memory length: 167760	epsilon: 0.8658332200029126
steps: 123	lr: 0.0001	evaluation reward: 1.68	
episode: 907	score: 2.0	memory length: 167958	epsilon: 0.8654411800029211
steps: 198	lr: 0.0001	evaluation reward: 1.68	
episode: 908	score: 0.0	memory length: 168081	epsilon: 0.8651976400029264
steps: 123	lr: 0.0001	evaluation reward: 1.66	
episode: 909	score: 2.0	memory length: 168260	epsilon: 0.8648432200029341
steps: 179	lr: 0.0001	evaluation reward: 1.68	
episode: 910	score: 1.0	memory length: 168411	epsilon: 0.8645442400029406
steps: 151	lr: 0.0001	evaluation reward: 1.65	
episode: 911	score: 2.0	memory length: 168608	epsilon: 0.8641541800029491
steps: 197	lr: 0.0001	evaluation reward: 1.67	
episode: 912	score: 2.0	memory length: 168806	epsilon: 0.8637621400029576
steps: 198	lr: 0.0001	evaluation reward: 1.68	
episode: 913	score: 3.0	memory length: 169014	epsilon: 0.8633503000029665
steps: 208	lr: 0.0001	evaluation reward: 1.71	
episode: 914	score: 1.0	memory length: 169165	epsilon: 0.863051320002973
steps: 151	lr: 0.0001	evaluation reward: 1.68	
episode: 915	score: 2.0	memory length: 169383	epsilon: 0.8626196800029824
steps: 218	lr: 0.0001	evaluation reward: 1.66	

episode: 916	score: 3.0	memory length: 169631	epsilon: 0.862128640002993
steps: 248	lr: 0.0001	evaluation reward: 1.68	
episode: 917	score: 2.0	memory length: 169828	epsilon: 0.8617385800030015
steps: 197	lr: 0.0001	evaluation reward: 1.69	
episode: 918	score: 2.0	memory length: 170013	epsilon: 0.8613722800030095
steps: 185	lr: 0.0001	evaluation reward: 1.7	
episode: 919	score: 0.0	memory length: 170136	epsilon: 0.8611287400030148
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 920	score: 7.0	memory length: 170532	epsilon: 0.8603446600030318
steps: 396	lr: 0.0001	evaluation reward: 1.71	
episode: 921	score: 3.0	memory length: 170781	epsilon: 0.8598516400030425
steps: 249	lr: 0.0001	evaluation reward: 1.74	
episode: 922	score: 2.0	memory length: 170979	epsilon: 0.859459600003051
steps: 198	lr: 0.0001	evaluation reward: 1.75	
episode: 923	score: 2.0	memory length: 171196	epsilon: 0.8590299400030603
steps: 217	lr: 0.0001	evaluation reward: 1.77	
episode: 924	score: 3.0	memory length: 171441	epsilon: 0.8585448400030709
steps: 245	lr: 0.0001	evaluation reward: 1.78	
episode: 925	score: 0.0	memory length: 171564	epsilon: 0.8583013000030761
steps: 123	lr: 0.0001	evaluation reward: 1.76	
episode: 926	score: 2.0	memory length: 171782	epsilon: 0.8578696600030855
steps: 218	lr: 0.0001	evaluation reward: 1.77	
episode: 927	score: 0.0	memory length: 171904	epsilon: 0.8576281000030908
steps: 122	lr: 0.0001	evaluation reward: 1.76	
episode: 928	score: 2.0	memory length: 172120	epsilon: 0.8572004200031
steps: 216	lr: 0.0001	evaluation reward: 1.78	
episode: 929	score: 2.0	memory length: 172318	epsilon: 0.8568083800031085
steps: 198	lr: 0.0001	evaluation reward: 1.77	
episode: 930	score: 0.0	memory length: 172441	epsilon: 0.8565648400031138
steps: 123	lr: 0.0001	evaluation reward: 1.73	
episode: 931	score: 2.0	memory length: 172639	epsilon: 0.8561728000031223
steps: 198	lr: 0.0001	evaluation reward: 1.72	
episode: 932	score: 2.0	memory length: 172857	epsilon: 0.8557411600031317
steps: 218	lr: 0.0001	evaluation reward: 1.72	
episode: 933	score: 2.0	memory length: 173054	epsilon: 0.8553511000031402
steps: 197	lr: 0.0001	evaluation reward: 1.73	
episode: 934	score: 2.0	memory length: 173251	epsilon: 0.8549610400031487
steps: 197	lr: 0.0001	evaluation reward: 1.72	
episode: 935	score: 0.0	memory length: 173374	epsilon: 0.8547175000031539
steps: 123	lr: 0.0001	evaluation reward: 1.71	
episode: 936	score: 1.0	memory length: 173543	epsilon: 0.8543828800031612
steps: 169	lr: 0.0001	evaluation reward: 1.67	
episode: 937	score: 1.0	memory length: 173693	epsilon: 0.8540858800031677
steps: 150	lr: 0.0001	evaluation reward: 1.67	
episode: 938	score: 2.0	memory length: 173913	epsilon: 0.8536502800031771
steps: 220	lr: 0.0001	evaluation reward: 1.68	
episode: 939	score: 2.0	memory length: 174111	epsilon: 0.8532582400031856
steps: 198	lr: 0.0001	evaluation reward: 1.7	

episode: 940	score: 3.0	memory length: 174357	epsilon: 0.8527711600031962
steps: 246	lr: 0.0001	evaluation reward: 1.7	
episode: 941	score: 1.0	memory length: 174525	epsilon: 0.8524385200032034
steps: 168	lr: 0.0001	evaluation reward: 1.7	
episode: 942	score: 8.0	memory length: 175009	epsilon: 0.8514802000032242
steps: 484	lr: 0.0001	evaluation reward: 1.77	
episode: 943	score: 0.0	memory length: 175132	epsilon: 0.8512366600032295
steps: 123	lr: 0.0001	evaluation reward: 1.76	
episode: 944	score: 2.0	memory length: 175347	epsilon: 0.8508109600032387
steps: 215	lr: 0.0001	evaluation reward: 1.76	
episode: 945	score: 3.0	memory length: 175594	epsilon: 0.8503219000032494
steps: 247	lr: 0.0001	evaluation reward: 1.78	
episode: 946	score: 2.0	memory length: 175811	epsilon: 0.8498922400032587
steps: 217	lr: 0.0001	evaluation reward: 1.8	
episode: 947	score: 2.0	memory length: 176008	epsilon: 0.8495021800032672
steps: 197	lr: 0.0001	evaluation reward: 1.78	
episode: 948	score: 2.0	memory length: 176223	epsilon: 0.8490764800032764
steps: 215	lr: 0.0001	evaluation reward: 1.78	
episode: 949	score: 2.0	memory length: 176441	epsilon: 0.8486448400032858
steps: 218	lr: 0.0001	evaluation reward: 1.78	
episode: 950	score: 2.0	memory length: 176658	epsilon: 0.8482151800032951
steps: 217	lr: 0.0001	evaluation reward: 1.78	
episode: 951	score: 1.0	memory length: 176827	epsilon: 0.8478805600033024
steps: 169	lr: 0.0001	evaluation reward: 1.78	
episode: 952	score: 1.0	memory length: 176978	epsilon: 0.8475815800033089
steps: 151	lr: 0.0001	evaluation reward: 1.78	
episode: 953	score: 2.0	memory length: 177199	epsilon: 0.8471440000033184
steps: 221	lr: 0.0001	evaluation reward: 1.76	
episode: 954	score: 4.0	memory length: 177494	epsilon: 0.846559900003331
steps: 295	lr: 0.0001	evaluation reward: 1.76	
episode: 955	score: 2.0	memory length: 177692	epsilon: 0.8461678600033395
steps: 198	lr: 0.0001	evaluation reward: 1.78	
episode: 956	score: 1.0	memory length: 177843	epsilon: 0.845868880003346
steps: 151	lr: 0.0001	evaluation reward: 1.79	
episode: 957	score: 0.0	memory length: 177966	epsilon: 0.8456253400033513
steps: 123	lr: 0.0001	evaluation reward: 1.79	
episode: 958	score: 3.0	memory length: 178233	epsilon: 0.8450966800033628
steps: 267	lr: 0.0001	evaluation reward: 1.81	
episode: 959	score: 6.0	memory length: 178595	epsilon: 0.8443799200033784
steps: 362	lr: 0.0001	evaluation reward: 1.87	
episode: 960	score: 3.0	memory length: 178840	epsilon: 0.8438948200033889
steps: 245	lr: 0.0001	evaluation reward: 1.9	
episode: 961	score: 2.0	memory length: 179037	epsilon: 0.8435047600033974
steps: 197	lr: 0.0001	evaluation reward: 1.91	
episode: 962	score: 4.0	memory length: 179335	epsilon: 0.8429147200034102
steps: 298	lr: 0.0001	evaluation reward: 1.93	
episode: 963	score: 4.0	memory length: 179648	epsilon: 0.8422949800034236
steps: 313	lr: 0.0001	evaluation reward: 1.95	

episode: 964	score: 0.0	memory length: 179771	epsilon: 0.8420514400034289
steps: 123	lr: 0.0001	evaluation reward: 1.94	
episode: 965	score: 0.0	memory length: 179893	epsilon: 0.8418098800034342
steps: 122	lr: 0.0001	evaluation reward: 1.92	
episode: 966	score: 0.0	memory length: 180016	epsilon: 0.8415663400034394
steps: 123	lr: 0.0001	evaluation reward: 1.86	
episode: 967	score: 3.0	memory length: 180264	epsilon: 0.8410753000034501
steps: 248	lr: 0.0001	evaluation reward: 1.88	
episode: 968	score: 2.0	memory length: 180482	epsilon: 0.8406436600034595
steps: 218	lr: 0.0001	evaluation reward: 1.88	
episode: 969	score: 1.0	memory length: 180633	epsilon: 0.840344680003466
steps: 151	lr: 0.0001	evaluation reward: 1.86	
episode: 970	score: 0.0	memory length: 180756	epsilon: 0.8401011400034712
steps: 123	lr: 0.0001	evaluation reward: 1.83	
episode: 971	score: 1.0	memory length: 180925	epsilon: 0.8397665200034785
steps: 169	lr: 0.0001	evaluation reward: 1.84	
episode: 972	score: 3.0	memory length: 181171	epsilon: 0.8392794400034891
steps: 246	lr: 0.0001	evaluation reward: 1.87	
episode: 973	score: 2.0	memory length: 181389	epsilon: 0.8388478000034985
steps: 218	lr: 0.0001	evaluation reward: 1.88	
episode: 974	score: 2.0	memory length: 181587	epsilon: 0.838455760003507
steps: 198	lr: 0.0001	evaluation reward: 1.9	
episode: 975	score: 2.0	memory length: 181805	epsilon: 0.8380241200035163
steps: 218	lr: 0.0001	evaluation reward: 1.89	
episode: 976	score: 3.0	memory length: 182052	epsilon: 0.837535060003527
steps: 247	lr: 0.0001	evaluation reward: 1.92	
episode: 977	score: 1.0	memory length: 182221	epsilon: 0.8372004400035342
steps: 169	lr: 0.0001	evaluation reward: 1.86	
episode: 978	score: 2.0	memory length: 182419	epsilon: 0.8368084000035427
steps: 198	lr: 0.0001	evaluation reward: 1.86	
episode: 979	score: 3.0	memory length: 182649	epsilon: 0.8363530000035526
steps: 230	lr: 0.0001	evaluation reward: 1.86	
episode: 980	score: 4.0	memory length: 182887	epsilon: 0.8358817600035628
steps: 238	lr: 0.0001	evaluation reward: 1.89	
episode: 981	score: 1.0	memory length: 183057	epsilon: 0.8355451600035702
steps: 170	lr: 0.0001	evaluation reward: 1.83	
episode: 982	score: 3.0	memory length: 183303	epsilon: 0.8350580800035807
steps: 246	lr: 0.0001	evaluation reward: 1.84	
episode: 983	score: 2.0	memory length: 183501	epsilon: 0.8346660400035892
steps: 198	lr: 0.0001	evaluation reward: 1.82	
episode: 984	score: 4.0	memory length: 183814	epsilon: 0.8340463000036027
steps: 313	lr: 0.0001	evaluation reward: 1.84	
episode: 985	score: 3.0	memory length: 184041	epsilon: 0.8335968400036124
steps: 227	lr: 0.0001	evaluation reward: 1.87	
episode: 986	score: 5.0	memory length: 184384	epsilon: 0.8329177000036272
steps: 343	lr: 0.0001	evaluation reward: 1.92	
episode: 987	score: 2.0	memory length: 184603	epsilon: 0.8324840800036366
steps: 219	lr: 0.0001	evaluation reward: 1.94	

episode: 988	score: 0.0	memory length: 184726	epsilon: 0.8322405400036419
steps: 123	lr: 0.0001	evaluation reward: 1.93	
episode: 989	score: 3.0	memory length: 184952	epsilon: 0.8317930600036516
steps: 226	lr: 0.0001	evaluation reward: 1.92	
episode: 990	score: 4.0	memory length: 185231	epsilon: 0.8312406400036636
steps: 279	lr: 0.0001	evaluation reward: 1.95	
episode: 991	score: 3.0	memory length: 185477	epsilon: 0.8307535600036742
steps: 246	lr: 0.0001	evaluation reward: 1.97	
episode: 992	score: 2.0	memory length: 185675	epsilon: 0.8303615200036827
steps: 198	lr: 0.0001	evaluation reward: 1.98	
episode: 993	score: 2.0	memory length: 185893	epsilon: 0.829929880003692
steps: 218	lr: 0.0001	evaluation reward: 1.98	
episode: 994	score: 1.0	memory length: 186063	epsilon: 0.8295932800036994
steps: 170	lr: 0.0001	evaluation reward: 1.96	
episode: 995	score: 1.0	memory length: 186214	epsilon: 0.8292943000037059
steps: 151	lr: 0.0001	evaluation reward: 1.96	
episode: 996	score: 5.0	memory length: 186539	epsilon: 0.8286508000037198
steps: 325	lr: 0.0001	evaluation reward: 1.99	
episode: 997	score: 3.0	memory length: 186771	epsilon: 0.8281914400037298
steps: 232	lr: 0.0001	evaluation reward: 2.01	
episode: 998	score: 4.0	memory length: 187067	epsilon: 0.8276053600037425
steps: 296	lr: 0.0001	evaluation reward: 2.05	
episode: 999	score: 0.0	memory length: 187190	epsilon: 0.8273618200037478
steps: 123	lr: 0.0001	evaluation reward: 2.04	
episode: 1000	score: 0.0	memory length: 187313	epsilon: 0.8271182800037531
steps: 123	lr: 0.0001	evaluation reward: 2.02	
episode: 1001	score: 0.0	memory length: 187435	epsilon: 0.8268767200037583
steps: 122	lr: 0.0001	evaluation reward: 1.98	
episode: 1002	score: 0.0	memory length: 187557	epsilon: 0.8266351600037636
steps: 122	lr: 0.0001	evaluation reward: 1.97	
episode: 1003	score: 2.0	memory length: 187755	epsilon: 0.8262431200037721
steps: 198	lr: 0.0001	evaluation reward: 1.98	
episode: 1004	score: 0.0	memory length: 187878	epsilon: 0.8259995800037774
steps: 123	lr: 0.0001	evaluation reward: 1.98	
episode: 1005	score: 1.0	memory length: 188029	epsilon: 0.8257006000037839
steps: 151	lr: 0.0001	evaluation reward: 1.99	
episode: 1006	score: 1.0	memory length: 188200	epsilon: 0.8253620200037912
steps: 171	lr: 0.0001	evaluation reward: 2.0	
episode: 1007	score: 3.0	memory length: 188448	epsilon: 0.8248709800038019
steps: 248	lr: 0.0001	evaluation reward: 2.01	
episode: 1008	score: 2.0	memory length: 188645	epsilon: 0.8244809200038103
steps: 197	lr: 0.0001	evaluation reward: 2.03	
episode: 1009	score: 1.0	memory length: 188813	epsilon: 0.8241482800038176
steps: 168	lr: 0.0001	evaluation reward: 2.02	
episode: 1010	score: 2.0	memory length: 189011	epsilon: 0.8237562400038261
steps: 198	lr: 0.0001	evaluation reward: 2.03	
episode: 1011	score: 0.0	memory length: 189134	epsilon: 0.8235127000038314
steps: 123	lr: 0.0001	evaluation reward: 2.01	

episode: 1012	score: 1.0	memory length: 189303	epsilon: 0.8231780800038386
steps: 169	lr: 0.0001	evaluation reward: 2.0	
episode: 1013	score: 3.0	memory length: 189570	epsilon: 0.8226494200038501
steps: 267	lr: 0.0001	evaluation reward: 2.0	
episode: 1014	score: 2.0	memory length: 189791	epsilon: 0.8222118400038596
steps: 221	lr: 0.0001	evaluation reward: 2.01	
episode: 1015	score: 2.0	memory length: 190009	epsilon: 0.821780200003869
steps: 218	lr: 0.0001	evaluation reward: 2.01	
episode: 1016	score: 2.0	memory length: 190207	epsilon: 0.8213881600038775
steps: 198	lr: 0.0001	evaluation reward: 2.0	
episode: 1017	score: 0.0	memory length: 190330	epsilon: 0.8211446200038828
steps: 123	lr: 0.0001	evaluation reward: 1.98	
episode: 1018	score: 2.0	memory length: 190548	epsilon: 0.8207129800038921
steps: 218	lr: 0.0001	evaluation reward: 1.98	
episode: 1019	score: 0.0	memory length: 190671	epsilon: 0.8204694400038974
steps: 123	lr: 0.0001	evaluation reward: 1.98	
episode: 1020	score: 3.0	memory length: 190900	epsilon: 0.8200160200039073
steps: 229	lr: 0.0001	evaluation reward: 1.94	
episode: 1021	score: 6.0	memory length: 191228	epsilon: 0.8193665800039214
steps: 328	lr: 0.0001	evaluation reward: 1.97	
episode: 1022	score: 2.0	memory length: 191425	epsilon: 0.8189765200039298
steps: 197	lr: 0.0001	evaluation reward: 1.97	
episode: 1023	score: 0.0	memory length: 191548	epsilon: 0.8187329800039351
steps: 123	lr: 0.0001	evaluation reward: 1.95	
episode: 1024	score: 2.0	memory length: 191746	epsilon: 0.8183409400039436
steps: 198	lr: 0.0001	evaluation reward: 1.94	
episode: 1025	score: 1.0	memory length: 191897	epsilon: 0.8180419600039501
steps: 151	lr: 0.0001	evaluation reward: 1.95	
episode: 1026	score: 2.0	memory length: 192115	epsilon: 0.8176103200039595
steps: 218	lr: 0.0001	evaluation reward: 1.95	
episode: 1027	score: 3.0	memory length: 192381	epsilon: 0.8170836400039709
steps: 266	lr: 0.0001	evaluation reward: 1.98	
episode: 1028	score: 2.0	memory length: 192598	epsilon: 0.8166539800039803
steps: 217	lr: 0.0001	evaluation reward: 1.98	
episode: 1029	score: 2.0	memory length: 192795	epsilon: 0.8162639200039887
steps: 197	lr: 0.0001	evaluation reward: 1.98	
episode: 1030	score: 2.0	memory length: 193012	epsilon: 0.8158342600039981
steps: 217	lr: 0.0001	evaluation reward: 2.0	
episode: 1031	score: 2.0	memory length: 193209	epsilon: 0.8154442000040065
steps: 197	lr: 0.0001	evaluation reward: 2.0	
episode: 1032	score: 1.0	memory length: 193360	epsilon: 0.815145220004013
steps: 151	lr: 0.0001	evaluation reward: 1.99	
episode: 1033	score: 2.0	memory length: 193558	epsilon: 0.8147531800040215
steps: 198	lr: 0.0001	evaluation reward: 1.99	
episode: 1034	score: 3.0	memory length: 193783	epsilon: 0.8143076800040312
steps: 225	lr: 0.0001	evaluation reward: 2.0	
episode: 1035	score: 1.0	memory length: 193933	epsilon: 0.8140106800040376
steps: 150	lr: 0.0001	evaluation reward: 2.01	

episode: 1036	score: 2.0	memory length: 194130	epsilon: 0.8136206200040461
steps: 197	lr: 0.0001	evaluation reward: 2.02	
episode: 1037	score: 3.0	memory length: 194376	epsilon: 0.8131335400040567
steps: 246	lr: 0.0001	evaluation reward: 2.04	
episode: 1038	score: 2.0	memory length: 194573	epsilon: 0.8127434800040652
steps: 197	lr: 0.0001	evaluation reward: 2.04	
episode: 1039	score: 2.0	memory length: 194790	epsilon: 0.8123138200040745
steps: 217	lr: 0.0001	evaluation reward: 2.04	
episode: 1040	score: 2.0	memory length: 195008	epsilon: 0.8118821800040839
steps: 218	lr: 0.0001	evaluation reward: 2.03	
episode: 1041	score: 2.0	memory length: 195206	epsilon: 0.8114901400040924
steps: 198	lr: 0.0001	evaluation reward: 2.04	
episode: 1042	score: 2.0	memory length: 195425	epsilon: 0.8110565200041018
steps: 219	lr: 0.0001	evaluation reward: 1.98	
episode: 1043	score: 2.0	memory length: 195642	epsilon: 0.8106268600041111
steps: 217	lr: 0.0001	evaluation reward: 2.0	
episode: 1044	score: 2.0	memory length: 195860	epsilon: 0.8101952200041205
steps: 218	lr: 0.0001	evaluation reward: 2.0	
episode: 1045	score: 1.0	memory length: 196011	epsilon: 0.809896240004127
steps: 151	lr: 0.0001	evaluation reward: 1.98	
episode: 1046	score: 2.0	memory length: 196208	epsilon: 0.8095061800041354
steps: 197	lr: 0.0001	evaluation reward: 1.98	
episode: 1047	score: 0.0	memory length: 196331	epsilon: 0.8092626400041407
steps: 123	lr: 0.0001	evaluation reward: 1.96	
episode: 1048	score: 2.0	memory length: 196528	epsilon: 0.8088725800041492
steps: 197	lr: 0.0001	evaluation reward: 1.96	
episode: 1049	score: 0.0	memory length: 196651	epsilon: 0.8086290400041545
steps: 123	lr: 0.0001	evaluation reward: 1.94	
episode: 1050	score: 4.0	memory length: 196927	epsilon: 0.8080825600041663
steps: 276	lr: 0.0001	evaluation reward: 1.96	
episode: 1051	score: 2.0	memory length: 197124	epsilon: 0.8076925000041748
steps: 197	lr: 0.0001	evaluation reward: 1.97	
episode: 1052	score: 2.0	memory length: 197322	epsilon: 0.8073004600041833
steps: 198	lr: 0.0001	evaluation reward: 1.98	
episode: 1053	score: 1.0	memory length: 197491	epsilon: 0.8069658400041906
steps: 169	lr: 0.0001	evaluation reward: 1.97	
episode: 1054	score: 0.0	memory length: 197614	epsilon: 0.8067223000041959
steps: 123	lr: 0.0001	evaluation reward: 1.93	
episode: 1055	score: 4.0	memory length: 197910	epsilon: 0.8061362200042086
steps: 296	lr: 0.0001	evaluation reward: 1.95	
episode: 1056	score: 3.0	memory length: 198135	epsilon: 0.8056907200042183
steps: 225	lr: 0.0001	evaluation reward: 1.97	
episode: 1057	score: 0.0	memory length: 198257	epsilon: 0.8054491600042235
steps: 122	lr: 0.0001	evaluation reward: 1.97	
episode: 1058	score: 0.0	memory length: 198380	epsilon: 0.8052056200042288
steps: 123	lr: 0.0001	evaluation reward: 1.94	
episode: 1059	score: 3.0	memory length: 198646	epsilon: 0.8046789400042402
steps: 266	lr: 0.0001	evaluation reward: 1.91	

episode: 1060	score: 2.0	memory length: 198844	epsilon: 0.8042869000042487
steps: 198	lr: 0.0001	evaluation reward: 1.9	
episode: 1061	score: 4.0	memory length: 199130	epsilon: 0.803720620004261
steps: 286	lr: 0.0001	evaluation reward: 1.92	
episode: 1062	score: 4.0	memory length: 199445	epsilon: 0.8030969200042746
steps: 315	lr: 0.0001	evaluation reward: 1.92	
episode: 1063	score: 2.0	memory length: 199663	epsilon: 0.8026652800042839
steps: 218	lr: 0.0001	evaluation reward: 1.9	
episode: 1064	score: 2.0	memory length: 199861	epsilon: 0.8022732400042925
steps: 198	lr: 0.0001	evaluation reward: 1.92	
episode: 1065	score: 2.0	memory length: 200061	epsilon: 0.801877240004301
steps: 200	lr: 0.0001	evaluation reward: 1.94	
episode: 1066	score: 0.0	memory length: 200184	epsilon: 0.8016337000043063
steps: 123	lr: 0.0001	evaluation reward: 1.94	
episode: 1067	score: 0.0	memory length: 200307	epsilon: 0.8013901600043116
steps: 123	lr: 0.0001	evaluation reward: 1.91	
episode: 1068	score: 2.0	memory length: 200527	epsilon: 0.8009545600043211
steps: 220	lr: 0.0001	evaluation reward: 1.91	
episode: 1069	score: 2.0	memory length: 200747	epsilon: 0.8005189600043305
steps: 220	lr: 0.0001	evaluation reward: 1.92	
episode: 1070	score: 2.0	memory length: 200948	epsilon: 0.8001209800043392
steps: 201	lr: 0.0001	evaluation reward: 1.94	
episode: 1071	score: 0.0	memory length: 201070	epsilon: 0.7998794200043444
steps: 122	lr: 0.0001	evaluation reward: 1.93	
episode: 1072	score: 2.0	memory length: 201267	epsilon: 0.7994893600043529
steps: 197	lr: 0.0001	evaluation reward: 1.92	
episode: 1073	score: 3.0	memory length: 201513	epsilon: 0.7990022800043635
steps: 246	lr: 0.0001	evaluation reward: 1.93	
episode: 1074	score: 0.0	memory length: 201635	epsilon: 0.7987607200043687
steps: 122	lr: 0.0001	evaluation reward: 1.91	
episode: 1075	score: 2.0	memory length: 201833	epsilon: 0.7983686800043772
steps: 198	lr: 0.0001	evaluation reward: 1.91	
episode: 1076	score: 4.0	memory length: 202112	epsilon: 0.7978162600043892
steps: 279	lr: 0.0001	evaluation reward: 1.92	
episode: 1077	score: 0.0	memory length: 202235	epsilon: 0.7975727200043945
steps: 123	lr: 0.0001	evaluation reward: 1.91	
episode: 1078	score: 0.0	memory length: 202358	epsilon: 0.7973291800043998
steps: 123	lr: 0.0001	evaluation reward: 1.89	
episode: 1079	score: 3.0	memory length: 202628	epsilon: 0.7967945800044114
steps: 270	lr: 0.0001	evaluation reward: 1.89	
episode: 1080	score: 0.0	memory length: 202750	epsilon: 0.7965530200044166
steps: 122	lr: 0.0001	evaluation reward: 1.85	
episode: 1081	score: 2.0	memory length: 202948	epsilon: 0.7961609800044251
steps: 198	lr: 0.0001	evaluation reward: 1.86	
episode: 1082	score: 3.0	memory length: 203215	epsilon: 0.7956323200044366
steps: 267	lr: 0.0001	evaluation reward: 1.86	
episode: 1083	score: 2.0	memory length: 203413	epsilon: 0.7952402800044451
steps: 198	lr: 0.0001	evaluation reward: 1.86	

episode: 1084	score: 2.0	memory length: 203611	epsilon: 0.7948482400044536
steps: 198	lr: 0.0001	evaluation reward: 1.84	
episode: 1085	score: 0.0	memory length: 203734	epsilon: 0.7946047000044589
steps: 123	lr: 0.0001	evaluation reward: 1.81	
episode: 1086	score: 2.0	memory length: 203931	epsilon: 0.7942146400044674
steps: 197	lr: 0.0001	evaluation reward: 1.78	
episode: 1087	score: 2.0	memory length: 204129	epsilon: 0.7938226000044759
steps: 198	lr: 0.0001	evaluation reward: 1.78	
episode: 1088	score: 2.0	memory length: 204327	epsilon: 0.7934305600044844
steps: 198	lr: 0.0001	evaluation reward: 1.8	
episode: 1089	score: 1.0	memory length: 204499	epsilon: 0.7930900000044918
steps: 172	lr: 0.0001	evaluation reward: 1.78	
episode: 1090	score: 0.0	memory length: 204621	epsilon: 0.7928484400044971
steps: 122	lr: 0.0001	evaluation reward: 1.74	
episode: 1091	score: 3.0	memory length: 204865	epsilon: 0.7923653200045075
steps: 244	lr: 0.0001	evaluation reward: 1.74	
episode: 1092	score: 2.0	memory length: 205063	epsilon: 0.7919732800045161
steps: 198	lr: 0.0001	evaluation reward: 1.74	
episode: 1093	score: 2.0	memory length: 205280	epsilon: 0.7915436200045254
steps: 217	lr: 0.0001	evaluation reward: 1.74	
episode: 1094	score: 0.0	memory length: 205403	epsilon: 0.7913000800045307
steps: 123	lr: 0.0001	evaluation reward: 1.73	
episode: 1095	score: 2.0	memory length: 205601	epsilon: 0.7909080400045392
steps: 198	lr: 0.0001	evaluation reward: 1.74	
episode: 1096	score: 3.0	memory length: 205829	epsilon: 0.790456600004549
steps: 228	lr: 0.0001	evaluation reward: 1.72	
episode: 1097	score: 2.0	memory length: 206050	epsilon: 0.7900190200045585
steps: 221	lr: 0.0001	evaluation reward: 1.71	
episode: 1098	score: 4.0	memory length: 206343	epsilon: 0.7894388800045711
steps: 293	lr: 0.0001	evaluation reward: 1.71	
episode: 1099	score: 3.0	memory length: 206590	epsilon: 0.7889498200045817
steps: 247	lr: 0.0001	evaluation reward: 1.74	
episode: 1100	score: 1.0	memory length: 206759	epsilon: 0.788615200004589
steps: 169	lr: 0.0001	evaluation reward: 1.75	
episode: 1101	score: 1.0	memory length: 206927	epsilon: 0.7882825600045962
steps: 168	lr: 0.0001	evaluation reward: 1.76	
episode: 1102	score: 3.0	memory length: 207153	epsilon: 0.7878350800046059
steps: 226	lr: 0.0001	evaluation reward: 1.79	
episode: 1103	score: 2.0	memory length: 207351	epsilon: 0.7874430400046144
steps: 198	lr: 0.0001	evaluation reward: 1.79	
episode: 1104	score: 3.0	memory length: 207577	epsilon: 0.7869955600046241
steps: 226	lr: 0.0001	evaluation reward: 1.82	
episode: 1105	score: 3.0	memory length: 207824	epsilon: 0.7865065000046347
steps: 247	lr: 0.0001	evaluation reward: 1.84	
episode: 1106	score: 3.0	memory length: 208054	epsilon: 0.7860511000046446
steps: 230	lr: 0.0001	evaluation reward: 1.86	
episode: 1107	score: 1.0	memory length: 208224	epsilon: 0.7857145000046519
steps: 170	lr: 0.0001	evaluation reward: 1.84	

episode: 1108	score: 1.0	memory length: 208375	epsilon: 0.7854155200046584
steps: 151	lr: 0.0001	evaluation reward: 1.83	
episode: 1109	score: 2.0	memory length: 208573	epsilon: 0.7850234800046669
steps: 198	lr: 0.0001	evaluation reward: 1.84	
episode: 1110	score: 5.0	memory length: 208921	epsilon: 0.7843344400046819
steps: 348	lr: 0.0001	evaluation reward: 1.87	
episode: 1111	score: 3.0	memory length: 209190	epsilon: 0.7838018200046935
steps: 269	lr: 0.0001	evaluation reward: 1.9	
episode: 1112	score: 4.0	memory length: 209488	epsilon: 0.7832117800047063
steps: 298	lr: 0.0001	evaluation reward: 1.93	
episode: 1113	score: 1.0	memory length: 209657	epsilon: 0.7828771600047135
steps: 169	lr: 0.0001	evaluation reward: 1.91	
episode: 1114	score: 2.0	memory length: 209872	epsilon: 0.7824514600047228
steps: 215	lr: 0.0001	evaluation reward: 1.91	
episode: 1115	score: 3.0	memory length: 210119	epsilon: 0.7819624000047334
steps: 247	lr: 0.0001	evaluation reward: 1.92	
episode: 1116	score: 4.0	memory length: 210415	epsilon: 0.7813763200047461
steps: 296	lr: 0.0001	evaluation reward: 1.94	
episode: 1117	score: 4.0	memory length: 210730	epsilon: 0.7807526200047596
steps: 315	lr: 0.0001	evaluation reward: 1.98	
episode: 1118	score: 2.0	memory length: 210928	epsilon: 0.7803605800047682
steps: 198	lr: 0.0001	evaluation reward: 1.98	
episode: 1119	score: 3.0	memory length: 211156	epsilon: 0.779909140004778
steps: 228	lr: 0.0001	evaluation reward: 2.01	
episode: 1120	score: 2.0	memory length: 211353	epsilon: 0.7795190800047864
steps: 197	lr: 0.0001	evaluation reward: 2.0	
episode: 1121	score: 5.0	memory length: 211683	epsilon: 0.7788656800048006
steps: 330	lr: 0.0001	evaluation reward: 1.99	
episode: 1122	score: 3.0	memory length: 211931	epsilon: 0.7783746400048113
steps: 248	lr: 0.0001	evaluation reward: 2.0	
episode: 1123	score: 2.0	memory length: 212129	epsilon: 0.7779826000048198
steps: 198	lr: 0.0001	evaluation reward: 2.02	
episode: 1124	score: 2.0	memory length: 212348	epsilon: 0.7775489800048292
steps: 219	lr: 0.0001	evaluation reward: 2.02	
episode: 1125	score: 1.0	memory length: 212517	epsilon: 0.7772143600048365
steps: 169	lr: 0.0001	evaluation reward: 2.02	
episode: 1126	score: 0.0	memory length: 212640	epsilon: 0.7769708200048417
steps: 123	lr: 0.0001	evaluation reward: 2.0	
episode: 1127	score: 1.0	memory length: 212791	epsilon: 0.7766718400048482
steps: 151	lr: 0.0001	evaluation reward: 1.98	
episode: 1128	score: 3.0	memory length: 213037	epsilon: 0.7761847600048588
steps: 246	lr: 0.0001	evaluation reward: 1.99	
episode: 1129	score: 4.0	memory length: 213317	epsilon: 0.7756303600048708
steps: 280	lr: 0.0001	evaluation reward: 2.01	
episode: 1130	score: 0.0	memory length: 213440	epsilon: 0.7753868200048761
steps: 123	lr: 0.0001	evaluation reward: 1.99	
episode: 1131	score: 2.0	memory length: 213637	epsilon: 0.7749967600048846
steps: 197	lr: 0.0001	evaluation reward: 1.99	

episode: 1132	score: 2.0	memory length: 213838	epsilon: 0.7745987800048932
steps: 201	lr: 0.0001	evaluation reward: 2.0	
episode: 1133	score: 3.0	memory length: 214067	epsilon: 0.7741453600049031
steps: 229	lr: 0.0001	evaluation reward: 2.01	
episode: 1134	score: 2.0	memory length: 214282	epsilon: 0.7737196600049123
steps: 215	lr: 0.0001	evaluation reward: 2.0	
episode: 1135	score: 0.0	memory length: 214405	epsilon: 0.7734761200049176
steps: 123	lr: 0.0001	evaluation reward: 1.99	
episode: 1136	score: 5.0	memory length: 214731	epsilon: 0.7728306400049316
steps: 326	lr: 0.0001	evaluation reward: 2.02	
episode: 1137	score: 0.0	memory length: 214853	epsilon: 0.7725890800049369
steps: 122	lr: 0.0001	evaluation reward: 1.99	
episode: 1138	score: 1.0	memory length: 215022	epsilon: 0.7722544600049441
steps: 169	lr: 0.0001	evaluation reward: 1.98	
episode: 1139	score: 2.0	memory length: 215220	epsilon: 0.7718624200049526
steps: 198	lr: 0.0001	evaluation reward: 1.98	
episode: 1140	score: 5.0	memory length: 215548	epsilon: 0.7712129800049667
steps: 328	lr: 0.0001	evaluation reward: 2.01	
episode: 1141	score: 1.0	memory length: 215717	epsilon: 0.770878360004974
steps: 169	lr: 0.0001	evaluation reward: 2.0	
episode: 1142	score: 4.0	memory length: 215989	epsilon: 0.7703398000049857
steps: 272	lr: 0.0001	evaluation reward: 2.02	
episode: 1143	score: 0.0	memory length: 216112	epsilon: 0.770096260004991
steps: 123	lr: 0.0001	evaluation reward: 2.0	
episode: 1144	score: 2.0	memory length: 216310	epsilon: 0.7697042200049995
steps: 198	lr: 0.0001	evaluation reward: 2.0	
episode: 1145	score: 10.0	memory length: 216748	epsilon:
0.7688369800050183	steps: 438	lr: 0.0001	evaluation reward: 2.09
episode: 1146	score: 2.0	memory length: 216946	epsilon: 0.7684449400050268
steps: 198	lr: 0.0001	evaluation reward: 2.09	
episode: 1147	score: 9.0	memory length: 217344	epsilon: 0.7676569000050439
steps: 398	lr: 0.0001	evaluation reward: 2.18	
episode: 1148	score: 2.0	memory length: 217542	epsilon: 0.7672648600050525
steps: 198	lr: 0.0001	evaluation reward: 2.18	
episode: 1149	score: 8.0	memory length: 218048	epsilon: 0.7662629800050742
steps: 506	lr: 0.0001	evaluation reward: 2.26	
episode: 1150	score: 0.0	memory length: 218171	epsilon: 0.7660194400050795
steps: 123	lr: 0.0001	evaluation reward: 2.22	
episode: 1151	score: 2.0	memory length: 218369	epsilon: 0.765627400005088
steps: 198	lr: 0.0001	evaluation reward: 2.22	
episode: 1152	score: 2.0	memory length: 218567	epsilon: 0.7652353600050965
steps: 198	lr: 0.0001	evaluation reward: 2.22	
episode: 1153	score: 4.0	memory length: 218841	epsilon: 0.7646928400051083
steps: 274	lr: 0.0001	evaluation reward: 2.25	
episode: 1154	score: 2.0	memory length: 219038	epsilon: 0.7643027800051168
steps: 197	lr: 0.0001	evaluation reward: 2.27	
episode: 1155	score: 4.0	memory length: 219336	epsilon: 0.7637127400051296
steps: 298	lr: 0.0001	evaluation reward: 2.27	

episode: 1156	score: 3.0	memory length: 219549	epsilon: 0.7632910000051387
steps: 213	lr: 0.0001	evaluation reward: 2.27	
episode: 1157	score: 4.0	memory length: 219827	epsilon: 0.7627405600051507
steps: 278	lr: 0.0001	evaluation reward: 2.31	
episode: 1158	score: 2.0	memory length: 220025	epsilon: 0.7623485200051592
steps: 198	lr: 0.0001	evaluation reward: 2.33	
episode: 1159	score: 3.0	memory length: 220252	epsilon: 0.7618990600051689
steps: 227	lr: 0.0001	evaluation reward: 2.33	
episode: 1160	score: 0.0	memory length: 220375	epsilon: 0.7616555200051742
steps: 123	lr: 0.0001	evaluation reward: 2.31	
episode: 1161	score: 5.0	memory length: 220685	epsilon: 0.7610417200051875
steps: 310	lr: 0.0001	evaluation reward: 2.32	
episode: 1162	score: 3.0	memory length: 220911	epsilon: 0.7605942400051973
steps: 226	lr: 0.0001	evaluation reward: 2.31	
episode: 1163	score: 1.0	memory length: 221082	epsilon: 0.7602556600052046
steps: 171	lr: 0.0001	evaluation reward: 2.3	
episode: 1164	score: 2.0	memory length: 221303	epsilon: 0.7598180800052141
steps: 221	lr: 0.0001	evaluation reward: 2.3	
episode: 1165	score: 4.0	memory length: 221596	epsilon: 0.7592379400052267
steps: 293	lr: 0.0001	evaluation reward: 2.32	
episode: 1166	score: 2.0	memory length: 221794	epsilon: 0.7588459000052352
steps: 198	lr: 0.0001	evaluation reward: 2.34	
episode: 1167	score: 4.0	memory length: 222072	epsilon: 0.7582954600052472
steps: 278	lr: 0.0001	evaluation reward: 2.38	
episode: 1168	score: 1.0	memory length: 222241	epsilon: 0.7579608400052544
steps: 169	lr: 0.0001	evaluation reward: 2.37	
episode: 1169	score: 2.0	memory length: 222438	epsilon: 0.7575707800052629
steps: 197	lr: 0.0001	evaluation reward: 2.37	
episode: 1170	score: 2.0	memory length: 222636	epsilon: 0.7571787400052714
steps: 198	lr: 0.0001	evaluation reward: 2.37	
episode: 1171	score: 3.0	memory length: 222864	epsilon: 0.7567273000052812
steps: 228	lr: 0.0001	evaluation reward: 2.4	
episode: 1172	score: 0.0	memory length: 222986	epsilon: 0.7564857400052865
steps: 122	lr: 0.0001	evaluation reward: 2.38	
episode: 1173	score: 2.0	memory length: 223183	epsilon: 0.7560956800052949
steps: 197	lr: 0.0001	evaluation reward: 2.37	
episode: 1174	score: 2.0	memory length: 223402	epsilon: 0.7556620600053043
steps: 219	lr: 0.0001	evaluation reward: 2.39	
episode: 1175	score: 2.0	memory length: 223599	epsilon: 0.7552720000053128
steps: 197	lr: 0.0001	evaluation reward: 2.39	
episode: 1176	score: 1.0	memory length: 223750	epsilon: 0.7549730200053193
steps: 151	lr: 0.0001	evaluation reward: 2.36	
episode: 1177	score: 1.0	memory length: 223900	epsilon: 0.7546760200053257
steps: 150	lr: 0.0001	evaluation reward: 2.37	
episode: 1178	score: 4.0	memory length: 224194	epsilon: 0.7540939000053384
steps: 294	lr: 0.0001	evaluation reward: 2.41	
episode: 1179	score: 1.0	memory length: 224345	epsilon: 0.7537949200053449
steps: 151	lr: 0.0001	evaluation reward: 2.39	

episode: 1180	score: 4.0	memory length: 224631	epsilon: 0.7532286400053572
steps: 286	lr: 0.0001	evaluation reward: 2.43	
episode: 1181	score: 4.0	memory length: 224941	epsilon: 0.7526148400053705
steps: 310	lr: 0.0001	evaluation reward: 2.45	
episode: 1182	score: 2.0	memory length: 225139	epsilon: 0.752222800005379
steps: 198	lr: 0.0001	evaluation reward: 2.44	
episode: 1183	score: 3.0	memory length: 225402	epsilon: 0.7517020600053903
steps: 263	lr: 0.0001	evaluation reward: 2.45	
episode: 1184	score: 5.0	memory length: 225747	epsilon: 0.7510189600054051
steps: 345	lr: 0.0001	evaluation reward: 2.48	
episode: 1185	score: 4.0	memory length: 226063	epsilon: 0.7503932800054187
steps: 316	lr: 0.0001	evaluation reward: 2.52	
episode: 1186	score: 1.0	memory length: 226214	epsilon: 0.7500943000054252
steps: 151	lr: 0.0001	evaluation reward: 2.51	
episode: 1187	score: 3.0	memory length: 226440	epsilon: 0.7496468200054349
steps: 226	lr: 0.0001	evaluation reward: 2.52	
episode: 1188	score: 3.0	memory length: 226687	epsilon: 0.7491577600054455
steps: 247	lr: 0.0001	evaluation reward: 2.53	
episode: 1189	score: 2.0	memory length: 226907	epsilon: 0.748722160005455
steps: 220	lr: 0.0001	evaluation reward: 2.54	
episode: 1190	score: 3.0	memory length: 227153	epsilon: 0.7482350800054656
steps: 246	lr: 0.0001	evaluation reward: 2.57	
episode: 1191	score: 4.0	memory length: 227429	epsilon: 0.7476886000054774
steps: 276	lr: 0.0001	evaluation reward: 2.58	
episode: 1192	score: 1.0	memory length: 227580	epsilon: 0.7473896200054839
steps: 151	lr: 0.0001	evaluation reward: 2.57	
episode: 1193	score: 4.0	memory length: 227876	epsilon: 0.7468035400054966
steps: 296	lr: 0.0001	evaluation reward: 2.59	
episode: 1194	score: 2.0	memory length: 228074	epsilon: 0.7464115000055052
steps: 198	lr: 0.0001	evaluation reward: 2.61	
episode: 1195	score: 2.0	memory length: 228292	epsilon: 0.7459798600055145
steps: 218	lr: 0.0001	evaluation reward: 2.61	
episode: 1196	score: 1.0	memory length: 228462	epsilon: 0.7456432600055218
steps: 170	lr: 0.0001	evaluation reward: 2.59	
episode: 1197	score: 3.0	memory length: 228708	epsilon: 0.7451561800055324
steps: 246	lr: 0.0001	evaluation reward: 2.6	
episode: 1198	score: 3.0	memory length: 228918	epsilon: 0.7447403800055414
steps: 210	lr: 0.0001	evaluation reward: 2.59	
episode: 1199	score: 3.0	memory length: 229163	epsilon: 0.744255280005552
steps: 245	lr: 0.0001	evaluation reward: 2.59	
episode: 1200	score: 2.0	memory length: 229361	epsilon: 0.7438632400055605
steps: 198	lr: 0.0001	evaluation reward: 2.6	
episode: 1201	score: 3.0	memory length: 229586	epsilon: 0.7434177400055701
steps: 225	lr: 0.0001	evaluation reward: 2.62	
episode: 1202	score: 4.0	memory length: 229881	epsilon: 0.7428336400055828
steps: 295	lr: 0.0001	evaluation reward: 2.63	
episode: 1203	score: 3.0	memory length: 230089	epsilon: 0.7424218000055918
steps: 208	lr: 0.0001	evaluation reward: 2.64	

episode: 1204	score: 2.0	memory length: 230307	epsilon: 0.7419901600056011
steps: 218	lr: 0.0001	evaluation reward: 2.63	
episode: 1205	score: 1.0	memory length: 230458	epsilon: 0.7416911800056076
steps: 151	lr: 0.0001	evaluation reward: 2.61	
episode: 1206	score: 4.0	memory length: 230775	epsilon: 0.7410635200056213
steps: 317	lr: 0.0001	evaluation reward: 2.62	
episode: 1207	score: 0.0	memory length: 230897	epsilon: 0.7408219600056265
steps: 122	lr: 0.0001	evaluation reward: 2.61	
episode: 1208	score: 2.0	memory length: 231115	epsilon: 0.7403903200056359
steps: 218	lr: 0.0001	evaluation reward: 2.62	
episode: 1209	score: 5.0	memory length: 231424	epsilon: 0.7397785000056492
steps: 309	lr: 0.0001	evaluation reward: 2.65	
episode: 1210	score: 3.0	memory length: 231655	epsilon: 0.7393211200056591
steps: 231	lr: 0.0001	evaluation reward: 2.63	
episode: 1211	score: 1.0	memory length: 231824	epsilon: 0.7389865000056663
steps: 169	lr: 0.0001	evaluation reward: 2.61	
episode: 1212	score: 3.0	memory length: 232053	epsilon: 0.7385330800056762
steps: 229	lr: 0.0001	evaluation reward: 2.6	
episode: 1213	score: 0.0	memory length: 232176	epsilon: 0.7382895400056815
steps: 123	lr: 0.0001	evaluation reward: 2.59	
episode: 1214	score: 3.0	memory length: 232402	epsilon: 0.7378420600056912
steps: 226	lr: 0.0001	evaluation reward: 2.6	
episode: 1215	score: 0.0	memory length: 232525	epsilon: 0.7375985200056965
steps: 123	lr: 0.0001	evaluation reward: 2.57	
episode: 1216	score: 3.0	memory length: 232772	epsilon: 0.7371094600057071
steps: 247	lr: 0.0001	evaluation reward: 2.56	
episode: 1217	score: 2.0	memory length: 232969	epsilon: 0.7367194000057156
steps: 197	lr: 0.0001	evaluation reward: 2.54	
episode: 1218	score: 1.0	memory length: 233141	epsilon: 0.736378840005723
steps: 172	lr: 0.0001	evaluation reward: 2.53	
episode: 1219	score: 1.0	memory length: 233310	epsilon: 0.7360442200057302
steps: 169	lr: 0.0001	evaluation reward: 2.51	
episode: 1220	score: 3.0	memory length: 233556	epsilon: 0.7355571400057408
steps: 246	lr: 0.0001	evaluation reward: 2.52	
episode: 1221	score: 2.0	memory length: 233755	epsilon: 0.7351631200057493
steps: 199	lr: 0.0001	evaluation reward: 2.49	
episode: 1222	score: 3.0	memory length: 234003	epsilon: 0.73467208000576
steps: 248	lr: 0.0001	evaluation reward: 2.49	
episode: 1223	score: 3.0	memory length: 234229	epsilon: 0.7342246000057697
steps: 226	lr: 0.0001	evaluation reward: 2.5	
episode: 1224	score: 6.0	memory length: 234606	epsilon: 0.7334781400057859
steps: 377	lr: 0.0001	evaluation reward: 2.54	
episode: 1225	score: 2.0	memory length: 234804	epsilon: 0.7330861000057944
steps: 198	lr: 0.0001	evaluation reward: 2.55	
episode: 1226	score: 2.0	memory length: 235001	epsilon: 0.7326960400058029
steps: 197	lr: 0.0001	evaluation reward: 2.57	
episode: 1227	score: 1.0	memory length: 235152	epsilon: 0.7323970600058094
steps: 151	lr: 0.0001	evaluation reward: 2.57	

episode: 1228	score: 1.0	memory length: 235303	epsilon: 0.7320980800058159
steps: 151	lr: 0.0001	evaluation reward: 2.55	
episode: 1229	score: 5.0	memory length: 235610	epsilon: 0.7314902200058291
steps: 307	lr: 0.0001	evaluation reward: 2.56	
episode: 1230	score: 3.0	memory length: 235835	epsilon: 0.7310447200058388
steps: 225	lr: 0.0001	evaluation reward: 2.59	
episode: 1231	score: 4.0	memory length: 236113	epsilon: 0.7304942800058507
steps: 278	lr: 0.0001	evaluation reward: 2.61	
episode: 1232	score: 2.0	memory length: 236311	epsilon: 0.7301022400058592
steps: 198	lr: 0.0001	evaluation reward: 2.61	
episode: 1233	score: 1.0	memory length: 236461	epsilon: 0.7298052400058657
steps: 150	lr: 0.0001	evaluation reward: 2.59	
episode: 1234	score: 1.0	memory length: 236630	epsilon: 0.7294706200058729
steps: 169	lr: 0.0001	evaluation reward: 2.58	
episode: 1235	score: 0.0	memory length: 236753	epsilon: 0.7292270800058782
steps: 123	lr: 0.0001	evaluation reward: 2.58	
episode: 1236	score: 1.0	memory length: 236922	epsilon: 0.7288924600058855
steps: 169	lr: 0.0001	evaluation reward: 2.54	
episode: 1237	score: 3.0	memory length: 237189	epsilon: 0.728363800005897
steps: 267	lr: 0.0001	evaluation reward: 2.57	
episode: 1238	score: 2.0	memory length: 237386	epsilon: 0.7279737400059054
steps: 197	lr: 0.0001	evaluation reward: 2.58	
episode: 1239	score: 3.0	memory length: 237650	epsilon: 0.7274510200059168
steps: 264	lr: 0.0001	evaluation reward: 2.59	
episode: 1240	score: 1.0	memory length: 237820	epsilon: 0.7271144200059241
steps: 170	lr: 0.0001	evaluation reward: 2.55	
episode: 1241	score: 0.0	memory length: 237942	epsilon: 0.7268728600059293
steps: 122	lr: 0.0001	evaluation reward: 2.54	
episode: 1242	score: 1.0	memory length: 238093	epsilon: 0.7265738800059358
steps: 151	lr: 0.0001	evaluation reward: 2.51	
episode: 1243	score: 3.0	memory length: 238339	epsilon: 0.7260868000059464
steps: 246	lr: 0.0001	evaluation reward: 2.54	
episode: 1244	score: 3.0	memory length: 238585	epsilon: 0.725599720005957
steps: 246	lr: 0.0001	evaluation reward: 2.55	
episode: 1245	score: 2.0	memory length: 238803	epsilon: 0.7251680800059663
steps: 218	lr: 0.0001	evaluation reward: 2.47	
episode: 1246	score: 1.0	memory length: 238954	epsilon: 0.7248691000059728
steps: 151	lr: 0.0001	evaluation reward: 2.46	
episode: 1247	score: 4.0	memory length: 239228	epsilon: 0.7243265800059846
steps: 274	lr: 0.0001	evaluation reward: 2.41	
episode: 1248	score: 3.0	memory length: 239476	epsilon: 0.7238355400059953
steps: 248	lr: 0.0001	evaluation reward: 2.42	
episode: 1249	score: 3.0	memory length: 239722	epsilon: 0.7233484600060058
steps: 246	lr: 0.0001	evaluation reward: 2.37	
episode: 1250	score: 0.0	memory length: 239844	epsilon: 0.7231069000060111
steps: 122	lr: 0.0001	evaluation reward: 2.37	
episode: 1251	score: 3.0	memory length: 240094	epsilon: 0.7226119000060218
steps: 250	lr: 0.0001	evaluation reward: 2.38	

episode: 1252	score: 1.0	memory length: 240245	epsilon: 0.7223129200060283
steps: 151	lr: 0.0001	evaluation reward: 2.37	
episode: 1253	score: 3.0	memory length: 240491	epsilon: 0.7218258400060389
steps: 246	lr: 0.0001	evaluation reward: 2.36	
episode: 1254	score: 3.0	memory length: 240719	epsilon: 0.7213744000060487
steps: 228	lr: 0.0001	evaluation reward: 2.37	
episode: 1255	score: 4.0	memory length: 240984	epsilon: 0.7208497000060601
steps: 265	lr: 0.0001	evaluation reward: 2.37	
episode: 1256	score: 4.0	memory length: 241258	epsilon: 0.7203071800060719
steps: 274	lr: 0.0001	evaluation reward: 2.38	
episode: 1257	score: 3.0	memory length: 241487	epsilon: 0.7198537600060817
steps: 229	lr: 0.0001	evaluation reward: 2.37	
episode: 1258	score: 2.0	memory length: 241685	epsilon: 0.7194617200060902
steps: 198	lr: 0.0001	evaluation reward: 2.37	
episode: 1259	score: 4.0	memory length: 241962	epsilon: 0.7189132600061021
steps: 277	lr: 0.0001	evaluation reward: 2.38	
episode: 1260	score: 0.0	memory length: 242085	epsilon: 0.7186697200061074
steps: 123	lr: 0.0001	evaluation reward: 2.38	
episode: 1261	score: 2.0	memory length: 242269	epsilon: 0.7183054000061153
steps: 184	lr: 0.0001	evaluation reward: 2.35	
episode: 1262	score: 3.0	memory length: 242516	epsilon: 0.7178163400061259
steps: 247	lr: 0.0001	evaluation reward: 2.35	
episode: 1263	score: 0.0	memory length: 242639	epsilon: 0.7175728000061312
steps: 123	lr: 0.0001	evaluation reward: 2.34	
episode: 1264	score: 6.0	memory length: 242890	epsilon: 0.717075820006142
steps: 251	lr: 0.0001	evaluation reward: 2.38	
episode: 1265	score: 1.0	memory length: 243040	epsilon: 0.7167788200061485
steps: 150	lr: 0.0001	evaluation reward: 2.35	
episode: 1266	score: 1.0	memory length: 243190	epsilon: 0.7164818200061549
steps: 150	lr: 0.0001	evaluation reward: 2.34	
episode: 1267	score: 1.0	memory length: 243341	epsilon: 0.7161828400061614
steps: 151	lr: 0.0001	evaluation reward: 2.31	
episode: 1268	score: 2.0	memory length: 243539	epsilon: 0.7157908000061699
steps: 198	lr: 0.0001	evaluation reward: 2.32	
episode: 1269	score: 2.0	memory length: 243737	epsilon: 0.7153987600061784
steps: 198	lr: 0.0001	evaluation reward: 2.32	
episode: 1270	score: 2.0	memory length: 243935	epsilon: 0.7150067200061869
steps: 198	lr: 0.0001	evaluation reward: 2.32	
episode: 1271	score: 2.0	memory length: 244152	epsilon: 0.7145770600061963
steps: 217	lr: 0.0001	evaluation reward: 2.31	
episode: 1272	score: 3.0	memory length: 244397	epsilon: 0.7140919600062068
steps: 245	lr: 0.0001	evaluation reward: 2.34	
episode: 1273	score: 1.0	memory length: 244548	epsilon: 0.7137929800062133
steps: 151	lr: 0.0001	evaluation reward: 2.33	
episode: 1274	score: 8.0	memory length: 244917	epsilon: 0.7130623600062291
steps: 369	lr: 0.0001	evaluation reward: 2.39	
episode: 1275	score: 3.0	memory length: 245143	epsilon: 0.7126148800062388
steps: 226	lr: 0.0001	evaluation reward: 2.4	

episode: 1276	score: 2.0	memory length: 245341	epsilon: 0.7122228400062474
steps: 198	lr: 0.0001	evaluation reward: 2.41	
episode: 1277	score: 4.0	memory length: 245623	epsilon: 0.7116644800062595
steps: 282	lr: 0.0001	evaluation reward: 2.44	
episode: 1278	score: 0.0	memory length: 245745	epsilon: 0.7114229200062647
steps: 122	lr: 0.0001	evaluation reward: 2.4	
episode: 1279	score: 3.0	memory length: 246012	epsilon: 0.7108942600062762
steps: 267	lr: 0.0001	evaluation reward: 2.42	
episode: 1280	score: 1.0	memory length: 246183	epsilon: 0.7105556800062836
steps: 171	lr: 0.0001	evaluation reward: 2.39	
episode: 1281	score: 1.0	memory length: 246333	epsilon: 0.71025868000629
steps: 150	lr: 0.0001	evaluation reward: 2.36	
episode: 1282	score: 0.0	memory length: 246456	epsilon: 0.7100151400062953
steps: 123	lr: 0.0001	evaluation reward: 2.34	
episode: 1283	score: 3.0	memory length: 246703	epsilon: 0.7095260800063059
steps: 247	lr: 0.0001	evaluation reward: 2.34	
episode: 1284	score: 1.0	memory length: 246853	epsilon: 0.7092290800063124
steps: 150	lr: 0.0001	evaluation reward: 2.3	
episode: 1285	score: 0.0	memory length: 246975	epsilon: 0.7089875200063176
steps: 122	lr: 0.0001	evaluation reward: 2.26	
episode: 1286	score: 3.0	memory length: 247221	epsilon: 0.7085004400063282
steps: 246	lr: 0.0001	evaluation reward: 2.28	
episode: 1287	score: 2.0	memory length: 247440	epsilon: 0.7080668200063376
steps: 219	lr: 0.0001	evaluation reward: 2.27	
episode: 1288	score: 4.0	memory length: 247698	epsilon: 0.7075559800063487
steps: 258	lr: 0.0001	evaluation reward: 2.28	
episode: 1289	score: 4.0	memory length: 247991	epsilon: 0.7069758400063613
steps: 293	lr: 0.0001	evaluation reward: 2.3	
episode: 1290	score: 1.0	memory length: 248160	epsilon: 0.7066412200063685
steps: 169	lr: 0.0001	evaluation reward: 2.28	
episode: 1291	score: 2.0	memory length: 248378	epsilon: 0.7062095800063779
steps: 218	lr: 0.0001	evaluation reward: 2.26	
episode: 1292	score: 0.0	memory length: 248500	epsilon: 0.7059680200063831
steps: 122	lr: 0.0001	evaluation reward: 2.25	
episode: 1293	score: 0.0	memory length: 248623	epsilon: 0.7057244800063884
steps: 123	lr: 0.0001	evaluation reward: 2.21	
episode: 1294	score: 2.0	memory length: 248823	epsilon: 0.705328480006397
steps: 200	lr: 0.0001	evaluation reward: 2.21	
episode: 1295	score: 0.0	memory length: 248945	epsilon: 0.7050869200064023
steps: 122	lr: 0.0001	evaluation reward: 2.19	
episode: 1296	score: 1.0	memory length: 249096	epsilon: 0.7047879400064088
steps: 151	lr: 0.0001	evaluation reward: 2.19	
episode: 1297	score: 0.0	memory length: 249218	epsilon: 0.704546380006414
steps: 122	lr: 0.0001	evaluation reward: 2.16	
episode: 1298	score: 2.0	memory length: 249437	epsilon: 0.7041127600064234
steps: 219	lr: 0.0001	evaluation reward: 2.15	
episode: 1299	score: 2.0	memory length: 249636	epsilon: 0.703718740006432
steps: 199	lr: 0.0001	evaluation reward: 2.14	

episode: 1300	score: 2.0	memory length: 249833	epsilon: 0.7033286800064404
steps: 197	lr: 0.0001	evaluation reward: 2.14	
episode: 1301	score: 2.0	memory length: 250030	epsilon: 0.7029386200064489
steps: 197	lr: 0.0001	evaluation reward: 2.13	
episode: 1302	score: 2.0	memory length: 250209	epsilon: 0.7025842000064566
steps: 179	lr: 0.0001	evaluation reward: 2.11	
episode: 1303	score: 4.0	memory length: 250484	epsilon: 0.7020397000064684
steps: 275	lr: 0.0001	evaluation reward: 2.12	
episode: 1304	score: 0.0	memory length: 250607	epsilon: 0.7017961600064737
steps: 123	lr: 0.0001	evaluation reward: 2.1	
episode: 1305	score: 3.0	memory length: 250855	epsilon: 0.7013051200064844
steps: 248	lr: 0.0001	evaluation reward: 2.12	
episode: 1306	score: 2.0	memory length: 251052	epsilon: 0.7009150600064928
steps: 197	lr: 0.0001	evaluation reward: 2.1	
episode: 1307	score: 3.0	memory length: 251277	epsilon: 0.7004695600065025
steps: 225	lr: 0.0001	evaluation reward: 2.13	
episode: 1308	score: 2.0	memory length: 251457	epsilon: 0.7001131600065102
steps: 180	lr: 0.0001	evaluation reward: 2.13	
episode: 1309	score: 4.0	memory length: 251753	epsilon: 0.699527080006523
steps: 296	lr: 0.0001	evaluation reward: 2.12	
episode: 1310	score: 3.0	memory length: 251964	epsilon: 0.699109300006532
steps: 211	lr: 0.0001	evaluation reward: 2.12	
episode: 1311	score: 2.0	memory length: 252183	epsilon: 0.6986756800065415
steps: 219	lr: 0.0001	evaluation reward: 2.13	
episode: 1312	score: 1.0	memory length: 252334	epsilon: 0.698376700006548
steps: 151	lr: 0.0001	evaluation reward: 2.11	
episode: 1313	score: 3.0	memory length: 252598	epsilon: 0.6978539800065593
steps: 264	lr: 0.0001	evaluation reward: 2.14	
episode: 1314	score: 1.0	memory length: 252767	epsilon: 0.6975193600065666
steps: 169	lr: 0.0001	evaluation reward: 2.12	
episode: 1315	score: 1.0	memory length: 252935	epsilon: 0.6971867200065738
steps: 168	lr: 0.0001	evaluation reward: 2.13	
episode: 1316	score: 4.0	memory length: 253248	epsilon: 0.6965669800065872
steps: 313	lr: 0.0001	evaluation reward: 2.14	
episode: 1317	score: 2.0	memory length: 253445	epsilon: 0.6961769200065957
steps: 197	lr: 0.0001	evaluation reward: 2.14	
episode: 1318	score: 0.0	memory length: 253567	epsilon: 0.695935360006601
steps: 122	lr: 0.0001	evaluation reward: 2.13	
episode: 1319	score: 0.0	memory length: 253689	epsilon: 0.6956938000066062
steps: 122	lr: 0.0001	evaluation reward: 2.12	
episode: 1320	score: 3.0	memory length: 253915	epsilon: 0.6952463200066159
steps: 226	lr: 0.0001	evaluation reward: 2.12	
episode: 1321	score: 2.0	memory length: 254133	epsilon: 0.6948146800066253
steps: 218	lr: 0.0001	evaluation reward: 2.12	
episode: 1322	score: 3.0	memory length: 254401	epsilon: 0.6942840400066368
steps: 268	lr: 0.0001	evaluation reward: 2.12	
episode: 1323	score: 3.0	memory length: 254627	epsilon: 0.6938365600066465
steps: 226	lr: 0.0001	evaluation reward: 2.12	

episode: 1324	score: 3.0	memory length: 254853	epsilon: 0.6933890800066562
steps: 226	lr: 0.0001	evaluation reward: 2.09	
episode: 1325	score: 2.0	memory length: 255054	epsilon: 0.6929911000066649
steps: 201	lr: 0.0001	evaluation reward: 2.09	
episode: 1326	score: 1.0	memory length: 255222	epsilon: 0.6926584600066721
steps: 168	lr: 0.0001	evaluation reward: 2.08	
episode: 1327	score: 2.0	memory length: 255440	epsilon: 0.6922268200066815
steps: 218	lr: 0.0001	evaluation reward: 2.09	
episode: 1328	score: 2.0	memory length: 255640	epsilon: 0.69183082000669
steps: 200	lr: 0.0001	evaluation reward: 2.1	
episode: 1329	score: 3.0	memory length: 255869	epsilon: 0.6913774000066999
steps: 229	lr: 0.0001	evaluation reward: 2.08	
episode: 1330	score: 3.0	memory length: 256084	epsilon: 0.6909517000067091
steps: 215	lr: 0.0001	evaluation reward: 2.08	
episode: 1331	score: 2.0	memory length: 256281	epsilon: 0.6905616400067176
steps: 197	lr: 0.0001	evaluation reward: 2.06	
episode: 1332	score: 2.0	memory length: 256497	epsilon: 0.6901339600067269
steps: 216	lr: 0.0001	evaluation reward: 2.06	
episode: 1333	score: 2.0	memory length: 256695	epsilon: 0.6897419200067354
steps: 198	lr: 0.0001	evaluation reward: 2.07	
episode: 1334	score: 0.0	memory length: 256817	epsilon: 0.6895003600067406
steps: 122	lr: 0.0001	evaluation reward: 2.06	
episode: 1335	score: 1.0	memory length: 256986	epsilon: 0.6891657400067479
steps: 169	lr: 0.0001	evaluation reward: 2.07	
episode: 1336	score: 0.0	memory length: 257108	epsilon: 0.6889241800067532
steps: 122	lr: 0.0001	evaluation reward: 2.06	
episode: 1337	score: 3.0	memory length: 257318	epsilon: 0.6885083800067622
steps: 210	lr: 0.0001	evaluation reward: 2.06	
episode: 1338	score: 3.0	memory length: 257564	epsilon: 0.6880213000067728
steps: 246	lr: 0.0001	evaluation reward: 2.07	
episode: 1339	score: 2.0	memory length: 257762	epsilon: 0.6876292600067813
steps: 198	lr: 0.0001	evaluation reward: 2.06	
episode: 1340	score: 0.0	memory length: 257885	epsilon: 0.6873857200067865
steps: 123	lr: 0.0001	evaluation reward: 2.05	
episode: 1341	score: 2.0	memory length: 258103	epsilon: 0.6869540800067959
steps: 218	lr: 0.0001	evaluation reward: 2.07	
episode: 1342	score: 2.0	memory length: 258301	epsilon: 0.6865620400068044
steps: 198	lr: 0.0001	evaluation reward: 2.08	
episode: 1343	score: 8.0	memory length: 258725	epsilon: 0.6857225200068227
steps: 424	lr: 0.0001	evaluation reward: 2.13	
episode: 1344	score: 0.0	memory length: 258847	epsilon: 0.6854809600068279
steps: 122	lr: 0.0001	evaluation reward: 2.1	
episode: 1345	score: 1.0	memory length: 259017	epsilon: 0.6851443600068352
steps: 170	lr: 0.0001	evaluation reward: 2.09	
episode: 1346	score: 3.0	memory length: 259243	epsilon: 0.6846968800068449
steps: 226	lr: 0.0001	evaluation reward: 2.11	
episode: 1347	score: 1.0	memory length: 259412	epsilon: 0.6843622600068522
steps: 169	lr: 0.0001	evaluation reward: 2.08	

episode: 1348	score: 3.0	memory length: 259659	epsilon: 0.6838732000068628
steps: 247	lr: 0.0001	evaluation reward: 2.08	
episode: 1349	score: 1.0	memory length: 259809	epsilon: 0.6835762000068693
steps: 150	lr: 0.0001	evaluation reward: 2.06	
episode: 1350	score: 0.0	memory length: 259931	epsilon: 0.6833346400068745
steps: 122	lr: 0.0001	evaluation reward: 2.06	
episode: 1351	score: 2.0	memory length: 260128	epsilon: 0.682944580006883
steps: 197	lr: 0.0001	evaluation reward: 2.05	
episode: 1352	score: 1.0	memory length: 260279	epsilon: 0.6826456000068895
steps: 151	lr: 0.0001	evaluation reward: 2.05	
episode: 1353	score: 4.0	memory length: 260538	epsilon: 0.6821327800069006
steps: 259	lr: 0.0001	evaluation reward: 2.06	
episode: 1354	score: 2.0	memory length: 260738	epsilon: 0.6817367800069092
steps: 200	lr: 0.0001	evaluation reward: 2.05	
episode: 1355	score: 0.0	memory length: 260861	epsilon: 0.6814932400069145
steps: 123	lr: 0.0001	evaluation reward: 2.01	
episode: 1356	score: 4.0	memory length: 261136	epsilon: 0.6809487400069263
steps: 275	lr: 0.0001	evaluation reward: 2.01	
episode: 1357	score: 2.0	memory length: 261334	epsilon: 0.6805567000069348
steps: 198	lr: 0.0001	evaluation reward: 2.0	
episode: 1358	score: 2.0	memory length: 261552	epsilon: 0.6801250600069442
steps: 218	lr: 0.0001	evaluation reward: 2.0	
episode: 1359	score: 4.0	memory length: 261851	epsilon: 0.679533040006957
steps: 299	lr: 0.0001	evaluation reward: 2.0	
episode: 1360	score: 3.0	memory length: 262095	epsilon: 0.6790499200069675
steps: 244	lr: 0.0001	evaluation reward: 2.03	
episode: 1361	score: 1.0	memory length: 262246	epsilon: 0.678750940006974
steps: 151	lr: 0.0001	evaluation reward: 2.02	
episode: 1362	score: 1.0	memory length: 262416	epsilon: 0.6784143400069813
steps: 170	lr: 0.0001	evaluation reward: 2.0	
episode: 1363	score: 4.0	memory length: 262694	epsilon: 0.6778639000069933
steps: 278	lr: 0.0001	evaluation reward: 2.04	
episode: 1364	score: 3.0	memory length: 262920	epsilon: 0.677416420007003
steps: 226	lr: 0.0001	evaluation reward: 2.01	
episode: 1365	score: 2.0	memory length: 263117	epsilon: 0.6770263600070114
steps: 197	lr: 0.0001	evaluation reward: 2.02	
episode: 1366	score: 4.0	memory length: 263414	epsilon: 0.6764383000070242
steps: 297	lr: 0.0001	evaluation reward: 2.05	
episode: 1367	score: 2.0	memory length: 263611	epsilon: 0.6760482400070327
steps: 197	lr: 0.0001	evaluation reward: 2.06	
episode: 1368	score: 0.0	memory length: 263734	epsilon: 0.675804700007038
steps: 123	lr: 0.0001	evaluation reward: 2.04	
episode: 1369	score: 3.0	memory length: 263960	epsilon: 0.6753572200070477
steps: 226	lr: 0.0001	evaluation reward: 2.05	
episode: 1370	score: 4.0	memory length: 264216	epsilon: 0.6748503400070587
steps: 256	lr: 0.0001	evaluation reward: 2.07	
episode: 1371	score: 2.0	memory length: 264434	epsilon: 0.674418700007068
steps: 218	lr: 0.0001	evaluation reward: 2.07	

episode: 1372	score: 3.0	memory length: 264677	epsilon: 0.6739375600070785
steps: 243	lr: 0.0001	evaluation reward: 2.07	
episode: 1373	score: 5.0	memory length: 264991	epsilon: 0.673315840007092
steps: 314	lr: 0.0001	evaluation reward: 2.11	
episode: 1374	score: 2.0	memory length: 265189	epsilon: 0.6729238000071005
steps: 198	lr: 0.0001	evaluation reward: 2.05	
episode: 1375	score: 1.0	memory length: 265340	epsilon: 0.672624820007107
steps: 151	lr: 0.0001	evaluation reward: 2.03	
episode: 1376	score: 3.0	memory length: 265606	epsilon: 0.6720981400071184
steps: 266	lr: 0.0001	evaluation reward: 2.04	
episode: 1377	score: 3.0	memory length: 265832	epsilon: 0.6716506600071281
steps: 226	lr: 0.0001	evaluation reward: 2.03	
episode: 1378	score: 5.0	memory length: 266192	epsilon: 0.6709378600071436
steps: 360	lr: 0.0001	evaluation reward: 2.08	
episode: 1379	score: 2.0	memory length: 266373	epsilon: 0.6705794800071514
steps: 181	lr: 0.0001	evaluation reward: 2.07	
episode: 1380	score: 2.0	memory length: 266555	epsilon: 0.6702191200071592
steps: 182	lr: 0.0001	evaluation reward: 2.08	
episode: 1381	score: 1.0	memory length: 266705	epsilon: 0.6699221200071657
steps: 150	lr: 0.0001	evaluation reward: 2.08	
episode: 1382	score: 0.0	memory length: 266828	epsilon: 0.669678580007171
steps: 123	lr: 0.0001	evaluation reward: 2.08	
episode: 1383	score: 1.0	memory length: 266980	epsilon: 0.6693776200071775
steps: 152	lr: 0.0001	evaluation reward: 2.06	
episode: 1384	score: 3.0	memory length: 267223	epsilon: 0.6688964800071879
steps: 243	lr: 0.0001	evaluation reward: 2.08	
episode: 1385	score: 8.0	memory length: 267647	epsilon: 0.6680569600072062
steps: 424	lr: 0.0001	evaluation reward: 2.16	
episode: 1386	score: 3.0	memory length: 267872	epsilon: 0.6676114600072158
steps: 225	lr: 0.0001	evaluation reward: 2.16	
episode: 1387	score: 3.0	memory length: 268139	epsilon: 0.6670828000072273
steps: 267	lr: 0.0001	evaluation reward: 2.17	
episode: 1388	score: 3.0	memory length: 268408	epsilon: 0.6665501800072389
steps: 269	lr: 0.0001	evaluation reward: 2.16	
episode: 1389	score: 0.0	memory length: 268531	epsilon: 0.6663066400072442
steps: 123	lr: 0.0001	evaluation reward: 2.12	
episode: 1390	score: 0.0	memory length: 268654	epsilon: 0.6660631000072494
steps: 123	lr: 0.0001	evaluation reward: 2.11	
episode: 1391	score: 3.0	memory length: 268902	epsilon: 0.6655720600072601
steps: 248	lr: 0.0001	evaluation reward: 2.12	
episode: 1392	score: 3.0	memory length: 269128	epsilon: 0.6651245800072698
steps: 226	lr: 0.0001	evaluation reward: 2.15	
episode: 1393	score: 6.0	memory length: 269497	epsilon: 0.6643939600072857
steps: 369	lr: 0.0001	evaluation reward: 2.21	
episode: 1394	score: 5.0	memory length: 269816	epsilon: 0.6637623400072994
steps: 319	lr: 0.0001	evaluation reward: 2.24	
episode: 1395	score: 2.0	memory length: 270016	epsilon: 0.663366340007308
steps: 200	lr: 0.0001	evaluation reward: 2.26	

episode: 1396	score: 3.0	memory length: 270260	epsilon: 0.6628832200073185
steps: 244	lr: 0.0001	evaluation reward: 2.28	
episode: 1397	score: 0.0	memory length: 270383	epsilon: 0.6626396800073238
steps: 123	lr: 0.0001	evaluation reward: 2.28	
episode: 1398	score: 2.0	memory length: 270604	epsilon: 0.6622021000073333
steps: 221	lr: 0.0001	evaluation reward: 2.28	
episode: 1399	score: 3.0	memory length: 270849	epsilon: 0.6617170000073438
steps: 245	lr: 0.0001	evaluation reward: 2.29	
episode: 1400	score: 3.0	memory length: 271100	epsilon: 0.6612200200073546
steps: 251	lr: 0.0001	evaluation reward: 2.3	
episode: 1401	score: 0.0	memory length: 271223	epsilon: 0.6609764800073599
steps: 123	lr: 0.0001	evaluation reward: 2.28	
episode: 1402	score: 3.0	memory length: 271466	epsilon: 0.6604953400073703
steps: 243	lr: 0.0001	evaluation reward: 2.29	
episode: 1403	score: 3.0	memory length: 271691	epsilon: 0.66004984000738
steps: 225	lr: 0.0001	evaluation reward: 2.28	
episode: 1404	score: 4.0	memory length: 271977	epsilon: 0.6594835600073923
steps: 286	lr: 0.0001	evaluation reward: 2.32	
episode: 1405	score: 4.0	memory length: 272253	epsilon: 0.6589370800074041
steps: 276	lr: 0.0001	evaluation reward: 2.33	
episode: 1406	score: 3.0	memory length: 272478	epsilon: 0.6584915800074138
steps: 225	lr: 0.0001	evaluation reward: 2.34	
episode: 1407	score: 3.0	memory length: 272724	epsilon: 0.6580045000074244
steps: 246	lr: 0.0001	evaluation reward: 2.34	
episode: 1408	score: 2.0	memory length: 272922	epsilon: 0.6576124600074329
steps: 198	lr: 0.0001	evaluation reward: 2.34	
episode: 1409	score: 3.0	memory length: 273148	epsilon: 0.6571649800074426
steps: 226	lr: 0.0001	evaluation reward: 2.33	
episode: 1410	score: 3.0	memory length: 273397	epsilon: 0.6566719600074533
steps: 249	lr: 0.0001	evaluation reward: 2.33	
episode: 1411	score: 7.0	memory length: 273784	epsilon: 0.65590570000747
steps: 387	lr: 0.0001	evaluation reward: 2.38	
episode: 1412	score: 6.0	memory length: 274034	epsilon: 0.6554107000074807
steps: 250	lr: 0.0001	evaluation reward: 2.43	
episode: 1413	score: 2.0	memory length: 274231	epsilon: 0.6550206400074892
steps: 197	lr: 0.0001	evaluation reward: 2.42	
episode: 1414	score: 6.0	memory length: 274605	epsilon: 0.6542801200075052
steps: 374	lr: 0.0001	evaluation reward: 2.47	
episode: 1415	score: 1.0	memory length: 274773	epsilon: 0.6539474800075125
steps: 168	lr: 0.0001	evaluation reward: 2.47	
episode: 1416	score: 1.0	memory length: 274941	epsilon: 0.6536148400075197
steps: 168	lr: 0.0001	evaluation reward: 2.44	
episode: 1417	score: 3.0	memory length: 275205	epsilon: 0.653092120007531
steps: 264	lr: 0.0001	evaluation reward: 2.45	
episode: 1418	score: 5.0	memory length: 275530	epsilon: 0.652448620007545
steps: 325	lr: 0.0001	evaluation reward: 2.5	
episode: 1419	score: 2.0	memory length: 275747	epsilon: 0.6520189600075543
steps: 217	lr: 0.0001	evaluation reward: 2.52	

episode: 1420	score: 1.0	memory length: 275897	epsilon: 0.6517219600075608
steps: 150	lr: 0.0001	evaluation reward: 2.5	
episode: 1421	score: 2.0	memory length: 276114	epsilon: 0.6512923000075701
steps: 217	lr: 0.0001	evaluation reward: 2.5	
episode: 1422	score: 7.0	memory length: 276501	epsilon: 0.6505260400075867
steps: 387	lr: 0.0001	evaluation reward: 2.54	
episode: 1423	score: 2.0	memory length: 276698	epsilon: 0.6501359800075952
steps: 197	lr: 0.0001	evaluation reward: 2.53	
episode: 1424	score: 4.0	memory length: 276977	epsilon: 0.6495835600076072
steps: 279	lr: 0.0001	evaluation reward: 2.54	
episode: 1425	score: 2.0	memory length: 277195	epsilon: 0.6491519200076166
steps: 218	lr: 0.0001	evaluation reward: 2.54	
episode: 1426	score: 1.0	memory length: 277346	epsilon: 0.6488529400076231
steps: 151	lr: 0.0001	evaluation reward: 2.54	
episode: 1427	score: 3.0	memory length: 277591	epsilon: 0.6483678400076336
steps: 245	lr: 0.0001	evaluation reward: 2.55	
episode: 1428	score: 2.0	memory length: 277789	epsilon: 0.6479758000076421
steps: 198	lr: 0.0001	evaluation reward: 2.55	
episode: 1429	score: 6.0	memory length: 278163	epsilon: 0.6472352800076582
steps: 374	lr: 0.0001	evaluation reward: 2.58	
episode: 1430	score: 3.0	memory length: 278389	epsilon: 0.6467878000076679
steps: 226	lr: 0.0001	evaluation reward: 2.58	
episode: 1431	score: 2.0	memory length: 278587	epsilon: 0.6463957600076764
steps: 198	lr: 0.0001	evaluation reward: 2.58	
episode: 1432	score: 0.0	memory length: 278709	epsilon: 0.6461542000076816
steps: 122	lr: 0.0001	evaluation reward: 2.56	
episode: 1433	score: 3.0	memory length: 278952	epsilon: 0.6456730600076921
steps: 243	lr: 0.0001	evaluation reward: 2.57	
episode: 1434	score: 2.0	memory length: 279150	epsilon: 0.6452810200077006
steps: 198	lr: 0.0001	evaluation reward: 2.59	
episode: 1435	score: 2.0	memory length: 279348	epsilon: 0.6448889800077091
steps: 198	lr: 0.0001	evaluation reward: 2.6	
episode: 1436	score: 2.0	memory length: 279548	epsilon: 0.6444929800077177
steps: 200	lr: 0.0001	evaluation reward: 2.62	
episode: 1437	score: 4.0	memory length: 279824	epsilon: 0.6439465000077296
steps: 276	lr: 0.0001	evaluation reward: 2.63	
episode: 1438	score: 2.0	memory length: 280022	epsilon: 0.6435544600077381
steps: 198	lr: 0.0001	evaluation reward: 2.62	
episode: 1439	score: 4.0	memory length: 280296	epsilon: 0.6430119400077499
steps: 274	lr: 0.0001	evaluation reward: 2.64	
episode: 1440	score: 3.0	memory length: 280522	epsilon: 0.6425644600077596
steps: 226	lr: 0.0001	evaluation reward: 2.67	
episode: 1441	score: 2.0	memory length: 280720	epsilon: 0.6421724200077681
steps: 198	lr: 0.0001	evaluation reward: 2.67	
episode: 1442	score: 3.0	memory length: 280964	epsilon: 0.6416893000077786
steps: 244	lr: 0.0001	evaluation reward: 2.68	
episode: 1443	score: 3.0	memory length: 281190	epsilon: 0.6412418200077883
steps: 226	lr: 0.0001	evaluation reward: 2.63	

episode: 1444	score: 2.0	memory length: 281387	epsilon: 0.6408517600077968
steps: 197	lr: 0.0001	evaluation reward: 2.65	
episode: 1445	score: 1.0	memory length: 281555	epsilon: 0.640519120007804
steps: 168	lr: 0.0001	evaluation reward: 2.65	
episode: 1446	score: 0.0	memory length: 281678	epsilon: 0.6402755800078093
steps: 123	lr: 0.0001	evaluation reward: 2.62	
episode: 1447	score: 2.0	memory length: 281876	epsilon: 0.6398835400078178
steps: 198	lr: 0.0001	evaluation reward: 2.63	
episode: 1448	score: 5.0	memory length: 282215	epsilon: 0.6392123200078323
steps: 339	lr: 0.0001	evaluation reward: 2.65	
episode: 1449	score: 4.0	memory length: 282481	epsilon: 0.6386856400078438
steps: 266	lr: 0.0001	evaluation reward: 2.68	
episode: 1450	score: 3.0	memory length: 282747	epsilon: 0.6381589600078552
steps: 266	lr: 0.0001	evaluation reward: 2.71	
episode: 1451	score: 2.0	memory length: 282928	epsilon: 0.637800580007863
steps: 181	lr: 0.0001	evaluation reward: 2.71	
episode: 1452	score: 3.0	memory length: 283154	epsilon: 0.6373531000078727
steps: 226	lr: 0.0001	evaluation reward: 2.73	
episode: 1453	score: 6.0	memory length: 283538	epsilon: 0.6365927800078892
steps: 384	lr: 0.0001	evaluation reward: 2.75	
episode: 1454	score: 1.0	memory length: 283689	epsilon: 0.6362938000078957
steps: 151	lr: 0.0001	evaluation reward: 2.74	
episode: 1455	score: 4.0	memory length: 283981	epsilon: 0.6357156400079083
steps: 292	lr: 0.0001	evaluation reward: 2.78	
episode: 1456	score: 0.0	memory length: 284104	epsilon: 0.6354721000079135
steps: 123	lr: 0.0001	evaluation reward: 2.74	
episode: 1457	score: 2.0	memory length: 284305	epsilon: 0.6350741200079222
steps: 201	lr: 0.0001	evaluation reward: 2.74	
episode: 1458	score: 0.0	memory length: 284428	epsilon: 0.6348305800079275
steps: 123	lr: 0.0001	evaluation reward: 2.72	
episode: 1459	score: 3.0	memory length: 284671	epsilon: 0.6343494400079379
steps: 243	lr: 0.0001	evaluation reward: 2.71	
episode: 1460	score: 1.0	memory length: 284839	epsilon: 0.6340168000079451
steps: 168	lr: 0.0001	evaluation reward: 2.69	
episode: 1461	score: 4.0	memory length: 285095	epsilon: 0.6335099200079561
steps: 256	lr: 0.0001	evaluation reward: 2.72	
episode: 1462	score: 0.0	memory length: 285218	epsilon: 0.6332663800079614
steps: 123	lr: 0.0001	evaluation reward: 2.71	
episode: 1463	score: 5.0	memory length: 285527	epsilon: 0.6326545600079747
steps: 309	lr: 0.0001	evaluation reward: 2.72	
episode: 1464	score: 4.0	memory length: 285820	epsilon: 0.6320744200079873
steps: 293	lr: 0.0001	evaluation reward: 2.73	
episode: 1465	score: 1.0	memory length: 285988	epsilon: 0.6317417800079945
steps: 168	lr: 0.0001	evaluation reward: 2.72	
episode: 1466	score: 2.0	memory length: 286186	epsilon: 0.631349740008003
steps: 198	lr: 0.0001	evaluation reward: 2.7	
episode: 1467	score: 2.0	memory length: 286368	epsilon: 0.6309893800080109
steps: 182	lr: 0.0001	evaluation reward: 2.7	

episode: 1468	score: 3.0	memory length: 286615	epsilon: 0.6305003200080215
steps: 247	lr: 0.0001	evaluation reward: 2.73	
episode: 1469	score: 2.0	memory length: 286836	epsilon: 0.630062740008031
steps: 221	lr: 0.0001	evaluation reward: 2.72	
episode: 1470	score: 2.0	memory length: 287034	epsilon: 0.6296707000080395
steps: 198	lr: 0.0001	evaluation reward: 2.7	
episode: 1471	score: 3.0	memory length: 287259	epsilon: 0.6292252000080492
steps: 225	lr: 0.0001	evaluation reward: 2.71	
episode: 1472	score: 4.0	memory length: 287576	epsilon: 0.6285975400080628
steps: 317	lr: 0.0001	evaluation reward: 2.72	
episode: 1473	score: 2.0	memory length: 287774	epsilon: 0.6282055000080713
steps: 198	lr: 0.0001	evaluation reward: 2.69	
episode: 1474	score: 2.0	memory length: 287971	epsilon: 0.6278154400080798
steps: 197	lr: 0.0001	evaluation reward: 2.69	
episode: 1475	score: 4.0	memory length: 288241	epsilon: 0.6272808400080914
steps: 270	lr: 0.0001	evaluation reward: 2.72	
episode: 1476	score: 2.0	memory length: 288439	epsilon: 0.6268888000080999
steps: 198	lr: 0.0001	evaluation reward: 2.71	
episode: 1477	score: 2.0	memory length: 288639	epsilon: 0.6264928000081085
steps: 200	lr: 0.0001	evaluation reward: 2.7	
episode: 1478	score: 2.0	memory length: 288836	epsilon: 0.6261027400081169
steps: 197	lr: 0.0001	evaluation reward: 2.67	
episode: 1479	score: 2.0	memory length: 289034	epsilon: 0.6257107000081255
steps: 198	lr: 0.0001	evaluation reward: 2.67	
episode: 1480	score: 4.0	memory length: 289294	epsilon: 0.6251959000081366
steps: 260	lr: 0.0001	evaluation reward: 2.69	
episode: 1481	score: 1.0	memory length: 289444	epsilon: 0.6248989000081431
steps: 150	lr: 0.0001	evaluation reward: 2.69	
episode: 1482	score: 1.0	memory length: 289616	epsilon: 0.6245583400081505
steps: 172	lr: 0.0001	evaluation reward: 2.7	
episode: 1483	score: 4.0	memory length: 289912	epsilon: 0.6239722600081632
steps: 296	lr: 0.0001	evaluation reward: 2.73	
episode: 1484	score: 2.0	memory length: 290130	epsilon: 0.6235406200081726
steps: 218	lr: 0.0001	evaluation reward: 2.72	
episode: 1485	score: 2.0	memory length: 290327	epsilon: 0.623150560008181
steps: 197	lr: 0.0001	evaluation reward: 2.66	
episode: 1486	score: 2.0	memory length: 290546	epsilon: 0.6227169400081904
steps: 219	lr: 0.0001	evaluation reward: 2.65	
episode: 1487	score: 0.0	memory length: 290668	epsilon: 0.6224753800081957
steps: 122	lr: 0.0001	evaluation reward: 2.62	
episode: 1488	score: 4.0	memory length: 290982	epsilon: 0.6218536600082092
steps: 314	lr: 0.0001	evaluation reward: 2.63	
episode: 1489	score: 3.0	memory length: 291208	epsilon: 0.6214061800082189
steps: 226	lr: 0.0001	evaluation reward: 2.66	
episode: 1490	score: 2.0	memory length: 291406	epsilon: 0.6210141400082274
steps: 198	lr: 0.0001	evaluation reward: 2.68	
episode: 1491	score: 2.0	memory length: 291603	epsilon: 0.6206240800082359
steps: 197	lr: 0.0001	evaluation reward: 2.67	

episode: 1492	score: 4.0	memory length: 291877	epsilon: 0.6200815600082477
steps: 274	lr: 0.0001	evaluation reward: 2.68	
episode: 1493	score: 2.0	memory length: 292058	epsilon: 0.6197231800082554
steps: 181	lr: 0.0001	evaluation reward: 2.64	
episode: 1494	score: 2.0	memory length: 292276	epsilon: 0.6192915400082648
steps: 218	lr: 0.0001	evaluation reward: 2.61	
episode: 1495	score: 2.0	memory length: 292475	epsilon: 0.6188975200082734
steps: 199	lr: 0.0001	evaluation reward: 2.61	
episode: 1496	score: 1.0	memory length: 292644	epsilon: 0.6185629000082806
steps: 169	lr: 0.0001	evaluation reward: 2.59	
episode: 1497	score: 2.0	memory length: 292862	epsilon: 0.61813126000829
steps: 218	lr: 0.0001	evaluation reward: 2.61	
episode: 1498	score: 4.0	memory length: 293156	epsilon: 0.6175491400083026
steps: 294	lr: 0.0001	evaluation reward: 2.63	
episode: 1499	score: 3.0	memory length: 293381	epsilon: 0.6171036400083123
steps: 225	lr: 0.0001	evaluation reward: 2.63	
episode: 1500	score: 1.0	memory length: 293532	epsilon: 0.6168046600083188
steps: 151	lr: 0.0001	evaluation reward: 2.61	
episode: 1501	score: 9.0	memory length: 293860	epsilon: 0.6161552200083329
steps: 328	lr: 0.0001	evaluation reward: 2.7	
episode: 1502	score: 4.0	memory length: 294156	epsilon: 0.6155691400083456
steps: 296	lr: 0.0001	evaluation reward: 2.71	
episode: 1503	score: 3.0	memory length: 294382	epsilon: 0.6151216600083553
steps: 226	lr: 0.0001	evaluation reward: 2.71	
episode: 1504	score: 6.0	memory length: 294776	epsilon: 0.6143415400083723
steps: 394	lr: 0.0001	evaluation reward: 2.73	
episode: 1505	score: 3.0	memory length: 295002	epsilon: 0.613894060008382
steps: 226	lr: 0.0001	evaluation reward: 2.72	
episode: 1506	score: 3.0	memory length: 295230	epsilon: 0.6134426200083918
steps: 228	lr: 0.0001	evaluation reward: 2.72	
episode: 1507	score: 0.0	memory length: 295352	epsilon: 0.613201060008397
steps: 122	lr: 0.0001	evaluation reward: 2.69	
episode: 1508	score: 2.0	memory length: 295550	epsilon: 0.6128090200084055
steps: 198	lr: 0.0001	evaluation reward: 2.69	
episode: 1509	score: 2.0	memory length: 295748	epsilon: 0.612416980008414
steps: 198	lr: 0.0001	evaluation reward: 2.68	
episode: 1510	score: 2.0	memory length: 295929	epsilon: 0.6120586000084218
steps: 181	lr: 0.0001	evaluation reward: 2.67	
episode: 1511	score: 3.0	memory length: 296155	epsilon: 0.6116111200084315
steps: 226	lr: 0.0001	evaluation reward: 2.63	
episode: 1512	score: 4.0	memory length: 296451	epsilon: 0.6110250400084443
steps: 296	lr: 0.0001	evaluation reward: 2.61	
episode: 1513	score: 4.0	memory length: 296708	epsilon: 0.6105161800084553
steps: 257	lr: 0.0001	evaluation reward: 2.63	
episode: 1514	score: 2.0	memory length: 296905	epsilon: 0.6101261200084638
steps: 197	lr: 0.0001	evaluation reward: 2.59	
episode: 1515	score: 2.0	memory length: 297102	epsilon: 0.6097360600084722
steps: 197	lr: 0.0001	evaluation reward: 2.6	

episode: 1516	score: 3.0	memory length: 297347	epsilon: 0.6092509600084828
steps: 245	lr: 0.0001	evaluation reward: 2.62	
episode: 1517	score: 1.0	memory length: 297497	epsilon: 0.6089539600084892
steps: 150	lr: 0.0001	evaluation reward: 2.6	
episode: 1518	score: 2.0	memory length: 297695	epsilon: 0.6085619200084977
steps: 198	lr: 0.0001	evaluation reward: 2.57	
episode: 1519	score: 4.0	memory length: 297961	epsilon: 0.6080352400085092
steps: 266	lr: 0.0001	evaluation reward: 2.59	
episode: 1520	score: 5.0	memory length: 298285	epsilon: 0.6073937200085231
steps: 324	lr: 0.0001	evaluation reward: 2.63	
episode: 1521	score: 3.0	memory length: 298531	epsilon: 0.6069066400085337
steps: 246	lr: 0.0001	evaluation reward: 2.64	
episode: 1522	score: 2.0	memory length: 298729	epsilon: 0.6065146000085422
steps: 198	lr: 0.0001	evaluation reward: 2.59	
episode: 1523	score: 4.0	memory length: 299021	epsilon: 0.6059364400085547
steps: 292	lr: 0.0001	evaluation reward: 2.61	
episode: 1524	score: 2.0	memory length: 299220	epsilon: 0.6055424200085633
steps: 199	lr: 0.0001	evaluation reward: 2.59	
episode: 1525	score: 3.0	memory length: 299445	epsilon: 0.605096920008573
steps: 225	lr: 0.0001	evaluation reward: 2.6	
episode: 1526	score: 2.0	memory length: 299644	epsilon: 0.6047029000085815
steps: 199	lr: 0.0001	evaluation reward: 2.61	
episode: 1527	score: 2.0	memory length: 299841	epsilon: 0.60431284000859
steps: 197	lr: 0.0001	evaluation reward: 2.6	
episode: 1528	score: 3.0	memory length: 300087	epsilon: 0.6038257600086006
steps: 246	lr: 0.0001	evaluation reward: 2.61	
episode: 1529	score: 2.0	memory length: 300285	epsilon: 0.6034337200086091
steps: 198	lr: 0.0001	evaluation reward: 2.57	
episode: 1530	score: 1.0	memory length: 300456	epsilon: 0.6030951400086164
steps: 171	lr: 0.0001	evaluation reward: 2.55	
episode: 1531	score: 5.0	memory length: 300777	epsilon: 0.6024595600086302
steps: 321	lr: 0.0001	evaluation reward: 2.58	
episode: 1532	score: 5.0	memory length: 301102	epsilon: 0.6018160600086442
steps: 325	lr: 0.0001	evaluation reward: 2.63	
episode: 1533	score: 5.0	memory length: 301428	epsilon: 0.6011705800086582
steps: 326	lr: 0.0001	evaluation reward: 2.65	
episode: 1534	score: 4.0	memory length: 301721	epsilon: 0.6005904400086708
steps: 293	lr: 0.0001	evaluation reward: 2.67	
episode: 1535	score: 3.0	memory length: 301946	epsilon: 0.6001449400086805
steps: 225	lr: 0.0001	evaluation reward: 2.68	
episode: 1536	score: 3.0	memory length: 302190	epsilon: 0.599661820008691
steps: 244	lr: 0.0001	evaluation reward: 2.69	
episode: 1537	score: 2.0	memory length: 302390	epsilon: 0.5992658200086995
steps: 200	lr: 0.0001	evaluation reward: 2.67	
episode: 1538	score: 2.0	memory length: 302588	epsilon: 0.5988737800087081
steps: 198	lr: 0.0001	evaluation reward: 2.67	
episode: 1539	score: 6.0	memory length: 302983	epsilon: 0.598091680008725
steps: 395	lr: 0.0001	evaluation reward: 2.69	

episode: 1540	score: 5.0	memory length: 303310	epsilon: 0.5974442200087391
steps: 327	lr: 0.0001	evaluation reward: 2.71	
episode: 1541	score: 3.0	memory length: 303521	epsilon: 0.5970264400087482
steps: 211	lr: 0.0001	evaluation reward: 2.72	
episode: 1542	score: 3.0	memory length: 303747	epsilon: 0.5965789600087579
steps: 226	lr: 0.0001	evaluation reward: 2.72	
episode: 1543	score: 4.0	memory length: 304025	epsilon: 0.5960285200087698
steps: 278	lr: 0.0001	evaluation reward: 2.73	
episode: 1544	score: 4.0	memory length: 304285	epsilon: 0.595513720008781
steps: 260	lr: 0.0001	evaluation reward: 2.75	
episode: 1545	score: 5.0	memory length: 304611	epsilon: 0.594868240008795
steps: 326	lr: 0.0001	evaluation reward: 2.79	
episode: 1546	score: 2.0	memory length: 304793	epsilon: 0.5945078800088028
steps: 182	lr: 0.0001	evaluation reward: 2.81	
episode: 1547	score: 4.0	memory length: 305068	epsilon: 0.5939633800088147
steps: 275	lr: 0.0001	evaluation reward: 2.83	
episode: 1548	score: 4.0	memory length: 305346	epsilon: 0.5934129400088266
steps: 278	lr: 0.0001	evaluation reward: 2.82	
episode: 1549	score: 1.0	memory length: 305497	epsilon: 0.5931139600088331
steps: 151	lr: 0.0001	evaluation reward: 2.79	
episode: 1550	score: 4.0	memory length: 305772	epsilon: 0.5925694600088449
steps: 275	lr: 0.0001	evaluation reward: 2.8	
episode: 1551	score: 1.0	memory length: 305923	epsilon: 0.5922704800088514
steps: 151	lr: 0.0001	evaluation reward: 2.79	
episode: 1552	score: 3.0	memory length: 306152	epsilon: 0.5918170600088613
steps: 229	lr: 0.0001	evaluation reward: 2.79	
episode: 1553	score: 4.0	memory length: 306426	epsilon: 0.591274540008873
steps: 274	lr: 0.0001	evaluation reward: 2.77	
episode: 1554	score: 1.0	memory length: 306597	epsilon: 0.5909359600088804
steps: 171	lr: 0.0001	evaluation reward: 2.77	
episode: 1555	score: 2.0	memory length: 306795	epsilon: 0.5905439200088889
steps: 198	lr: 0.0001	evaluation reward: 2.75	
episode: 1556	score: 4.0	memory length: 307070	epsilon: 0.5899994200089007
steps: 275	lr: 0.0001	evaluation reward: 2.79	
episode: 1557	score: 4.0	memory length: 307362	epsilon: 0.5894212600089133
steps: 292	lr: 0.0001	evaluation reward: 2.81	
episode: 1558	score: 2.0	memory length: 307580	epsilon: 0.5889896200089226
steps: 218	lr: 0.0001	evaluation reward: 2.83	
episode: 1559	score: 6.0	memory length: 307994	epsilon: 0.5881699000089404
steps: 414	lr: 0.0001	evaluation reward: 2.86	
episode: 1560	score: 2.0	memory length: 308192	epsilon: 0.5877778600089489
steps: 198	lr: 0.0001	evaluation reward: 2.87	
episode: 1561	score: 2.0	memory length: 308390	epsilon: 0.5873858200089574
steps: 198	lr: 0.0001	evaluation reward: 2.85	
episode: 1562	score: 4.0	memory length: 308683	epsilon: 0.58680568000897
steps: 293	lr: 0.0001	evaluation reward: 2.89	
episode: 1563	score: 2.0	memory length: 308881	epsilon: 0.5864136400089786
steps: 198	lr: 0.0001	evaluation reward: 2.86	

episode: 1564	score: 0.0	memory length: 309004	epsilon: 0.5861701000089838
steps: 123	lr: 0.0001	evaluation reward: 2.82	
episode: 1565	score: 2.0	memory length: 309201	epsilon: 0.5857800400089923
steps: 197	lr: 0.0001	evaluation reward: 2.83	
episode: 1566	score: 0.0	memory length: 309324	epsilon: 0.5855365000089976
steps: 123	lr: 0.0001	evaluation reward: 2.81	
episode: 1567	score: 3.0	memory length: 309570	epsilon: 0.5850494200090082
steps: 246	lr: 0.0001	evaluation reward: 2.82	
episode: 1568	score: 3.0	memory length: 309796	epsilon: 0.5846019400090179
steps: 226	lr: 0.0001	evaluation reward: 2.82	
episode: 1569	score: 4.0	memory length: 310072	epsilon: 0.5840554600090297
steps: 276	lr: 0.0001	evaluation reward: 2.84	
episode: 1570	score: 5.0	memory length: 310397	epsilon: 0.5834119600090437
steps: 325	lr: 0.0001	evaluation reward: 2.87	
episode: 1571	score: 3.0	memory length: 310623	epsilon: 0.5829644800090534
steps: 226	lr: 0.0001	evaluation reward: 2.87	
episode: 1572	score: 5.0	memory length: 310947	epsilon: 0.5823229600090674
steps: 324	lr: 0.0001	evaluation reward: 2.88	
episode: 1573	score: 2.0	memory length: 311145	epsilon: 0.5819309200090759
steps: 198	lr: 0.0001	evaluation reward: 2.88	
episode: 1574	score: 2.0	memory length: 311343	epsilon: 0.5815388800090844
steps: 198	lr: 0.0001	evaluation reward: 2.88	
episode: 1575	score: 2.0	memory length: 311525	epsilon: 0.5811785200090922
steps: 182	lr: 0.0001	evaluation reward: 2.86	
episode: 1576	score: 5.0	memory length: 311834	epsilon: 0.5805667000091055
steps: 309	lr: 0.0001	evaluation reward: 2.89	
episode: 1577	score: 0.0	memory length: 311956	epsilon: 0.5803251400091107
steps: 122	lr: 0.0001	evaluation reward: 2.87	
episode: 1578	score: 4.0	memory length: 312251	epsilon: 0.5797410400091234
steps: 295	lr: 0.0001	evaluation reward: 2.89	
episode: 1579	score: 0.0	memory length: 312374	epsilon: 0.5794975000091287
steps: 123	lr: 0.0001	evaluation reward: 2.87	
episode: 1580	score: 4.0	memory length: 312652	epsilon: 0.5789470600091406
steps: 278	lr: 0.0001	evaluation reward: 2.87	
episode: 1581	score: 2.0	memory length: 312834	epsilon: 0.5785867000091485
steps: 182	lr: 0.0001	evaluation reward: 2.88	
episode: 1582	score: 3.0	memory length: 313081	epsilon: 0.5780976400091591
steps: 247	lr: 0.0001	evaluation reward: 2.9	
episode: 1583	score: 1.0	memory length: 313232	epsilon: 0.5777986600091656
steps: 151	lr: 0.0001	evaluation reward: 2.87	
episode: 1584	score: 4.0	memory length: 313489	epsilon: 0.5772898000091766
steps: 257	lr: 0.0001	evaluation reward: 2.89	
episode: 1585	score: 3.0	memory length: 313738	epsilon: 0.5767967800091873
steps: 249	lr: 0.0001	evaluation reward: 2.9	
episode: 1586	score: 3.0	memory length: 313963	epsilon: 0.576351280009197
steps: 225	lr: 0.0001	evaluation reward: 2.91	
episode: 1587	score: 3.0	memory length: 314189	epsilon: 0.5759038000092067
steps: 226	lr: 0.0001	evaluation reward: 2.94	

episode: 1588	score: 3.0	memory length: 314419	epsilon: 0.5754484000092166
steps: 230	lr: 0.0001	evaluation reward: 2.93	
episode: 1589	score: 3.0	memory length: 314645	epsilon: 0.5750009200092263
steps: 226	lr: 0.0001	evaluation reward: 2.93	
episode: 1590	score: 3.0	memory length: 314871	epsilon: 0.574553440009236
steps: 226	lr: 0.0001	evaluation reward: 2.94	
episode: 1591	score: 3.0	memory length: 315097	epsilon: 0.5741059600092457
steps: 226	lr: 0.0001	evaluation reward: 2.95	
episode: 1592	score: 3.0	memory length: 315322	epsilon: 0.5736604600092554
steps: 225	lr: 0.0001	evaluation reward: 2.94	
episode: 1593	score: 4.0	memory length: 315597	epsilon: 0.5731159600092672
steps: 275	lr: 0.0001	evaluation reward: 2.96	
episode: 1594	score: 3.0	memory length: 315822	epsilon: 0.5726704600092769
steps: 225	lr: 0.0001	evaluation reward: 2.97	
episode: 1595	score: 3.0	memory length: 316050	epsilon: 0.5722190200092867
steps: 228	lr: 0.0001	evaluation reward: 2.98	
episode: 1596	score: 0.0	memory length: 316173	epsilon: 0.571975480009292
steps: 123	lr: 0.0001	evaluation reward: 2.97	
episode: 1597	score: 3.0	memory length: 316383	epsilon: 0.571559680009301
steps: 210	lr: 0.0001	evaluation reward: 2.98	
episode: 1598	score: 3.0	memory length: 316609	epsilon: 0.5711122000093107
steps: 226	lr: 0.0001	evaluation reward: 2.97	
episode: 1599	score: 2.0	memory length: 316791	epsilon: 0.5707518400093186
steps: 182	lr: 0.0001	evaluation reward: 2.96	
episode: 1600	score: 5.0	memory length: 317115	epsilon: 0.5701103200093325
steps: 324	lr: 0.0001	evaluation reward: 3.0	
episode: 1601	score: 2.0	memory length: 317331	epsilon: 0.5696826400093418
steps: 216	lr: 0.0001	evaluation reward: 2.93	
episode: 1602	score: 4.0	memory length: 317647	epsilon: 0.5690569600093554
steps: 316	lr: 0.0001	evaluation reward: 2.93	
episode: 1603	score: 3.0	memory length: 317892	epsilon: 0.5685718600093659
steps: 245	lr: 0.0001	evaluation reward: 2.93	
episode: 1604	score: 1.0	memory length: 318062	epsilon: 0.5682352600093732
steps: 170	lr: 0.0001	evaluation reward: 2.88	
episode: 1605	score: 2.0	memory length: 318242	epsilon: 0.5678788600093809
steps: 180	lr: 0.0001	evaluation reward: 2.87	
episode: 1606	score: 3.0	memory length: 318468	epsilon: 0.5674313800093906
steps: 226	lr: 0.0001	evaluation reward: 2.87	
episode: 1607	score: 5.0	memory length: 318832	epsilon: 0.5667106600094063
steps: 364	lr: 0.0001	evaluation reward: 2.92	
episode: 1608	score: 3.0	memory length: 319045	epsilon: 0.5662889200094154
steps: 213	lr: 0.0001	evaluation reward: 2.93	
episode: 1609	score: 2.0	memory length: 319243	epsilon: 0.565896880009424
steps: 198	lr: 0.0001	evaluation reward: 2.93	
episode: 1610	score: 4.0	memory length: 319518	epsilon: 0.5653523800094358
steps: 275	lr: 0.0001	evaluation reward: 2.95	
episode: 1611	score: 1.0	memory length: 319686	epsilon: 0.565019740009443
steps: 168	lr: 0.0001	evaluation reward: 2.93	

episode: 1612	score: 3.0	memory length: 319912	epsilon: 0.5645722600094527
steps: 226	lr: 0.0001	evaluation reward: 2.92	
episode: 1613	score: 3.0	memory length: 320122	epsilon: 0.5641564600094617
steps: 210	lr: 0.0001	evaluation reward: 2.91	
episode: 1614	score: 0.0	memory length: 320245	epsilon: 0.563912920009467
steps: 123	lr: 0.0001	evaluation reward: 2.89	
episode: 1615	score: 3.0	memory length: 320475	epsilon: 0.5634575200094769
steps: 230	lr: 0.0001	evaluation reward: 2.9	
episode: 1616	score: 3.0	memory length: 320701	epsilon: 0.5630100400094866
steps: 226	lr: 0.0001	evaluation reward: 2.9	
episode: 1617	score: 3.0	memory length: 320926	epsilon: 0.5625645400094963
steps: 225	lr: 0.0001	evaluation reward: 2.92	
episode: 1618	score: 3.0	memory length: 321152	epsilon: 0.562117060009506
steps: 226	lr: 0.0001	evaluation reward: 2.93	
episode: 1619	score: 2.0	memory length: 321352	epsilon: 0.5617210600095146
steps: 200	lr: 0.0001	evaluation reward: 2.91	
episode: 1620	score: 4.0	memory length: 321608	epsilon: 0.5612141800095256
steps: 256	lr: 0.0001	evaluation reward: 2.9	
episode: 1621	score: 2.0	memory length: 321826	epsilon: 0.560782540009535
steps: 218	lr: 0.0001	evaluation reward: 2.89	
episode: 1622	score: 5.0	memory length: 322132	epsilon: 0.5601766600095481
steps: 306	lr: 0.0001	evaluation reward: 2.92	
episode: 1623	score: 3.0	memory length: 322358	epsilon: 0.5597291800095578
steps: 226	lr: 0.0001	evaluation reward: 2.91	
episode: 1624	score: 1.0	memory length: 322509	epsilon: 0.5594302000095643
steps: 151	lr: 0.0001	evaluation reward: 2.9	
episode: 1625	score: 0.0	memory length: 322632	epsilon: 0.5591866600095696
steps: 123	lr: 0.0001	evaluation reward: 2.87	
episode: 1626	score: 2.0	memory length: 322814	epsilon: 0.5588263000095774
steps: 182	lr: 0.0001	evaluation reward: 2.87	
episode: 1627	score: 2.0	memory length: 323032	epsilon: 0.5583946600095868
steps: 218	lr: 0.0001	evaluation reward: 2.87	
episode: 1628	score: 6.0	memory length: 323418	epsilon: 0.5576303800096034
steps: 386	lr: 0.0001	evaluation reward: 2.9	
episode: 1629	score: 4.0	memory length: 323677	epsilon: 0.5571175600096145
steps: 259	lr: 0.0001	evaluation reward: 2.92	
episode: 1630	score: 3.0	memory length: 323888	epsilon: 0.5566997800096236
steps: 211	lr: 0.0001	evaluation reward: 2.94	
episode: 1631	score: 5.0	memory length: 324213	epsilon: 0.5560562800096376
steps: 325	lr: 0.0001	evaluation reward: 2.94	
episode: 1632	score: 1.0	memory length: 324382	epsilon: 0.5557216600096448
steps: 169	lr: 0.0001	evaluation reward: 2.9	
episode: 1633	score: 3.0	memory length: 324609	epsilon: 0.5552722000096546
steps: 227	lr: 0.0001	evaluation reward: 2.88	
episode: 1634	score: 1.0	memory length: 324777	epsilon: 0.5549395600096618
steps: 168	lr: 0.0001	evaluation reward: 2.85	
episode: 1635	score: 3.0	memory length: 324990	epsilon: 0.554517820009671
steps: 213	lr: 0.0001	evaluation reward: 2.85	

episode: 1636	score: 5.0	memory length: 325297	epsilon: 0.5539099600096842
steps: 307	lr: 0.0001	evaluation reward: 2.87	
episode: 1637	score: 2.0	memory length: 325495	epsilon: 0.5535179200096927
steps: 198	lr: 0.0001	evaluation reward: 2.87	
episode: 1638	score: 3.0	memory length: 325721	epsilon: 0.5530704400097024
steps: 226	lr: 0.0001	evaluation reward: 2.88	
episode: 1639	score: 1.0	memory length: 325872	epsilon: 0.5527714600097089
steps: 151	lr: 0.0001	evaluation reward: 2.83	
episode: 1640	score: 2.0	memory length: 326070	epsilon: 0.5523794200097174
steps: 198	lr: 0.0001	evaluation reward: 2.8	
episode: 1641	score: 4.0	memory length: 326366	epsilon: 0.5517933400097301
steps: 296	lr: 0.0001	evaluation reward: 2.81	
episode: 1642	score: 3.0	memory length: 326592	epsilon: 0.5513458600097398
steps: 226	lr: 0.0001	evaluation reward: 2.81	
episode: 1643	score: 3.0	memory length: 326821	epsilon: 0.5508924400097497
steps: 229	lr: 0.0001	evaluation reward: 2.8	
episode: 1644	score: 2.0	memory length: 327018	epsilon: 0.5505023800097582
steps: 197	lr: 0.0001	evaluation reward: 2.78	
episode: 1645	score: 0.0	memory length: 327141	epsilon: 0.5502588400097634
steps: 123	lr: 0.0001	evaluation reward: 2.73	
episode: 1646	score: 4.0	memory length: 327418	epsilon: 0.5497103800097753
steps: 277	lr: 0.0001	evaluation reward: 2.75	
episode: 1647	score: 2.0	memory length: 327615	epsilon: 0.5493203200097838
steps: 197	lr: 0.0001	evaluation reward: 2.73	
episode: 1648	score: 3.0	memory length: 327841	epsilon: 0.5488728400097935
steps: 226	lr: 0.0001	evaluation reward: 2.72	
episode: 1649	score: 2.0	memory length: 328022	epsilon: 0.5485144600098013
steps: 181	lr: 0.0001	evaluation reward: 2.73	
episode: 1650	score: 1.0	memory length: 328192	epsilon: 0.5481778600098086
steps: 170	lr: 0.0001	evaluation reward: 2.7	
episode: 1651	score: 3.0	memory length: 328441	epsilon: 0.5476848400098193
steps: 249	lr: 0.0001	evaluation reward: 2.72	
episode: 1652	score: 4.0	memory length: 328703	epsilon: 0.5471660800098306
steps: 262	lr: 0.0001	evaluation reward: 2.73	
episode: 1653	score: 3.0	memory length: 328915	epsilon: 0.5467463200098397
steps: 212	lr: 0.0001	evaluation reward: 2.72	
episode: 1654	score: 4.0	memory length: 329189	epsilon: 0.5462038000098515
steps: 274	lr: 0.0001	evaluation reward: 2.75	
episode: 1655	score: 2.0	memory length: 329405	epsilon: 0.5457761200098608
steps: 216	lr: 0.0001	evaluation reward: 2.75	
episode: 1656	score: 3.0	memory length: 329649	epsilon: 0.5452930000098712
steps: 244	lr: 0.0001	evaluation reward: 2.74	
episode: 1657	score: 3.0	memory length: 329875	epsilon: 0.544845520009881
steps: 226	lr: 0.0001	evaluation reward: 2.73	
episode: 1658	score: 2.0	memory length: 330096	epsilon: 0.5444079400098905
steps: 221	lr: 0.0001	evaluation reward: 2.73	
episode: 1659	score: 1.0	memory length: 330247	epsilon: 0.544108960009897
steps: 151	lr: 0.0001	evaluation reward: 2.68	

episode: 1660	score: 0.0	memory length: 330369	epsilon: 0.5438674000099022
steps: 122	lr: 0.0001	evaluation reward: 2.66	
episode: 1661	score: 3.0	memory length: 330595	epsilon: 0.5434199200099119
steps: 226	lr: 0.0001	evaluation reward: 2.67	
episode: 1662	score: 0.0	memory length: 330717	epsilon: 0.5431783600099171
steps: 122	lr: 0.0001	evaluation reward: 2.63	
episode: 1663	score: 2.0	memory length: 330915	epsilon: 0.5427863200099257
steps: 198	lr: 0.0001	evaluation reward: 2.63	
episode: 1664	score: 3.0	memory length: 331160	epsilon: 0.5423012200099362
steps: 245	lr: 0.0001	evaluation reward: 2.66	
episode: 1665	score: 2.0	memory length: 331360	epsilon: 0.5419052200099448
steps: 200	lr: 0.0001	evaluation reward: 2.66	
episode: 1666	score: 4.0	memory length: 331636	epsilon: 0.5413587400099567
steps: 276	lr: 0.0001	evaluation reward: 2.7	
episode: 1667	score: 4.0	memory length: 331893	epsilon: 0.5408498800099677
steps: 257	lr: 0.0001	evaluation reward: 2.71	
episode: 1668	score: 2.0	memory length: 332091	epsilon: 0.5404578400099762
steps: 198	lr: 0.0001	evaluation reward: 2.7	
episode: 1669	score: 3.0	memory length: 332340	epsilon: 0.5399648200099869
steps: 249	lr: 0.0001	evaluation reward: 2.69	
episode: 1670	score: 3.0	memory length: 332553	epsilon: 0.5395430800099961
steps: 213	lr: 0.0001	evaluation reward: 2.67	
episode: 1671	score: 0.0	memory length: 332675	epsilon: 0.5393015200100013
steps: 122	lr: 0.0001	evaluation reward: 2.64	
episode: 1672	score: 7.0	memory length: 333114	epsilon: 0.5384323000100202
steps: 439	lr: 0.0001	evaluation reward: 2.66	
episode: 1673	score: 1.0	memory length: 333265	epsilon: 0.5381333200100267
steps: 151	lr: 0.0001	evaluation reward: 2.65	
episode: 1674	score: 3.0	memory length: 333494	epsilon: 0.5376799000100365
steps: 229	lr: 0.0001	evaluation reward: 2.66	
episode: 1675	score: 2.0	memory length: 333676	epsilon: 0.5373195400100443
steps: 182	lr: 0.0001	evaluation reward: 2.66	
episode: 1676	score: 4.0	memory length: 333953	epsilon: 0.5367710800100562
steps: 277	lr: 0.0001	evaluation reward: 2.65	
episode: 1677	score: 5.0	memory length: 334278	epsilon: 0.5361275800100702
steps: 325	lr: 0.0001	evaluation reward: 2.7	
episode: 1678	score: 5.0	memory length: 334594	epsilon: 0.5355019000100838
steps: 316	lr: 0.0001	evaluation reward: 2.71	
episode: 1679	score: 5.0	memory length: 334870	epsilon: 0.5349554200100957
steps: 276	lr: 0.0001	evaluation reward: 2.76	
episode: 1680	score: 1.0	memory length: 335040	epsilon: 0.534618820010103
steps: 170	lr: 0.0001	evaluation reward: 2.73	
episode: 1681	score: 1.0	memory length: 335191	epsilon: 0.5343198400101095
steps: 151	lr: 0.0001	evaluation reward: 2.72	
episode: 1682	score: 2.0	memory length: 335388	epsilon: 0.5339297800101179
steps: 197	lr: 0.0001	evaluation reward: 2.71	
episode: 1683	score: 0.0	memory length: 335510	epsilon: 0.5336882200101232
steps: 122	lr: 0.0001	evaluation reward: 2.7	

episode: 1684	score: 3.0	memory length: 335756	epsilon: 0.5332011400101337
steps: 246	lr: 0.0001	evaluation reward: 2.69	
episode: 1685	score: 3.0	memory length: 335985	epsilon: 0.5327477200101436
steps: 229	lr: 0.0001	evaluation reward: 2.69	
episode: 1686	score: 2.0	memory length: 336164	epsilon: 0.5323933000101513
steps: 179	lr: 0.0001	evaluation reward: 2.68	
episode: 1687	score: 4.0	memory length: 336460	epsilon: 0.531807220010164
steps: 296	lr: 0.0001	evaluation reward: 2.69	
episode: 1688	score: 0.0	memory length: 336583	epsilon: 0.5315636800101693
steps: 123	lr: 0.0001	evaluation reward: 2.66	
episode: 1689	score: 3.0	memory length: 336815	epsilon: 0.5311043200101793
steps: 232	lr: 0.0001	evaluation reward: 2.66	
episode: 1690	score: 3.0	memory length: 337041	epsilon: 0.530656840010189
steps: 226	lr: 0.0001	evaluation reward: 2.66	
episode: 1691	score: 3.0	memory length: 337271	epsilon: 0.5302014400101989
steps: 230	lr: 0.0001	evaluation reward: 2.66	
episode: 1692	score: 1.0	memory length: 337422	epsilon: 0.5299024600102054
steps: 151	lr: 0.0001	evaluation reward: 2.64	
episode: 1693	score: 2.0	memory length: 337620	epsilon: 0.5295104200102139
steps: 198	lr: 0.0001	evaluation reward: 2.62	
episode: 1694	score: 6.0	memory length: 337962	epsilon: 0.5288332600102286
steps: 342	lr: 0.0001	evaluation reward: 2.65	
episode: 1695	score: 3.0	memory length: 338175	epsilon: 0.5284115200102377
steps: 213	lr: 0.0001	evaluation reward: 2.65	
episode: 1696	score: 5.0	memory length: 338461	epsilon: 0.52784524001025
steps: 286	lr: 0.0001	evaluation reward: 2.7	
episode: 1697	score: 2.0	memory length: 338659	epsilon: 0.5274532000102585
steps: 198	lr: 0.0001	evaluation reward: 2.69	
episode: 1698	score: 3.0	memory length: 338885	epsilon: 0.5270057200102682
steps: 226	lr: 0.0001	evaluation reward: 2.69	
episode: 1699	score: 2.0	memory length: 339066	epsilon: 0.526647340010276
steps: 181	lr: 0.0001	evaluation reward: 2.69	
episode: 1700	score: 4.0	memory length: 339347	epsilon: 0.5260909600102881
steps: 281	lr: 0.0001	evaluation reward: 2.68	
episode: 1701	score: 3.0	memory length: 339577	epsilon: 0.525635560010298
steps: 230	lr: 0.0001	evaluation reward: 2.69	
episode: 1702	score: 2.0	memory length: 339775	epsilon: 0.5252435200103065
steps: 198	lr: 0.0001	evaluation reward: 2.67	
episode: 1703	score: 3.0	memory length: 339988	epsilon: 0.5248217800103157
steps: 213	lr: 0.0001	evaluation reward: 2.67	
episode: 1704	score: 2.0	memory length: 340170	epsilon: 0.5244614200103235
steps: 182	lr: 0.0001	evaluation reward: 2.68	
episode: 1705	score: 5.0	memory length: 340460	epsilon: 0.5238872200103359
steps: 290	lr: 0.0001	evaluation reward: 2.71	
episode: 1706	score: 1.0	memory length: 340611	epsilon: 0.5235882400103424
steps: 151	lr: 0.0001	evaluation reward: 2.69	
episode: 1707	score: 3.0	memory length: 340824	epsilon: 0.5231665000103516
steps: 213	lr: 0.0001	evaluation reward: 2.67	

episode: 1708	score: 3.0	memory length: 341071	epsilon: 0.5226774400103622
steps: 247	lr: 0.0001	evaluation reward: 2.67	
episode: 1709	score: 3.0	memory length: 341300	epsilon: 0.522224020010372
steps: 229	lr: 0.0001	evaluation reward: 2.68	
episode: 1710	score: 5.0	memory length: 341579	epsilon: 0.521671600010384
steps: 279	lr: 0.0001	evaluation reward: 2.69	
episode: 1711	score: 2.0	memory length: 341760	epsilon: 0.5213132200103918
steps: 181	lr: 0.0001	evaluation reward: 2.7	
episode: 1712	score: 4.0	memory length: 342074	epsilon: 0.5206915000104053
steps: 314	lr: 0.0001	evaluation reward: 2.71	
episode: 1713	score: 3.0	memory length: 342319	epsilon: 0.5202064000104158
steps: 245	lr: 0.0001	evaluation reward: 2.71	
episode: 1714	score: 4.0	memory length: 342609	epsilon: 0.5196322000104283
steps: 290	lr: 0.0001	evaluation reward: 2.75	
episode: 1715	score: 3.0	memory length: 342820	epsilon: 0.5192144200104374
steps: 211	lr: 0.0001	evaluation reward: 2.75	
episode: 1716	score: 2.0	memory length: 343001	epsilon: 0.5188560400104452
steps: 181	lr: 0.0001	evaluation reward: 2.74	
episode: 1717	score: 3.0	memory length: 343246	epsilon: 0.5183709400104557
steps: 245	lr: 0.0001	evaluation reward: 2.74	
episode: 1718	score: 3.0	memory length: 343457	epsilon: 0.5179531600104648
steps: 211	lr: 0.0001	evaluation reward: 2.74	
episode: 1719	score: 2.0	memory length: 343639	epsilon: 0.5175928000104726
steps: 182	lr: 0.0001	evaluation reward: 2.74	
episode: 1720	score: 2.0	memory length: 343819	epsilon: 0.5172364000104803
steps: 180	lr: 0.0001	evaluation reward: 2.72	
episode: 1721	score: 2.0	memory length: 343999	epsilon: 0.5168800000104881
steps: 180	lr: 0.0001	evaluation reward: 2.72	
episode: 1722	score: 2.0	memory length: 344197	epsilon: 0.5164879600104966
steps: 198	lr: 0.0001	evaluation reward: 2.69	
episode: 1723	score: 2.0	memory length: 344395	epsilon: 0.5160959200105051
steps: 198	lr: 0.0001	evaluation reward: 2.68	
episode: 1724	score: 1.0	memory length: 344545	epsilon: 0.5157989200105115
steps: 150	lr: 0.0001	evaluation reward: 2.68	
episode: 1725	score: 3.0	memory length: 344791	epsilon: 0.5153118400105221
steps: 246	lr: 0.0001	evaluation reward: 2.71	
episode: 1726	score: 2.0	memory length: 344972	epsilon: 0.5149534600105299
steps: 181	lr: 0.0001	evaluation reward: 2.71	
episode: 1727	score: 4.0	memory length: 345268	epsilon: 0.5143673800105426
steps: 296	lr: 0.0001	evaluation reward: 2.73	
episode: 1728	score: 4.0	memory length: 345544	epsilon: 0.5138209000105545
steps: 276	lr: 0.0001	evaluation reward: 2.71	
episode: 1729	score: 2.0	memory length: 345741	epsilon: 0.5134308400105629
steps: 197	lr: 0.0001	evaluation reward: 2.69	
episode: 1730	score: 5.0	memory length: 346086	epsilon: 0.5127477400105778
steps: 345	lr: 0.0001	evaluation reward: 2.71	
episode: 1731	score: 2.0	memory length: 346283	epsilon: 0.5123576800105862
steps: 197	lr: 0.0001	evaluation reward: 2.68	

episode: 1732	score: 5.0	memory length: 346608	epsilon: 0.5117141800106002
steps: 325	lr: 0.0001	evaluation reward: 2.72	
episode: 1733	score: 5.0	memory length: 346915	epsilon: 0.5111063200106134
steps: 307	lr: 0.0001	evaluation reward: 2.74	
episode: 1734	score: 4.0	memory length: 347190	epsilon: 0.5105618200106252
steps: 275	lr: 0.0001	evaluation reward: 2.77	
episode: 1735	score: 2.0	memory length: 347390	epsilon: 0.5101658200106338
steps: 200	lr: 0.0001	evaluation reward: 2.76	
episode: 1736	score: 5.0	memory length: 347757	epsilon: 0.5094391600106496
steps: 367	lr: 0.0001	evaluation reward: 2.76	
episode: 1737	score: 3.0	memory length: 347985	epsilon: 0.5089877200106594
steps: 228	lr: 0.0001	evaluation reward: 2.77	
episode: 1738	score: 2.0	memory length: 348167	epsilon: 0.5086273600106672
steps: 182	lr: 0.0001	evaluation reward: 2.76	
episode: 1739	score: 2.0	memory length: 348365	epsilon: 0.5082353200106757
steps: 198	lr: 0.0001	evaluation reward: 2.77	
episode: 1740	score: 3.0	memory length: 348609	epsilon: 0.5077522000106862
steps: 244	lr: 0.0001	evaluation reward: 2.78	
episode: 1741	score: 5.0	memory length: 348920	epsilon: 0.5071364200106996
steps: 311	lr: 0.0001	evaluation reward: 2.79	
episode: 1742	score: 2.0	memory length: 349102	epsilon: 0.5067760600107074
steps: 182	lr: 0.0001	evaluation reward: 2.78	
episode: 1743	score: 3.0	memory length: 349332	epsilon: 0.5063206600107173
steps: 230	lr: 0.0001	evaluation reward: 2.78	
episode: 1744	score: 4.0	memory length: 349620	epsilon: 0.5057504200107297
steps: 288	lr: 0.0001	evaluation reward: 2.8	
episode: 1745	score: 2.0	memory length: 349801	epsilon: 0.5053920400107375
steps: 181	lr: 0.0001	evaluation reward: 2.82	
episode: 1746	score: 3.0	memory length: 350027	epsilon: 0.5049445600107472
steps: 226	lr: 0.0001	evaluation reward: 2.81	
episode: 1747	score: 3.0	memory length: 350256	epsilon: 0.504491140010757
steps: 229	lr: 0.0001	evaluation reward: 2.82	
episode: 1748	score: 2.0	memory length: 350454	epsilon: 0.5040991000107655
steps: 198	lr: 0.0001	evaluation reward: 2.81	
episode: 1749	score: 3.0	memory length: 350685	epsilon: 0.5036417200107755
steps: 231	lr: 0.0001	evaluation reward: 2.82	
episode: 1750	score: 2.0	memory length: 350882	epsilon: 0.5032516600107839
steps: 197	lr: 0.0001	evaluation reward: 2.83	
episode: 1751	score: 5.0	memory length: 351198	epsilon: 0.5026259800107975
steps: 316	lr: 0.0001	evaluation reward: 2.85	
episode: 1752	score: 2.0	memory length: 351396	epsilon: 0.502233940010806
steps: 198	lr: 0.0001	evaluation reward: 2.83	
episode: 1753	score: 8.0	memory length: 351812	epsilon: 0.5014102600108239
steps: 416	lr: 0.0001	evaluation reward: 2.88	
episode: 1754	score: 1.0	memory length: 351983	epsilon: 0.5010716800108312
steps: 171	lr: 0.0001	evaluation reward: 2.85	
episode: 1755	score: 7.0	memory length: 352360	epsilon: 0.5003252200108474
steps: 377	lr: 0.0001	evaluation reward: 2.9	

episode: 1756 score: 2.0 memory length: 352558 epsilon: 0.4999331800108541
 steps: 198 lr: 0.0001 evaluation reward: 2.89
 episode: 1757 score: 4.0 memory length: 352809 epsilon:
 0.49943620001085093 steps: 251 lr: 0.0001 evaluation reward: 2.9
 episode: 1758 score: 1.0 memory length: 352959 epsilon:
 0.49913920001084905 steps: 150 lr: 0.0001 evaluation reward: 2.89
 episode: 1759 score: 5.0 memory length: 353264 epsilon:
 0.49853530001084523 steps: 305 lr: 0.0001 evaluation reward: 2.93
 episode: 1760 score: 3.0 memory length: 353509 epsilon:
 0.49805020001084216 steps: 245 lr: 0.0001 evaluation reward: 2.96
 episode: 1761 score: 3.0 memory length: 353756 epsilon:
 0.49756114001083906 steps: 247 lr: 0.0001 evaluation reward: 2.96
 episode: 1762 score: 2.0 memory length: 353955 epsilon:
 0.49716712001083657 steps: 199 lr: 0.0001 evaluation reward: 2.98
 episode: 1763 score: 4.0 memory length: 354230 epsilon: 0.4966226200108331
 steps: 275 lr: 0.0001 evaluation reward: 3.0
 episode: 1764 score: 2.0 memory length: 354428 epsilon:
 0.49623058001083065 steps: 198 lr: 0.0001 evaluation reward: 2.99
 episode: 1765 score: 2.0 memory length: 354646 epsilon: 0.4957989400108279
 steps: 218 lr: 0.0001 evaluation reward: 2.99
 episode: 1766 score: 4.0 memory length: 354898 epsilon:
 0.49529998001082476 steps: 252 lr: 0.0001 evaluation reward: 2.99
 episode: 1767 score: 1.0 memory length: 355049 epsilon:
 0.49500100001082287 steps: 151 lr: 0.0001 evaluation reward: 2.96
 episode: 1768 score: 5.0 memory length: 355373 epsilon: 0.4943594800108188
 steps: 324 lr: 0.0001 evaluation reward: 2.99
 episode: 1769 score: 2.0 memory length: 355570 epsilon:
 0.49396942001081634 steps: 197 lr: 0.0001 evaluation reward: 2.98
 episode: 1770 score: 2.0 memory length: 355752 epsilon:
 0.49360906001081406 steps: 182 lr: 0.0001 evaluation reward: 2.97
 episode: 1771 score: 4.0 memory length: 356009 epsilon:
 0.49310020001081084 steps: 257 lr: 0.0001 evaluation reward: 3.01
 episode: 1772 score: 5.0 memory length: 356299 epsilon: 0.4925260000108072
 steps: 290 lr: 0.0001 evaluation reward: 2.99
 episode: 1773 score: 1.0 memory length: 356449 epsilon:
 0.49222900001080533 steps: 150 lr: 0.0001 evaluation reward: 2.99
 episode: 1774 score: 2.0 memory length: 356631 epsilon:
 0.49186864001080305 steps: 182 lr: 0.0001 evaluation reward: 2.98
 episode: 1775 score: 5.0 memory length: 356943 epsilon:
 0.49125088001079914 steps: 312 lr: 0.0001 evaluation reward: 3.01
 episode: 1776 score: 4.0 memory length: 357241 epsilon: 0.4906608400107954
 steps: 298 lr: 0.0001 evaluation reward: 3.01
 episode: 1777 score: 3.0 memory length: 357489 epsilon: 0.4901698000107923
 steps: 248 lr: 0.0001 evaluation reward: 2.99
 episode: 1778 score: 4.0 memory length: 357742 epsilon:
 0.48966886001078913 steps: 253 lr: 0.0001 evaluation reward: 2.98
 episode: 1779 score: 0.0 memory length: 357864 epsilon: 0.4894273000107876
 steps: 122 lr: 0.0001 evaluation reward: 2.93

episode: 1780 score: 4.0 memory length: 358160 epsilon: 0.4888412200107839
 steps: 296 lr: 0.0001 evaluation reward: 2.96
 episode: 1781 score: 3.0 memory length: 358371 epsilon:
 0.48842344001078125 steps: 211 lr: 0.0001 evaluation reward: 2.98
 episode: 1782 score: 3.0 memory length: 358596 epsilon:
 0.48797794001077843 steps: 225 lr: 0.0001 evaluation reward: 2.99
 episode: 1783 score: 2.0 memory length: 358778 epsilon:
 0.48761758001077615 steps: 182 lr: 0.0001 evaluation reward: 3.01
 episode: 1784 score: 3.0 memory length: 359025 epsilon:
 0.48712852001077306 steps: 247 lr: 0.0001 evaluation reward: 3.01
 episode: 1785 score: 3.0 memory length: 359253 epsilon: 0.4866770800107702
 steps: 228 lr: 0.0001 evaluation reward: 3.01
 episode: 1786 score: 2.0 memory length: 359435 epsilon: 0.4863167200107679
 steps: 182 lr: 0.0001 evaluation reward: 3.01
 episode: 1787 score: 3.0 memory length: 359647 epsilon:
 0.48589696001076527 steps: 212 lr: 0.0001 evaluation reward: 3.0
 episode: 1788 score: 2.0 memory length: 359845 epsilon: 0.4855049200107628
 steps: 198 lr: 0.0001 evaluation reward: 3.02
 episode: 1789 score: 2.0 memory length: 360027 epsilon: 0.4851445600107605
 steps: 182 lr: 0.0001 evaluation reward: 3.01
 episode: 1790 score: 3.0 memory length: 360272 epsilon:
 0.48465946001075744 steps: 245 lr: 0.0001 evaluation reward: 3.01
 episode: 1791 score: 2.0 memory length: 360452 epsilon: 0.4843030600107552
 steps: 180 lr: 0.0001 evaluation reward: 3.0
 episode: 1792 score: 4.0 memory length: 360729 epsilon: 0.4837546000107517
 steps: 277 lr: 0.0001 evaluation reward: 3.03
 episode: 1793 score: 1.0 memory length: 360880 epsilon: 0.4834556200107498
 steps: 151 lr: 0.0001 evaluation reward: 3.02
 episode: 1794 score: 2.0 memory length: 361061 epsilon:
 0.48309724001074755 steps: 181 lr: 0.0001 evaluation reward: 2.98
 episode: 1795 score: 3.0 memory length: 361309 epsilon:
 0.48260620001074445 steps: 248 lr: 0.0001 evaluation reward: 2.98
 episode: 1796 score: 4.0 memory length: 361607 epsilon: 0.4820161600107407
 steps: 298 lr: 0.0001 evaluation reward: 2.97
 episode: 1797 score: 3.0 memory length: 361833 epsilon: 0.4815686800107379
 steps: 226 lr: 0.0001 evaluation reward: 2.98
 episode: 1798 score: 2.0 memory length: 362030 epsilon: 0.4811786200107354
 steps: 197 lr: 0.0001 evaluation reward: 2.97
 episode: 1799 score: 3.0 memory length: 362256 epsilon: 0.4807311400107326
 steps: 226 lr: 0.0001 evaluation reward: 2.98
 episode: 1800 score: 3.0 memory length: 362507 epsilon:
 0.48023416001072944 steps: 251 lr: 0.0001 evaluation reward: 2.97
 episode: 1801 score: 1.0 memory length: 362659 epsilon:
 0.47993320001072753 steps: 152 lr: 0.0001 evaluation reward: 2.95
 episode: 1802 score: 2.0 memory length: 362841 epsilon:
 0.47957284001072525 steps: 182 lr: 0.0001 evaluation reward: 2.95
 episode: 1803 score: 3.0 memory length: 363072 epsilon:
 0.47911546001072236 steps: 231 lr: 0.0001 evaluation reward: 2.95

episode: 1804 score: 3.0 memory length: 363283 epsilon: 0.4786976800107197
 steps: 211 lr: 0.0001 evaluation reward: 2.96
 episode: 1805 score: 2.0 memory length: 363483 epsilon: 0.4783016800107172
 steps: 200 lr: 0.0001 evaluation reward: 2.93
 episode: 1806 score: 4.0 memory length: 363749 epsilon: 0.4777750000107139
 steps: 266 lr: 0.0001 evaluation reward: 2.96
 episode: 1807 score: 4.0 memory length: 364026 epsilon: 0.4772265400107104
 steps: 277 lr: 0.0001 evaluation reward: 2.97
 episode: 1808 score: 6.0 memory length: 364402 epsilon: 0.4764820600107057
 steps: 376 lr: 0.0001 evaluation reward: 3.0
 episode: 1809 score: 0.0 memory length: 364525 epsilon:
 0.47623852001070416 steps: 123 lr: 0.0001 evaluation reward: 2.97
 episode: 1810 score: 2.0 memory length: 364722 epsilon: 0.4758484600107017
 steps: 197 lr: 0.0001 evaluation reward: 2.94
 episode: 1811 score: 4.0 memory length: 364981 epsilon:
 0.47533564001069845 steps: 259 lr: 0.0001 evaluation reward: 2.96
 episode: 1812 score: 5.0 memory length: 365297 epsilon: 0.4747099600106945
 steps: 316 lr: 0.0001 evaluation reward: 2.97
 episode: 1813 score: 3.0 memory length: 365523 epsilon:
 0.47426248001069166 steps: 226 lr: 0.0001 evaluation reward: 2.97
 episode: 1814 score: 2.0 memory length: 365721 epsilon: 0.4738704400106892
 steps: 198 lr: 0.0001 evaluation reward: 2.95
 episode: 1815 score: 3.0 memory length: 365934 epsilon: 0.4734487000106865
 steps: 213 lr: 0.0001 evaluation reward: 2.95
 episode: 1816 score: 4.0 memory length: 366250 epsilon:
 0.47282302001068255 steps: 316 lr: 0.0001 evaluation reward: 2.97
 episode: 1817 score: 2.0 memory length: 366448 epsilon:
 0.47243098001068007 steps: 198 lr: 0.0001 evaluation reward: 2.96
 episode: 1818 score: 2.0 memory length: 366627 epsilon: 0.4720765600106778
 steps: 179 lr: 0.0001 evaluation reward: 2.95
 episode: 1819 score: 4.0 memory length: 366871 epsilon:
 0.47159344001067477 steps: 244 lr: 0.0001 evaluation reward: 2.97
 episode: 1820 score: 4.0 memory length: 367145 epsilon:
 0.47105092001067134 steps: 274 lr: 0.0001 evaluation reward: 2.99
 episode: 1821 score: 4.0 memory length: 367400 epsilon:
 0.47054602001066814 steps: 255 lr: 0.0001 evaluation reward: 3.01
 episode: 1822 score: 4.0 memory length: 367675 epsilon: 0.4700015200106647
 steps: 275 lr: 0.0001 evaluation reward: 3.03
 episode: 1823 score: 5.0 memory length: 368002 epsilon: 0.4693540600106606
 steps: 327 lr: 0.0001 evaluation reward: 3.06
 episode: 1824 score: 4.0 memory length: 368277 epsilon:
 0.46880956001065716 steps: 275 lr: 0.0001 evaluation reward: 3.09
 episode: 1825 score: 3.0 memory length: 368523 epsilon: 0.4683224800106541
 steps: 246 lr: 0.0001 evaluation reward: 3.09
 episode: 1826 score: 3.0 memory length: 368768 epsilon: 0.467837380010651
 steps: 245 lr: 0.0001 evaluation reward: 3.1
 episode: 1827 score: 2.0 memory length: 368949 epsilon:
 0.46747900001064874 steps: 181 lr: 0.0001 evaluation reward: 3.08

episode: 1828 score: 3.0 memory length: 369175 epsilon: 0.4670315200106459
 steps: 226 lr: 0.0001 evaluation reward: 3.07
 episode: 1829 score: 2.0 memory length: 369373 epsilon: 0.4666394800106434
 steps: 198 lr: 0.0001 evaluation reward: 3.07
 episode: 1830 score: 2.0 memory length: 369555 epsilon:
 0.46627912001064115 steps: 182 lr: 0.0001 evaluation reward: 3.04
 episode: 1831 score: 3.0 memory length: 369781 epsilon: 0.4658316400106383
 steps: 226 lr: 0.0001 evaluation reward: 3.05
 episode: 1832 score: 4.0 memory length: 370058 epsilon:
 0.46528318001063484 steps: 277 lr: 0.0001 evaluation reward: 3.04
 episode: 1833 score: 3.0 memory length: 370268 epsilon: 0.4648673800106322
 steps: 210 lr: 0.0001 evaluation reward: 3.02
 episode: 1834 score: 3.0 memory length: 370497 epsilon:
 0.46441396001062935 steps: 229 lr: 0.0001 evaluation reward: 3.01
 episode: 1835 score: 3.0 memory length: 370726 epsilon: 0.4639605400106265
 steps: 229 lr: 0.0001 evaluation reward: 3.02
 episode: 1836 score: 2.0 memory length: 370908 epsilon: 0.4636001800106242
 steps: 182 lr: 0.0001 evaluation reward: 2.99
 episode: 1837 score: 3.0 memory length: 371135 epsilon:
 0.46315072001062135 steps: 227 lr: 0.0001 evaluation reward: 2.99
 episode: 1838 score: 3.0 memory length: 371361 epsilon: 0.4627032400106185
 steps: 226 lr: 0.0001 evaluation reward: 3.0
 episode: 1839 score: 3.0 memory length: 371586 epsilon: 0.4622577400106157
 steps: 225 lr: 0.0001 evaluation reward: 3.01
 episode: 1840 score: 3.0 memory length: 371812 epsilon:
 0.46181026001061287 steps: 226 lr: 0.0001 evaluation reward: 3.01
 episode: 1841 score: 2.0 memory length: 372010 epsilon: 0.4614182200106104
 steps: 198 lr: 0.0001 evaluation reward: 2.98
 episode: 1842 score: 3.0 memory length: 372256 epsilon: 0.4609311400106073
 steps: 246 lr: 0.0001 evaluation reward: 2.99
 episode: 1843 score: 2.0 memory length: 372454 epsilon:
 0.46053910001060483 steps: 198 lr: 0.0001 evaluation reward: 2.98
 episode: 1844 score: 2.0 memory length: 372635 epsilon:
 0.46018072001060256 steps: 181 lr: 0.0001 evaluation reward: 2.96
 episode: 1845 score: 3.0 memory length: 372864 epsilon: 0.4597273000105997
 steps: 229 lr: 0.0001 evaluation reward: 2.97
 episode: 1846 score: 3.0 memory length: 373076 epsilon:
 0.45930754001059704 steps: 212 lr: 0.0001 evaluation reward: 2.97
 episode: 1847 score: 4.0 memory length: 373330 epsilon:
 0.45880462001059386 steps: 254 lr: 0.0001 evaluation reward: 2.98
 episode: 1848 score: 0.0 memory length: 373452 epsilon: 0.4585630600105923
 steps: 122 lr: 0.0001 evaluation reward: 2.96
 episode: 1849 score: 0.0 memory length: 373575 epsilon: 0.4583195200105908
 steps: 123 lr: 0.0001 evaluation reward: 2.93
 episode: 1850 score: 3.0 memory length: 373788 epsilon: 0.4578977800105881
 steps: 213 lr: 0.0001 evaluation reward: 2.94
 episode: 1851 score: 3.0 memory length: 374035 epsilon: 0.457408720010585
 steps: 247 lr: 0.0001 evaluation reward: 2.92

episode: 1852 score: 3.0 memory length: 374260 epsilon: 0.4569632200105822
 steps: 225 lr: 0.0001 evaluation reward: 2.93
 episode: 1853 score: 3.0 memory length: 374490 epsilon: 0.4565078200105793
 steps: 230 lr: 0.0001 evaluation reward: 2.88
 episode: 1854 score: 1.0 memory length: 374661 epsilon: 0.4561692400105772
 steps: 171 lr: 0.0001 evaluation reward: 2.88
 episode: 1855 score: 3.0 memory length: 374891 epsilon: 0.4557138400105743
 steps: 230 lr: 0.0001 evaluation reward: 2.84
 episode: 1856 score: 3.0 memory length: 375137 epsilon: 0.4552267600105712
 steps: 246 lr: 0.0001 evaluation reward: 2.85
 episode: 1857 score: 2.0 memory length: 375336 epsilon: 0.4548327400105687
 steps: 199 lr: 0.0001 evaluation reward: 2.83
 episode: 1858 score: 6.0 memory length: 375729 epsilon: 0.4540546000105638
 steps: 393 lr: 0.0001 evaluation reward: 2.88
 episode: 1859 score: 3.0 memory length: 375974 epsilon:
 0.45356950001056073 steps: 245 lr: 0.0001 evaluation reward: 2.86
 episode: 1860 score: 1.0 memory length: 376126 epsilon:
 0.45326854001055883 steps: 152 lr: 0.0001 evaluation reward: 2.84
 episode: 1861 score: 1.0 memory length: 376277 epsilon:
 0.45296956001055694 steps: 151 lr: 0.0001 evaluation reward: 2.82
 episode: 1862 score: 2.0 memory length: 376456 epsilon: 0.4526151400105547
 steps: 179 lr: 0.0001 evaluation reward: 2.82
 episode: 1863 score: 3.0 memory length: 376701 epsilon: 0.4521300400105516
 steps: 245 lr: 0.0001 evaluation reward: 2.81
 episode: 1864 score: 2.0 memory length: 376898 epsilon:
 0.45173998001054916 steps: 197 lr: 0.0001 evaluation reward: 2.81
 episode: 1865 score: 2.0 memory length: 377095 epsilon: 0.4513499200105467
 steps: 197 lr: 0.0001 evaluation reward: 2.81
 episode: 1866 score: 3.0 memory length: 377320 epsilon:
 0.45090442001054387 steps: 225 lr: 0.0001 evaluation reward: 2.8
 episode: 1867 score: 2.0 memory length: 377518 epsilon: 0.4505123800105414
 steps: 198 lr: 0.0001 evaluation reward: 2.81
 episode: 1868 score: 2.0 memory length: 377715 epsilon: 0.4501223200105389
 steps: 197 lr: 0.0001 evaluation reward: 2.78
 episode: 1869 score: 2.0 memory length: 377896 epsilon:
 0.44976394001053666 steps: 181 lr: 0.0001 evaluation reward: 2.78
 episode: 1870 score: 4.0 memory length: 378171 epsilon: 0.4492194400105332
 steps: 275 lr: 0.0001 evaluation reward: 2.8
 episode: 1871 score: 2.0 memory length: 378369 epsilon:
 0.44882740001053073 steps: 198 lr: 0.0001 evaluation reward: 2.78
 episode: 1872 score: 3.0 memory length: 378594 epsilon: 0.4483819000105279
 steps: 225 lr: 0.0001 evaluation reward: 2.76
 episode: 1873 score: 2.0 memory length: 378791 epsilon:
 0.44799184001052544 steps: 197 lr: 0.0001 evaluation reward: 2.77
 episode: 1874 score: 3.0 memory length: 379019 epsilon: 0.4475404000105226
 steps: 228 lr: 0.0001 evaluation reward: 2.78
 episode: 1875 score: 2.0 memory length: 379234 epsilon: 0.4471147000105199
 steps: 215 lr: 0.0001 evaluation reward: 2.75

episode: 1876 score: 2.0 memory length: 379434 epsilon: 0.4467187000105174
steps: 200 lr: 0.0001 evaluation reward: 2.73
episode: 1877 score: 4.0 memory length: 379729 epsilon: 0.4461346000105137
steps: 295 lr: 0.0001 evaluation reward: 2.74
episode: 1878 score: 1.0 memory length: 379879 epsilon: 0.4458376000105118
steps: 150 lr: 0.0001 evaluation reward: 2.71
episode: 1879 score: 5.0 memory length: 380217 epsilon: 0.4451683600105076
steps: 338 lr: 0.0001 evaluation reward: 2.76
episode: 1880 score: 9.0 memory length: 380587 epsilon:
0.44443576001050294 steps: 370 lr: 0.0001 evaluation reward: 2.81
episode: 1881 score: 2.0 memory length: 380784 epsilon: 0.4440457000105005
steps: 197 lr: 0.0001 evaluation reward: 2.8
episode: 1882 score: 2.0 memory length: 380965 epsilon: 0.4436873200104982
steps: 181 lr: 0.0001 evaluation reward: 2.79
episode: 1883 score: 3.0 memory length: 381216 epsilon:
0.44319034001049507 steps: 251 lr: 0.0001 evaluation reward: 2.8
episode: 1884 score: 4.0 memory length: 381491 epsilon: 0.4426458400104916
steps: 275 lr: 0.0001 evaluation reward: 2.81
episode: 1885 score: 4.0 memory length: 381765 epsilon: 0.4421033200104882
steps: 274 lr: 0.0001 evaluation reward: 2.82
episode: 1886 score: 2.0 memory length: 381963 epsilon: 0.4417112800104857
steps: 198 lr: 0.0001 evaluation reward: 2.82
episode: 1887 score: 1.0 memory length: 382133 epsilon: 0.4413746800104836
steps: 170 lr: 0.0001 evaluation reward: 2.8
episode: 1888 score: 4.0 memory length: 382429 epsilon:
0.44078860001047987 steps: 296 lr: 0.0001 evaluation reward: 2.82
episode: 1889 score: 2.0 memory length: 382627 epsilon: 0.4403965600104774
steps: 198 lr: 0.0001 evaluation reward: 2.82
episode: 1890 score: 4.0 memory length: 382888 epsilon: 0.4398797800104741
steps: 261 lr: 0.0001 evaluation reward: 2.83
episode: 1891 score: 4.0 memory length: 383144 epsilon: 0.4393729000104709
steps: 256 lr: 0.0001 evaluation reward: 2.85
episode: 1892 score: 2.0 memory length: 383342 epsilon:
0.43898086001046843 steps: 198 lr: 0.0001 evaluation reward: 2.83
episode: 1893 score: 2.0 memory length: 383559 epsilon: 0.4385512000104657
steps: 217 lr: 0.0001 evaluation reward: 2.84
episode: 1894 score: 2.0 memory length: 383757 epsilon:
0.43815916001046323 steps: 198 lr: 0.0001 evaluation reward: 2.84
episode: 1895 score: 3.0 memory length: 383984 epsilon: 0.4377097000104604
steps: 227 lr: 0.0001 evaluation reward: 2.84
episode: 1896 score: 2.0 memory length: 384181 epsilon: 0.4373196400104579
steps: 197 lr: 0.0001 evaluation reward: 2.82
episode: 1897 score: 2.0 memory length: 384378 epsilon:
0.43692958001045545 steps: 197 lr: 0.0001 evaluation reward: 2.81
episode: 1898 score: 3.0 memory length: 384591 epsilon: 0.4365078400104528
steps: 213 lr: 0.0001 evaluation reward: 2.82
episode: 1899 score: 3.0 memory length: 384821 epsilon: 0.4360524400104499
steps: 230 lr: 0.0001 evaluation reward: 2.82

episode: 1900 score: 1.0 memory length: 384990 epsilon: 0.4357178200104478
 steps: 169 lr: 0.0001 evaluation reward: 2.8
 episode: 1901 score: 3.0 memory length: 385219 epsilon: 0.4352644000104449
 steps: 229 lr: 0.0001 evaluation reward: 2.82
 episode: 1902 score: 2.0 memory length: 385399 epsilon:
 0.43490800001044266 steps: 180 lr: 0.0001 evaluation reward: 2.82
 episode: 1903 score: 3.0 memory length: 385624 epsilon:
 0.43446250001043984 steps: 225 lr: 0.0001 evaluation reward: 2.82
 episode: 1904 score: 3.0 memory length: 385870 epsilon:
 0.43397542001043676 steps: 246 lr: 0.0001 evaluation reward: 2.82
 episode: 1905 score: 2.0 memory length: 386051 epsilon: 0.4336170400104345
 steps: 181 lr: 0.0001 evaluation reward: 2.82
 episode: 1906 score: 2.0 memory length: 386232 epsilon:
 0.43325866001043223 steps: 181 lr: 0.0001 evaluation reward: 2.8
 episode: 1907 score: 3.0 memory length: 386457 epsilon: 0.4328131600104294
 steps: 225 lr: 0.0001 evaluation reward: 2.79
 episode: 1908 score: 1.0 memory length: 386608 epsilon: 0.4325141800104275
 steps: 151 lr: 0.0001 evaluation reward: 2.74
 episode: 1909 score: 2.0 memory length: 386808 epsilon: 0.432118180010425
 steps: 200 lr: 0.0001 evaluation reward: 2.76
 episode: 1910 score: 2.0 memory length: 386989 epsilon:
 0.43175980001042275 steps: 181 lr: 0.0001 evaluation reward: 2.76
 episode: 1911 score: 5.0 memory length: 387268 epsilon:
 0.43120738001041925 steps: 279 lr: 0.0001 evaluation reward: 2.77
 episode: 1912 score: 5.0 memory length: 387573 epsilon:
 0.43060348001041543 steps: 305 lr: 0.0001 evaluation reward: 2.77
 episode: 1913 score: 2.0 memory length: 387754 epsilon:
 0.43024510001041316 steps: 181 lr: 0.0001 evaluation reward: 2.76
 episode: 1914 score: 3.0 memory length: 387979 epsilon:
 0.42979960001041034 steps: 225 lr: 0.0001 evaluation reward: 2.77
 episode: 1915 score: 2.0 memory length: 388177 epsilon:
 0.42940756001040786 steps: 198 lr: 0.0001 evaluation reward: 2.76
 episode: 1916 score: 7.0 memory length: 388522 epsilon:
 0.42872446001040354 steps: 345 lr: 0.0001 evaluation reward: 2.79
 episode: 1917 score: 3.0 memory length: 388748 epsilon: 0.4282769800104007
 steps: 226 lr: 0.0001 evaluation reward: 2.8
 episode: 1918 score: 2.0 memory length: 388945 epsilon:
 0.42788692001039824 steps: 197 lr: 0.0001 evaluation reward: 2.8
 episode: 1919 score: 3.0 memory length: 389191 epsilon:
 0.42739984001039516 steps: 246 lr: 0.0001 evaluation reward: 2.79
 episode: 1920 score: 5.0 memory length: 389492 epsilon: 0.4268038600103914
 steps: 301 lr: 0.0001 evaluation reward: 2.8
 episode: 1921 score: 2.0 memory length: 389689 epsilon: 0.4264138000103889
 steps: 197 lr: 0.0001 evaluation reward: 2.78
 episode: 1922 score: 2.0 memory length: 389871 epsilon:
 0.42605344001038664 steps: 182 lr: 0.0001 evaluation reward: 2.76
 episode: 1923 score: 5.0 memory length: 390198 epsilon:
 0.42540598001038255 steps: 327 lr: 0.0001 evaluation reward: 2.76

episode: 1924 score: 2.0 memory length: 390380 epsilon:
 0.42504562001038027 steps: 182 lr: 0.0001 evaluation reward: 2.74
 episode: 1925 score: 3.0 memory length: 390593 epsilon: 0.4246238800103776
 steps: 213 lr: 0.0001 evaluation reward: 2.74
 episode: 1926 score: 3.0 memory length: 390820 epsilon:
 0.42417442001037475 steps: 227 lr: 0.0001 evaluation reward: 2.74
 episode: 1927 score: 3.0 memory length: 391033 epsilon: 0.4237526800103721
 steps: 213 lr: 0.0001 evaluation reward: 2.75
 episode: 1928 score: 3.0 memory length: 391258 epsilon:
 0.42330718001036927 steps: 225 lr: 0.0001 evaluation reward: 2.75
 episode: 1929 score: 4.0 memory length: 391534 epsilon: 0.4227607000103658
 steps: 276 lr: 0.0001 evaluation reward: 2.77
 episode: 1930 score: 6.0 memory length: 391890 epsilon:
 0.42205582001036135 steps: 356 lr: 0.0001 evaluation reward: 2.81
 episode: 1931 score: 3.0 memory length: 392103 epsilon: 0.4216340800103587
 steps: 213 lr: 0.0001 evaluation reward: 2.81
 episode: 1932 score: 1.0 memory length: 392254 epsilon: 0.4213351000103568
 steps: 151 lr: 0.0001 evaluation reward: 2.78
 episode: 1933 score: 5.0 memory length: 392587 epsilon: 0.4206757600103526
 steps: 333 lr: 0.0001 evaluation reward: 2.8
 episode: 1934 score: 4.0 memory length: 392842 epsilon: 0.4201708600103494
 steps: 255 lr: 0.0001 evaluation reward: 2.81
 episode: 1935 score: 5.0 memory length: 393166 epsilon:
 0.41952934001034536 steps: 324 lr: 0.0001 evaluation reward: 2.83
 episode: 1936 score: 2.0 memory length: 393381 epsilon:
 0.41910364001034267 steps: 215 lr: 0.0001 evaluation reward: 2.83
 episode: 1937 score: 2.0 memory length: 393578 epsilon: 0.4187135800103402
 steps: 197 lr: 0.0001 evaluation reward: 2.82
 episode: 1938 score: 3.0 memory length: 393826 epsilon: 0.4182225400103371
 steps: 248 lr: 0.0001 evaluation reward: 2.82
 episode: 1939 score: 0.0 memory length: 393948 epsilon:
 0.41798098001033557 steps: 122 lr: 0.0001 evaluation reward: 2.79
 episode: 1940 score: 2.0 memory length: 394145 epsilon: 0.4175909200103331
 steps: 197 lr: 0.0001 evaluation reward: 2.78
 episode: 1941 score: 4.0 memory length: 394421 epsilon:
 0.41704444001032964 steps: 276 lr: 0.0001 evaluation reward: 2.8
 episode: 1942 score: 4.0 memory length: 394696 epsilon: 0.4164999400103262
 steps: 275 lr: 0.0001 evaluation reward: 2.81
 episode: 1943 score: 1.0 memory length: 394847 epsilon: 0.4162009600103243
 steps: 151 lr: 0.0001 evaluation reward: 2.8
 episode: 1944 score: 3.0 memory length: 395072 epsilon: 0.4157554600103215
 steps: 225 lr: 0.0001 evaluation reward: 2.81
 episode: 1945 score: 2.0 memory length: 395272 epsilon: 0.415359460010319
 steps: 200 lr: 0.0001 evaluation reward: 2.8
 episode: 1946 score: 3.0 memory length: 395498 epsilon:
 0.41491198001031615 steps: 226 lr: 0.0001 evaluation reward: 2.8
 episode: 1947 score: 3.0 memory length: 395724 epsilon: 0.4144645000103133
 steps: 226 lr: 0.0001 evaluation reward: 2.79

episode: 1948 score: 4.0 memory length: 395998 epsilon: 0.4139219800103099
 steps: 274 lr: 0.0001 evaluation reward: 2.83
 episode: 1949 score: 3.0 memory length: 396211 epsilon: 0.4135002400103072
 steps: 213 lr: 0.0001 evaluation reward: 2.86
 episode: 1950 score: 3.0 memory length: 396421 epsilon: 0.4130844400103046
 steps: 210 lr: 0.0001 evaluation reward: 2.86
 episode: 1951 score: 5.0 memory length: 396745 epsilon:
 0.41244292001030053 steps: 324 lr: 0.0001 evaluation reward: 2.88
 episode: 1952 score: 4.0 memory length: 397040 epsilon:
 0.41185882001029683 steps: 295 lr: 0.0001 evaluation reward: 2.89
 episode: 1953 score: 2.0 memory length: 397222 epsilon:
 0.41149846001029455 steps: 182 lr: 0.0001 evaluation reward: 2.88
 episode: 1954 score: 4.0 memory length: 397481 epsilon: 0.4109856400102913
 steps: 259 lr: 0.0001 evaluation reward: 2.91
 episode: 1955 score: 2.0 memory length: 397682 epsilon: 0.4105876600102888
 steps: 201 lr: 0.0001 evaluation reward: 2.9
 episode: 1956 score: 2.0 memory length: 397879 epsilon: 0.4101976000102863
 steps: 197 lr: 0.0001 evaluation reward: 2.89
 episode: 1957 score: 4.0 memory length: 398175 epsilon: 0.4096115200102826
 steps: 296 lr: 0.0001 evaluation reward: 2.91
 episode: 1958 score: 3.0 memory length: 398421 epsilon:
 0.40912444001027953 steps: 246 lr: 0.0001 evaluation reward: 2.88
 episode: 1959 score: 4.0 memory length: 398663 epsilon: 0.4086452800102765
 steps: 242 lr: 0.0001 evaluation reward: 2.89
 episode: 1960 score: 4.0 memory length: 398938 epsilon:
 0.40810078001027306 steps: 275 lr: 0.0001 evaluation reward: 2.92
 episode: 1961 score: 3.0 memory length: 399164 epsilon: 0.4076533000102702
 steps: 226 lr: 0.0001 evaluation reward: 2.94
 episode: 1962 score: 2.0 memory length: 399346 epsilon:
 0.40729294001026795 steps: 182 lr: 0.0001 evaluation reward: 2.94
 episode: 1963 score: 3.0 memory length: 399574 epsilon: 0.4068415000102651
 steps: 228 lr: 0.0001 evaluation reward: 2.94
 episode: 1964 score: 2.0 memory length: 399755 epsilon: 0.4064831200102628
 steps: 181 lr: 0.0001 evaluation reward: 2.94
 episode: 1965 score: 3.0 memory length: 399966 epsilon: 0.4060653400102602
 steps: 211 lr: 0.0001 evaluation reward: 2.95
 episode: 1966 score: 3.0 memory length: 400178 epsilon: 0.4056455800102575
 steps: 212 lr: 0.0001 evaluation reward: 2.95
 episode: 1967 score: 3.0 memory length: 400408 epsilon:
 0.40519018001025464 steps: 230 lr: 0.0001 evaluation reward: 2.96
 episode: 1968 score: 2.0 memory length: 400607 epsilon:
 0.40479616001025215 steps: 199 lr: 0.0001 evaluation reward: 2.96
 episode: 1969 score: 3.0 memory length: 400833 epsilon: 0.4043486800102493
 steps: 226 lr: 0.0001 evaluation reward: 2.97
 episode: 1970 score: 2.0 memory length: 401013 epsilon:
 0.40399228001024706 steps: 180 lr: 0.0001 evaluation reward: 2.95
 episode: 1971 score: 2.0 memory length: 401194 epsilon: 0.4036339000102448
 steps: 181 lr: 0.0001 evaluation reward: 2.95

episode: 1972 score: 7.0 memory length: 401541 epsilon:
 0.40294684001024045 steps: 347 lr: 0.0001 evaluation reward: 2.99
 episode: 1973 score: 4.0 memory length: 401807 epsilon: 0.4024201600102371
 steps: 266 lr: 0.0001 evaluation reward: 3.01
 episode: 1974 score: 3.0 memory length: 402020 epsilon:
 0.40199842001023445 steps: 213 lr: 0.0001 evaluation reward: 3.01
 episode: 1975 score: 3.0 memory length: 402232 epsilon: 0.4015786600102318
 steps: 212 lr: 0.0001 evaluation reward: 3.02
 episode: 1976 score: 6.0 memory length: 402550 epsilon: 0.4009490200102278
 steps: 318 lr: 0.0001 evaluation reward: 3.06
 episode: 1977 score: 2.0 memory length: 402732 epsilon:
 0.40058866001022553 steps: 182 lr: 0.0001 evaluation reward: 3.04
 episode: 1978 score: 3.0 memory length: 402943 epsilon: 0.4001708800102229
 steps: 211 lr: 0.0001 evaluation reward: 3.06
 episode: 1979 score: 4.0 memory length: 403203 epsilon: 0.3996560800102196
 steps: 260 lr: 0.0001 evaluation reward: 3.05
 episode: 1980 score: 3.0 memory length: 403416 epsilon:
 0.39923434001021696 steps: 213 lr: 0.0001 evaluation reward: 2.99
 episode: 1981 score: 3.0 memory length: 403629 epsilon: 0.3988126000102143
 steps: 213 lr: 0.0001 evaluation reward: 3.0
 episode: 1982 score: 2.0 memory length: 403826 epsilon: 0.3984225400102118
 steps: 197 lr: 0.0001 evaluation reward: 3.0
 episode: 1983 score: 2.0 memory length: 404006 epsilon:
 0.39806614001020957 steps: 180 lr: 0.0001 evaluation reward: 2.99
 episode: 1984 score: 3.0 memory length: 404232 epsilon:
 0.39761866001020674 steps: 226 lr: 0.0001 evaluation reward: 2.98
 episode: 1985 score: 2.0 memory length: 404431 epsilon:
 0.39722464001020424 steps: 199 lr: 0.0001 evaluation reward: 2.96
 episode: 1986 score: 4.0 memory length: 404700 epsilon: 0.3966920200102009
 steps: 269 lr: 0.0001 evaluation reward: 2.98
 episode: 1987 score: 7.0 memory length: 405126 epsilon:
 0.39584854001019554 steps: 426 lr: 0.0001 evaluation reward: 3.04
 episode: 1988 score: 0.0 memory length: 405248 epsilon: 0.395606980010194
 steps: 122 lr: 0.0001 evaluation reward: 3.0
 episode: 1989 score: 3.0 memory length: 405460 epsilon:
 0.39518722001019135 steps: 212 lr: 0.0001 evaluation reward: 3.01
 episode: 1990 score: 4.0 memory length: 405711 epsilon: 0.3946902400101882
 steps: 251 lr: 0.0001 evaluation reward: 3.01
 episode: 1991 score: 3.0 memory length: 405924 epsilon:
 0.39426850001018554 steps: 213 lr: 0.0001 evaluation reward: 3.0
 episode: 1992 score: 4.0 memory length: 406176 epsilon: 0.3937695400101824
 steps: 252 lr: 0.0001 evaluation reward: 3.02
 episode: 1993 score: 2.0 memory length: 406358 epsilon: 0.3934091800101801
 steps: 182 lr: 0.0001 evaluation reward: 3.02
 episode: 1994 score: 4.0 memory length: 406618 epsilon:
 0.39289438001017685 steps: 260 lr: 0.0001 evaluation reward: 3.04
 episode: 1995 score: 3.0 memory length: 406847 epsilon: 0.392440960010174
 steps: 229 lr: 0.0001 evaluation reward: 3.04

episode: 1996 score: 3.0 memory length: 407076 epsilon: 0.3919875400101711
 steps: 229 lr: 0.0001 evaluation reward: 3.05
 episode: 1997 score: 7.0 memory length: 407508 epsilon: 0.3911321800101657
 steps: 432 lr: 0.0001 evaluation reward: 3.1
 episode: 1998 score: 4.0 memory length: 407774 epsilon:
 0.39060550001016237 steps: 266 lr: 0.0001 evaluation reward: 3.11
 episode: 1999 score: 5.0 memory length: 408070 epsilon:
 0.39001942001015866 steps: 296 lr: 0.0001 evaluation reward: 3.13
 episode: 2000 score: 4.0 memory length: 408350 epsilon:
 0.38946502001015515 steps: 280 lr: 0.0001 evaluation reward: 3.16
 episode: 2001 score: 3.0 memory length: 408579 epsilon: 0.3890116000101523
 steps: 229 lr: 0.0001 evaluation reward: 3.16
 episode: 2002 score: 3.0 memory length: 408805 epsilon:
 0.38856412001014945 steps: 226 lr: 0.0001 evaluation reward: 3.17
 episode: 2003 score: 2.0 memory length: 409023 epsilon: 0.3881324800101467
 steps: 218 lr: 0.0001 evaluation reward: 3.16
 episode: 2004 score: 3.0 memory length: 409249 epsilon: 0.3876850000101439
 steps: 226 lr: 0.0001 evaluation reward: 3.16
 episode: 2005 score: 2.0 memory length: 409431 epsilon: 0.3873246400101416
 steps: 182 lr: 0.0001 evaluation reward: 3.16
 episode: 2006 score: 3.0 memory length: 409657 epsilon: 0.3868771600101388
 steps: 226 lr: 0.0001 evaluation reward: 3.17
 episode: 2007 score: 4.0 memory length: 409934 epsilon: 0.3863287000101353
 steps: 277 lr: 0.0001 evaluation reward: 3.18
 episode: 2008 score: 3.0 memory length: 410160 epsilon: 0.3858812200101325
 steps: 226 lr: 0.0001 evaluation reward: 3.2
 episode: 2009 score: 3.0 memory length: 410386 epsilon:
 0.38543374001012964 steps: 226 lr: 0.0001 evaluation reward: 3.21
 episode: 2010 score: 5.0 memory length: 410730 epsilon:
 0.38475262001012533 steps: 344 lr: 0.0001 evaluation reward: 3.24
 episode: 2011 score: 4.0 memory length: 410989 epsilon: 0.3842398000101221
 steps: 259 lr: 0.0001 evaluation reward: 3.23
 episode: 2012 score: 3.0 memory length: 411237 epsilon: 0.383748760010119
 steps: 248 lr: 0.0001 evaluation reward: 3.21
 episode: 2013 score: 2.0 memory length: 411419 epsilon: 0.3833884000101167
 steps: 182 lr: 0.0001 evaluation reward: 3.21
 episode: 2014 score: 3.0 memory length: 411631 epsilon:
 0.38296864001011405 steps: 212 lr: 0.0001 evaluation reward: 3.21
 episode: 2015 score: 2.0 memory length: 411812 epsilon: 0.3826102600101118
 steps: 181 lr: 0.0001 evaluation reward: 3.21
 episode: 2016 score: 5.0 memory length: 412157 epsilon:
 0.38192716001010746 steps: 345 lr: 0.0001 evaluation reward: 3.19
 episode: 2017 score: 3.0 memory length: 412404 epsilon:
 0.38143810001010436 steps: 247 lr: 0.0001 evaluation reward: 3.19
 episode: 2018 score: 3.0 memory length: 412630 epsilon:
 0.38099062001010153 steps: 226 lr: 0.0001 evaluation reward: 3.2
 episode: 2019 score: 4.0 memory length: 412892 epsilon:
 0.38047186001009825 steps: 262 lr: 0.0001 evaluation reward: 3.21

episode: 2020 score: 3.0 memory length: 413104 epsilon: 0.3800521000100956
 steps: 212 lr: 0.0001 evaluation reward: 3.19
 episode: 2021 score: 3.0 memory length: 413315 epsilon:
 0.37963432001009295 steps: 211 lr: 0.0001 evaluation reward: 3.2
 episode: 2022 score: 4.0 memory length: 413590 epsilon: 0.3790898200100895
 steps: 275 lr: 0.0001 evaluation reward: 3.22
 episode: 2023 score: 3.0 memory length: 413800 epsilon: 0.3786740200100869
 steps: 210 lr: 0.0001 evaluation reward: 3.2
 episode: 2024 score: 5.0 memory length: 414124 epsilon: 0.3780325000100828
 steps: 324 lr: 0.0001 evaluation reward: 3.23
 episode: 2025 score: 2.0 memory length: 414322 epsilon:
 0.37764046001008034 steps: 198 lr: 0.0001 evaluation reward: 3.22
 episode: 2026 score: 1.0 memory length: 414473 epsilon:
 0.37734148001007844 steps: 151 lr: 0.0001 evaluation reward: 3.2
 episode: 2027 score: 2.0 memory length: 414654 epsilon: 0.3769831000100762
 steps: 181 lr: 0.0001 evaluation reward: 3.19
 episode: 2028 score: 2.0 memory length: 414852 epsilon: 0.3765910600100737
 steps: 198 lr: 0.0001 evaluation reward: 3.18
 episode: 2029 score: 3.0 memory length: 415078 epsilon:
 0.37614358001007087 steps: 226 lr: 0.0001 evaluation reward: 3.17
 episode: 2030 score: 3.0 memory length: 415306 epsilon: 0.375692140010068
 steps: 228 lr: 0.0001 evaluation reward: 3.14
 episode: 2031 score: 6.0 memory length: 415640 epsilon: 0.3750308200100638
 steps: 334 lr: 0.0001 evaluation reward: 3.17
 episode: 2032 score: 3.0 memory length: 415866 epsilon: 0.374583340010061
 steps: 226 lr: 0.0001 evaluation reward: 3.19
 episode: 2033 score: 4.0 memory length: 416143 epsilon: 0.3740348800100575
 steps: 277 lr: 0.0001 evaluation reward: 3.18
 episode: 2034 score: 4.0 memory length: 416383 epsilon: 0.3735596800100545
 steps: 240 lr: 0.0001 evaluation reward: 3.18
 episode: 2035 score: 2.0 memory length: 416580 epsilon:
 0.37316962001005205 steps: 197 lr: 0.0001 evaluation reward: 3.15
 episode: 2036 score: 3.0 memory length: 416826 epsilon:
 0.37268254001004897 steps: 246 lr: 0.0001 evaluation reward: 3.16
 episode: 2037 score: 2.0 memory length: 417008 epsilon: 0.3723221800100467
 steps: 182 lr: 0.0001 evaluation reward: 3.16
 episode: 2038 score: 5.0 memory length: 417351 epsilon: 0.3716430400100424
 steps: 343 lr: 0.0001 evaluation reward: 3.18
 episode: 2039 score: 5.0 memory length: 417654 epsilon: 0.3710431000100386
 steps: 303 lr: 0.0001 evaluation reward: 3.23
 episode: 2040 score: 3.0 memory length: 417879 epsilon: 0.3705976000100358
 steps: 225 lr: 0.0001 evaluation reward: 3.24
 episode: 2041 score: 3.0 memory length: 418105 epsilon:
 0.37015012001003295 steps: 226 lr: 0.0001 evaluation reward: 3.23
 episode: 2042 score: 3.0 memory length: 418331 epsilon: 0.3697026400100301
 steps: 226 lr: 0.0001 evaluation reward: 3.22
 episode: 2043 score: 5.0 memory length: 418637 epsilon: 0.3690967600100263
 steps: 306 lr: 0.0001 evaluation reward: 3.26

episode: 2044 score: 4.0 memory length: 418896 epsilon:
 0.36858394001002304 steps: 259 lr: 0.0001 evaluation reward: 3.27
 episode: 2045 score: 3.0 memory length: 419142 epsilon:
 0.36809686001001996 steps: 246 lr: 0.0001 evaluation reward: 3.28
 episode: 2046 score: 3.0 memory length: 419353 epsilon: 0.3676790800100173
 steps: 211 lr: 0.0001 evaluation reward: 3.28
 episode: 2047 score: 5.0 memory length: 419676 epsilon:
 0.36703954001001327 steps: 323 lr: 0.0001 evaluation reward: 3.3
 episode: 2048 score: 7.0 memory length: 420072 epsilon: 0.3662554600100083
 steps: 396 lr: 0.0001 evaluation reward: 3.33
 episode: 2049 score: 5.0 memory length: 420398 epsilon: 0.3656099800100042
 steps: 326 lr: 0.0001 evaluation reward: 3.35
 episode: 2050 score: 5.0 memory length: 420723 epsilon:
 0.36496648001000015 steps: 325 lr: 0.0001 evaluation reward: 3.37
 episode: 2051 score: 4.0 memory length: 420984 epsilon: 0.3644497000099969
 steps: 261 lr: 0.0001 evaluation reward: 3.36
 episode: 2052 score: 2.0 memory length: 421166 epsilon: 0.3640893400099946
 steps: 182 lr: 0.0001 evaluation reward: 3.34
 episode: 2053 score: 2.0 memory length: 421346 epsilon:
 0.36373294000999234 steps: 180 lr: 0.0001 evaluation reward: 3.34
 episode: 2054 score: 3.0 memory length: 421555 epsilon: 0.3633191200099897
 steps: 209 lr: 0.0001 evaluation reward: 3.33
 episode: 2055 score: 4.0 memory length: 421855 epsilon:
 0.36272512000998597 steps: 300 lr: 0.0001 evaluation reward: 3.35
 episode: 2056 score: 2.0 memory length: 422036 epsilon: 0.3623667400099837
 steps: 181 lr: 0.0001 evaluation reward: 3.35
 episode: 2057 score: 3.0 memory length: 422283 epsilon: 0.3618776800099806
 steps: 247 lr: 0.0001 evaluation reward: 3.34
 episode: 2058 score: 3.0 memory length: 422533 epsilon: 0.3613826800099775
 steps: 250 lr: 0.0001 evaluation reward: 3.34
 episode: 2059 score: 3.0 memory length: 422745 epsilon: 0.3609629200099748
 steps: 212 lr: 0.0001 evaluation reward: 3.33
 episode: 2060 score: 4.0 memory length: 422999 epsilon:
 0.36046000000997164 steps: 254 lr: 0.0001 evaluation reward: 3.33
 episode: 2061 score: 5.0 memory length: 423306 epsilon: 0.3598521400099678
 steps: 307 lr: 0.0001 evaluation reward: 3.35
 episode: 2062 score: 3.0 memory length: 423554 epsilon: 0.3593611000099647
 steps: 248 lr: 0.0001 evaluation reward: 3.36
 episode: 2063 score: 3.0 memory length: 423780 epsilon:
 0.35891362000996185 steps: 226 lr: 0.0001 evaluation reward: 3.36
 episode: 2064 score: 3.0 memory length: 423990 epsilon: 0.3584978200099592
 steps: 210 lr: 0.0001 evaluation reward: 3.37
 episode: 2065 score: 3.0 memory length: 424220 epsilon:
 0.35804242000995634 steps: 230 lr: 0.0001 evaluation reward: 3.37
 episode: 2066 score: 2.0 memory length: 424401 epsilon: 0.3576840400099541
 steps: 181 lr: 0.0001 evaluation reward: 3.36
 episode: 2067 score: 3.0 memory length: 424628 epsilon:
 0.35723458000995123 steps: 227 lr: 0.0001 evaluation reward: 3.36

episode: 2068 score: 6.0 memory length: 424981 epsilon: 0.3565356400099468
 steps: 353 lr: 0.0001 evaluation reward: 3.4
 episode: 2069 score: 3.0 memory length: 425209 epsilon:
 0.35608420000994395 steps: 228 lr: 0.0001 evaluation reward: 3.4
 episode: 2070 score: 4.0 memory length: 425488 epsilon:
 0.35553178000994046 steps: 279 lr: 0.0001 evaluation reward: 3.42
 episode: 2071 score: 4.0 memory length: 425745 epsilon:
 0.35502292000993724 steps: 257 lr: 0.0001 evaluation reward: 3.44
 episode: 2072 score: 4.0 memory length: 426022 epsilon:
 0.35447446000993377 steps: 277 lr: 0.0001 evaluation reward: 3.41
 episode: 2073 score: 3.0 memory length: 426249 epsilon: 0.3540250000099309
 steps: 227 lr: 0.0001 evaluation reward: 3.4
 episode: 2074 score: 3.0 memory length: 426478 epsilon:
 0.35357158000992805 steps: 229 lr: 0.0001 evaluation reward: 3.4
 episode: 2075 score: 4.0 memory length: 426772 epsilon:
 0.35298946000992437 steps: 294 lr: 0.0001 evaluation reward: 3.41
 episode: 2076 score: 2.0 memory length: 426952 epsilon: 0.3526330600099221
 steps: 180 lr: 0.0001 evaluation reward: 3.37
 episode: 2077 score: 4.0 memory length: 427211 epsilon: 0.3521202400099189
 steps: 259 lr: 0.0001 evaluation reward: 3.39
 episode: 2078 score: 6.0 memory length: 427565 epsilon:
 0.35141932000991444 steps: 354 lr: 0.0001 evaluation reward: 3.42
 episode: 2079 score: 3.0 memory length: 427777 epsilon: 0.3509995600099118
 steps: 212 lr: 0.0001 evaluation reward: 3.41
 episode: 2080 score: 3.0 memory length: 428006 epsilon: 0.3505461400099089
 steps: 229 lr: 0.0001 evaluation reward: 3.41
 episode: 2081 score: 3.0 memory length: 428217 epsilon:
 0.35012836000990627 steps: 211 lr: 0.0001 evaluation reward: 3.41
 episode: 2082 score: 3.0 memory length: 428442 epsilon:
 0.34968286000990345 steps: 225 lr: 0.0001 evaluation reward: 3.42
 episode: 2083 score: 3.0 memory length: 428688 epsilon:
 0.34919578000990037 steps: 246 lr: 0.0001 evaluation reward: 3.43
 episode: 2084 score: 4.0 memory length: 428963 epsilon: 0.3486512800098969
 steps: 275 lr: 0.0001 evaluation reward: 3.44
 episode: 2085 score: 2.0 memory length: 429160 epsilon:
 0.34826122000989446 steps: 197 lr: 0.0001 evaluation reward: 3.44
 episode: 2086 score: 4.0 memory length: 429438 epsilon: 0.347710780009891
 steps: 278 lr: 0.0001 evaluation reward: 3.44
 episode: 2087 score: 4.0 memory length: 429700 epsilon: 0.3471920200098877
 steps: 262 lr: 0.0001 evaluation reward: 3.41
 episode: 2088 score: 3.0 memory length: 429946 epsilon: 0.3467049400098846
 steps: 246 lr: 0.0001 evaluation reward: 3.44
 episode: 2089 score: 3.0 memory length: 430159 epsilon:
 0.34628320000988194 steps: 213 lr: 0.0001 evaluation reward: 3.44
 episode: 2090 score: 5.0 memory length: 430445 epsilon:
 0.34571692000987836 steps: 286 lr: 0.0001 evaluation reward: 3.45
 episode: 2091 score: 3.0 memory length: 430675 epsilon: 0.3452615200098755
 steps: 230 lr: 0.0001 evaluation reward: 3.45

episode: 2092 score: 4.0 memory length: 430950 epsilon:
0.34471702000987203 steps: 275 lr: 0.0001 evaluation reward: 3.45
episode: 2093 score: 3.0 memory length: 431161 epsilon: 0.3442992400098694
steps: 211 lr: 0.0001 evaluation reward: 3.46
episode: 2094 score: 7.0 memory length: 431584 epsilon: 0.3434617000098641
steps: 423 lr: 0.0001 evaluation reward: 3.49
episode: 2095 score: 2.0 memory length: 431766 epsilon: 0.3431013400098618
steps: 182 lr: 0.0001 evaluation reward: 3.48
episode: 2096 score: 3.0 memory length: 432013 epsilon: 0.3426122800098587
steps: 247 lr: 0.0001 evaluation reward: 3.48
episode: 2097 score: 5.0 memory length: 432325 epsilon: 0.3419945200098548
steps: 312 lr: 0.0001 evaluation reward: 3.46
episode: 2098 score: 4.0 memory length: 432600 epsilon:
0.34145002000985136 steps: 275 lr: 0.0001 evaluation reward: 3.46
episode: 2099 score: 5.0 memory length: 432925 epsilon: 0.3408065200098473
steps: 325 lr: 0.0001 evaluation reward: 3.46
episode: 2100 score: 5.0 memory length: 433270 epsilon:
0.34012342000984297 steps: 345 lr: 0.0001 evaluation reward: 3.47
episode: 2101 score: 5.0 memory length: 433579 epsilon: 0.3395116000098391
steps: 309 lr: 0.0001 evaluation reward: 3.49
episode: 2102 score: 3.0 memory length: 433827 epsilon: 0.339020560009836
steps: 248 lr: 0.0001 evaluation reward: 3.49
episode: 2103 score: 1.0 memory length: 433978 epsilon: 0.3387215800098341
steps: 151 lr: 0.0001 evaluation reward: 3.48
episode: 2104 score: 3.0 memory length: 434191 epsilon:
0.33829984000983143 steps: 213 lr: 0.0001 evaluation reward: 3.48
episode: 2105 score: 2.0 memory length: 434372 epsilon:
0.33794146000982916 steps: 181 lr: 0.0001 evaluation reward: 3.48
episode: 2106 score: 3.0 memory length: 434598 epsilon:
0.33749398000982633 steps: 226 lr: 0.0001 evaluation reward: 3.48
episode: 2107 score: 2.0 memory length: 434797 epsilon:
0.33709996000982384 steps: 199 lr: 0.0001 evaluation reward: 3.46
episode: 2108 score: 4.0 memory length: 435071 epsilon: 0.3365574400098204
steps: 274 lr: 0.0001 evaluation reward: 3.47
episode: 2109 score: 3.0 memory length: 435297 epsilon: 0.3361099600098176
steps: 226 lr: 0.0001 evaluation reward: 3.47
episode: 2110 score: 4.0 memory length: 435557 epsilon: 0.3355951600098143
steps: 260 lr: 0.0001 evaluation reward: 3.46
episode: 2111 score: 6.0 memory length: 435884 epsilon: 0.3349477000098102
steps: 327 lr: 0.0001 evaluation reward: 3.48
episode: 2112 score: 7.0 memory length: 436272 epsilon:
0.33417946000980536 steps: 388 lr: 0.0001 evaluation reward: 3.52
episode: 2113 score: 5.0 memory length: 436582 epsilon: 0.3335656600098015
steps: 310 lr: 0.0001 evaluation reward: 3.55
episode: 2114 score: 3.0 memory length: 436792 epsilon:
0.33314986000979885 steps: 210 lr: 0.0001 evaluation reward: 3.55
episode: 2115 score: 4.0 memory length: 437087 epsilon:
0.33256576000979515 steps: 295 lr: 0.0001 evaluation reward: 3.57

episode: 2116 score: 2.0 memory length: 437269 epsilon: 0.3322054000097929
 steps: 182 lr: 0.0001 evaluation reward: 3.54
 episode: 2117 score: 3.0 memory length: 437495 epsilon:
 0.33175792000979004 steps: 226 lr: 0.0001 evaluation reward: 3.54
 episode: 2118 score: 3.0 memory length: 437720 epsilon: 0.3313124200097872
 steps: 225 lr: 0.0001 evaluation reward: 3.54
 episode: 2119 score: 2.0 memory length: 437901 epsilon:
 0.33095404000978496 steps: 181 lr: 0.0001 evaluation reward: 3.52
 episode: 2120 score: 2.0 memory length: 438082 epsilon: 0.3305956600097827
 steps: 181 lr: 0.0001 evaluation reward: 3.51
 episode: 2121 score: 2.0 memory length: 438264 epsilon: 0.3302353000097804
 steps: 182 lr: 0.0001 evaluation reward: 3.5
 episode: 2122 score: 3.0 memory length: 438495 epsilon: 0.3297779200097775
 steps: 231 lr: 0.0001 evaluation reward: 3.49
 episode: 2123 score: 4.0 memory length: 438772 epsilon:
 0.32922946000977404 steps: 277 lr: 0.0001 evaluation reward: 3.5
 episode: 2124 score: 3.0 memory length: 438985 epsilon: 0.3288077200097714
 steps: 213 lr: 0.0001 evaluation reward: 3.48
 episode: 2125 score: 5.0 memory length: 439281 epsilon:
 0.32822164000976767 steps: 296 lr: 0.0001 evaluation reward: 3.51
 episode: 2126 score: 3.0 memory length: 439509 epsilon: 0.3277702000097648
 steps: 228 lr: 0.0001 evaluation reward: 3.53
 episode: 2127 score: 2.0 memory length: 439691 epsilon:
 0.32740984000976253 steps: 182 lr: 0.0001 evaluation reward: 3.53
 episode: 2128 score: 3.0 memory length: 439939 epsilon: 0.3269188000097594
 steps: 248 lr: 0.0001 evaluation reward: 3.54
 episode: 2129 score: 4.0 memory length: 440201 epsilon:
 0.32640004000975614 steps: 262 lr: 0.0001 evaluation reward: 3.55
 episode: 2130 score: 2.0 memory length: 440381 epsilon: 0.3260436400097539
 steps: 180 lr: 0.0001 evaluation reward: 3.54
 episode: 2131 score: 2.0 memory length: 440581 epsilon: 0.3256476400097514
 steps: 200 lr: 0.0001 evaluation reward: 3.5
 episode: 2132 score: 4.0 memory length: 440847 epsilon:
 0.32512096000974805 steps: 266 lr: 0.0001 evaluation reward: 3.51
 episode: 2133 score: 6.0 memory length: 441220 epsilon: 0.3243824200097434
 steps: 373 lr: 0.0001 evaluation reward: 3.53
 episode: 2134 score: 7.0 memory length: 441595 epsilon: 0.3236399200097387
 steps: 375 lr: 0.0001 evaluation reward: 3.56
 episode: 2135 score: 2.0 memory length: 441795 epsilon: 0.3232439200097362
 steps: 200 lr: 0.0001 evaluation reward: 3.56
 episode: 2136 score: 3.0 memory length: 442006 epsilon:
 0.32282614000973353 steps: 211 lr: 0.0001 evaluation reward: 3.56
 episode: 2137 score: 3.0 memory length: 442219 epsilon:
 0.32240440000973086 steps: 213 lr: 0.0001 evaluation reward: 3.57
 episode: 2138 score: 3.0 memory length: 442447 epsilon: 0.321952960009728
 steps: 228 lr: 0.0001 evaluation reward: 3.55
 episode: 2139 score: 3.0 memory length: 442660 epsilon:
 0.32153122000972534 steps: 213 lr: 0.0001 evaluation reward: 3.53

episode: 2140 score: 3.0 memory length: 442906 epsilon:
 0.32104414000972226 steps: 246 lr: 0.0001 evaluation reward: 3.53
 episode: 2141 score: 3.0 memory length: 443136 epsilon: 0.3205887400097194
 steps: 230 lr: 0.0001 evaluation reward: 3.53
 episode: 2142 score: 3.0 memory length: 443383 epsilon: 0.3200996800097163
 steps: 247 lr: 0.0001 evaluation reward: 3.53
 episode: 2143 score: 5.0 memory length: 443664 epsilon:
 0.31954330000971276 steps: 281 lr: 0.0001 evaluation reward: 3.53
 episode: 2144 score: 4.0 memory length: 443938 epsilon:
 0.31900078000970933 steps: 274 lr: 0.0001 evaluation reward: 3.53
 episode: 2145 score: 4.0 memory length: 444200 epsilon:
 0.31848202000970605 steps: 262 lr: 0.0001 evaluation reward: 3.54
 episode: 2146 score: 4.0 memory length: 444474 epsilon: 0.3179395000097026
 steps: 274 lr: 0.0001 evaluation reward: 3.55
 episode: 2147 score: 2.0 memory length: 444655 epsilon:
 0.31758112000970035 steps: 181 lr: 0.0001 evaluation reward: 3.52
 episode: 2148 score: 3.0 memory length: 444901 epsilon:
 0.31709404000969726 steps: 246 lr: 0.0001 evaluation reward: 3.48
 episode: 2149 score: 3.0 memory length: 445111 epsilon:
 0.31667824000969463 steps: 210 lr: 0.0001 evaluation reward: 3.46
 episode: 2150 score: 2.0 memory length: 445290 epsilon: 0.3163238200096924
 steps: 179 lr: 0.0001 evaluation reward: 3.43
 episode: 2151 score: 4.0 memory length: 445585 epsilon: 0.3157397200096887
 steps: 295 lr: 0.0001 evaluation reward: 3.43
 episode: 2152 score: 4.0 memory length: 445897 epsilon: 0.3151219600096848
 steps: 312 lr: 0.0001 evaluation reward: 3.45
 episode: 2153 score: 3.0 memory length: 446127 epsilon: 0.3146665600096819
 steps: 230 lr: 0.0001 evaluation reward: 3.46
 episode: 2154 score: 3.0 memory length: 446356 epsilon:
 0.31421314000967904 steps: 229 lr: 0.0001 evaluation reward: 3.46
 episode: 2155 score: 4.0 memory length: 446631 epsilon: 0.3136686400096756
 steps: 275 lr: 0.0001 evaluation reward: 3.46
 episode: 2156 score: 2.0 memory length: 446812 epsilon: 0.3133102600096733
 steps: 181 lr: 0.0001 evaluation reward: 3.46
 episode: 2157 score: 3.0 memory length: 447022 epsilon: 0.3128944600096707
 steps: 210 lr: 0.0001 evaluation reward: 3.46
 episode: 2158 score: 3.0 memory length: 447247 epsilon: 0.3124489600096679
 steps: 225 lr: 0.0001 evaluation reward: 3.46
 episode: 2159 score: 2.0 memory length: 447428 epsilon: 0.3120905800096656
 steps: 181 lr: 0.0001 evaluation reward: 3.45
 episode: 2160 score: 7.0 memory length: 447849 epsilon:
 0.31125700000966033 steps: 421 lr: 0.0001 evaluation reward: 3.48
 episode: 2161 score: 9.0 memory length: 448328 epsilon:
 0.31030858000965433 steps: 479 lr: 0.0001 evaluation reward: 3.52
 episode: 2162 score: 2.0 memory length: 448509 epsilon:
 0.30995020000965207 steps: 181 lr: 0.0001 evaluation reward: 3.51
 episode: 2163 score: 4.0 memory length: 448784 epsilon: 0.3094057000096486
 steps: 275 lr: 0.0001 evaluation reward: 3.52

episode: 2164 score: 3.0 memory length: 449010 epsilon: 0.3089582200096458
 steps: 226 lr: 0.0001 evaluation reward: 3.52
 episode: 2165 score: 4.0 memory length: 449268 epsilon:
 0.30844738000964256 steps: 258 lr: 0.0001 evaluation reward: 3.53
 episode: 2166 score: 4.0 memory length: 449527 epsilon: 0.3079345600096393
 steps: 259 lr: 0.0001 evaluation reward: 3.55
 episode: 2167 score: 3.0 memory length: 449738 epsilon:
 0.30751678000963667 steps: 211 lr: 0.0001 evaluation reward: 3.55
 episode: 2168 score: 4.0 memory length: 450011 epsilon:
 0.30697624000963325 steps: 273 lr: 0.0001 evaluation reward: 3.53
 episode: 2169 score: 4.0 memory length: 450288 epsilon: 0.3064277800096298
 steps: 277 lr: 0.0001 evaluation reward: 3.54
 episode: 2170 score: 5.0 memory length: 450612 epsilon: 0.3057862600096257
 steps: 324 lr: 0.0001 evaluation reward: 3.55
 episode: 2171 score: 2.0 memory length: 450810 epsilon:
 0.30539422000962324 steps: 198 lr: 0.0001 evaluation reward: 3.53
 episode: 2172 score: 4.0 memory length: 451085 epsilon: 0.3048497200096198
 steps: 275 lr: 0.0001 evaluation reward: 3.53
 episode: 2173 score: 3.0 memory length: 451297 epsilon:
 0.30442996000961714 steps: 212 lr: 0.0001 evaluation reward: 3.53
 episode: 2174 score: 3.0 memory length: 451543 epsilon:
 0.30394288000961406 steps: 246 lr: 0.0001 evaluation reward: 3.53
 episode: 2175 score: 4.0 memory length: 451820 epsilon: 0.3033944200096106
 steps: 277 lr: 0.0001 evaluation reward: 3.53
 episode: 2176 score: 3.0 memory length: 452048 epsilon:
 0.30294298000960773 steps: 228 lr: 0.0001 evaluation reward: 3.54
 episode: 2177 score: 5.0 memory length: 452357 epsilon:
 0.30233116000960386 steps: 309 lr: 0.0001 evaluation reward: 3.55
 episode: 2178 score: 2.0 memory length: 452539 epsilon: 0.3019708000096016
 steps: 182 lr: 0.0001 evaluation reward: 3.51
 episode: 2179 score: 2.0 memory length: 452719 epsilon: 0.3016144000095993
 steps: 180 lr: 0.0001 evaluation reward: 3.5
 episode: 2180 score: 6.0 memory length: 453063 epsilon: 0.300933280009595
 steps: 344 lr: 0.0001 evaluation reward: 3.53
 episode: 2181 score: 2.0 memory length: 453244 epsilon:
 0.30057490000959275 steps: 181 lr: 0.0001 evaluation reward: 3.52
 episode: 2182 score: 3.0 memory length: 453473 epsilon: 0.3001214800095899
 steps: 229 lr: 0.0001 evaluation reward: 3.52
 episode: 2183 score: 4.0 memory length: 453735 epsilon: 0.2996027200095866
 steps: 262 lr: 0.0001 evaluation reward: 3.53
 episode: 2184 score: 2.0 memory length: 453935 epsilon: 0.2992067200095841
 steps: 200 lr: 0.0001 evaluation reward: 3.51
 episode: 2185 score: 3.0 memory length: 454148 epsilon: 0.2987849800095814
 steps: 213 lr: 0.0001 evaluation reward: 3.52
 episode: 2186 score: 3.0 memory length: 454360 epsilon:
 0.29836522000957877 steps: 212 lr: 0.0001 evaluation reward: 3.51
 episode: 2187 score: 4.0 memory length: 454638 epsilon: 0.2978147800095753
 steps: 278 lr: 0.0001 evaluation reward: 3.51

episode: 2188 score: 2.0 memory length: 454820 epsilon: 0.297454420009573
 steps: 182 lr: 0.0001 evaluation reward: 3.5
 episode: 2189 score: 5.0 memory length: 455164 epsilon: 0.2967733000095687
 steps: 344 lr: 0.0001 evaluation reward: 3.52
 episode: 2190 score: 4.0 memory length: 455422 epsilon:
 0.29626246000956546 steps: 258 lr: 0.0001 evaluation reward: 3.51
 episode: 2191 score: 4.0 memory length: 455719 epsilon:
 0.29567440000956174 steps: 297 lr: 0.0001 evaluation reward: 3.52
 episode: 2192 score: 3.0 memory length: 455965 epsilon:
 0.29518732000955866 steps: 246 lr: 0.0001 evaluation reward: 3.51
 episode: 2193 score: 5.0 memory length: 456291 epsilon: 0.2945418400095546
 steps: 326 lr: 0.0001 evaluation reward: 3.53
 episode: 2194 score: 10.0 memory length: 456666 epsilon:
 0.2937993400095499 steps: 375 lr: 0.0001 evaluation reward: 3.56
 episode: 2195 score: 4.0 memory length: 456928 epsilon: 0.2932805800095466
 steps: 262 lr: 0.0001 evaluation reward: 3.58
 episode: 2196 score: 5.0 memory length: 457254 epsilon: 0.2926351000095425
 steps: 326 lr: 0.0001 evaluation reward: 3.6
 episode: 2197 score: 5.0 memory length: 457535 epsilon: 0.292078720009539
 steps: 281 lr: 0.0001 evaluation reward: 3.6
 episode: 2198 score: 2.0 memory length: 457717 epsilon: 0.2917183600095367
 steps: 182 lr: 0.0001 evaluation reward: 3.58
 episode: 2199 score: 2.0 memory length: 457937 epsilon:
 0.29128276000953396 steps: 220 lr: 0.0001 evaluation reward: 3.55
 episode: 2200 score: 4.0 memory length: 458197 epsilon: 0.2907679600095307
 steps: 260 lr: 0.0001 evaluation reward: 3.54
 episode: 2201 score: 3.0 memory length: 458424 epsilon:
 0.29031850000952786 steps: 227 lr: 0.0001 evaluation reward: 3.52
 episode: 2202 score: 4.0 memory length: 458704 epsilon:
 0.28976410000952435 steps: 280 lr: 0.0001 evaluation reward: 3.53
 episode: 2203 score: 2.0 memory length: 458884 epsilon: 0.2894077000095221
 steps: 180 lr: 0.0001 evaluation reward: 3.54
 episode: 2204 score: 9.0 memory length: 459312 epsilon:
 0.28856026000951673 steps: 428 lr: 0.0001 evaluation reward: 3.6
 episode: 2205 score: 2.0 memory length: 459509 epsilon:
 0.28817020000951427 steps: 197 lr: 0.0001 evaluation reward: 3.6
 episode: 2206 score: 6.0 memory length: 459834 epsilon: 0.2875267000095102
 steps: 325 lr: 0.0001 evaluation reward: 3.63
 episode: 2207 score: 6.0 memory length: 460230 epsilon:
 0.28674262000950523 steps: 396 lr: 0.0001 evaluation reward: 3.67
 episode: 2208 score: 4.0 memory length: 460525 epsilon:
 0.28615852000950154 steps: 295 lr: 0.0001 evaluation reward: 3.67
 episode: 2209 score: 2.0 memory length: 460722 epsilon:
 0.28576846000949907 steps: 197 lr: 0.0001 evaluation reward: 3.66
 episode: 2210 score: 4.0 memory length: 460977 epsilon: 0.2852635600094959
 steps: 255 lr: 0.0001 evaluation reward: 3.66
 episode: 2211 score: 3.0 memory length: 461207 epsilon: 0.284808160009493
 steps: 230 lr: 0.0001 evaluation reward: 3.63

episode: 2212 score: 2.0 memory length: 461389 epsilon: 0.2844478000094907
 steps: 182 lr: 0.0001 evaluation reward: 3.58
 episode: 2213 score: 2.0 memory length: 461570 epsilon:
 0.28408942000948845 steps: 181 lr: 0.0001 evaluation reward: 3.55
 episode: 2214 score: 7.0 memory length: 461956 epsilon: 0.2833251400094836
 steps: 386 lr: 0.0001 evaluation reward: 3.59
 episode: 2215 score: 3.0 memory length: 462168 epsilon:
 0.28290538000948096 steps: 212 lr: 0.0001 evaluation reward: 3.58
 episode: 2216 score: 5.0 memory length: 462493 epsilon: 0.2822618800094769
 steps: 325 lr: 0.0001 evaluation reward: 3.61
 episode: 2217 score: 2.0 memory length: 462674 epsilon: 0.2819035000094746
 steps: 181 lr: 0.0001 evaluation reward: 3.6
 episode: 2218 score: 2.0 memory length: 462856 epsilon:
 0.28154314000947234 steps: 182 lr: 0.0001 evaluation reward: 3.59
 episode: 2219 score: 3.0 memory length: 463063 epsilon:
 0.28113328000946974 steps: 207 lr: 0.0001 evaluation reward: 3.6
 episode: 2220 score: 2.0 memory length: 463243 epsilon: 0.2807768800094675
 steps: 180 lr: 0.0001 evaluation reward: 3.6
 episode: 2221 score: 3.0 memory length: 463471 epsilon:
 0.28032544000946463 steps: 228 lr: 0.0001 evaluation reward: 3.61
 episode: 2222 score: 4.0 memory length: 463730 epsilon: 0.2798126200094614
 steps: 259 lr: 0.0001 evaluation reward: 3.62
 episode: 2223 score: 4.0 memory length: 464005 epsilon:
 0.27926812000945794 steps: 275 lr: 0.0001 evaluation reward: 3.62
 episode: 2224 score: 5.0 memory length: 464311 epsilon: 0.2786622400094541
 steps: 306 lr: 0.0001 evaluation reward: 3.64
 episode: 2225 score: 5.0 memory length: 464606 epsilon: 0.2780781400094504
 steps: 295 lr: 0.0001 evaluation reward: 3.64
 episode: 2226 score: 2.0 memory length: 464787 epsilon:
 0.27771976000944815 steps: 181 lr: 0.0001 evaluation reward: 3.63
 episode: 2227 score: 5.0 memory length: 465111 epsilon: 0.2770782400094441
 steps: 324 lr: 0.0001 evaluation reward: 3.66
 episode: 2228 score: 5.0 memory length: 465419 epsilon:
 0.27646840000944023 steps: 308 lr: 0.0001 evaluation reward: 3.68
 episode: 2229 score: 3.0 memory length: 465629 epsilon: 0.2760526000094376
 steps: 210 lr: 0.0001 evaluation reward: 3.67
 episode: 2230 score: 2.0 memory length: 465811 epsilon: 0.2756922400094353
 steps: 182 lr: 0.0001 evaluation reward: 3.67
 episode: 2231 score: 3.0 memory length: 466024 epsilon:
 0.27527050000943265 steps: 213 lr: 0.0001 evaluation reward: 3.68
 episode: 2232 score: 4.0 memory length: 466281 epsilon:
 0.27476164000942943 steps: 257 lr: 0.0001 evaluation reward: 3.68
 episode: 2233 score: 2.0 memory length: 466462 epsilon:
 0.27440326000942716 steps: 181 lr: 0.0001 evaluation reward: 3.64
 episode: 2234 score: 5.0 memory length: 466771 epsilon: 0.2737914400094233
 steps: 309 lr: 0.0001 evaluation reward: 3.62
 episode: 2235 score: 4.0 memory length: 467067 epsilon: 0.2732053600094196
 steps: 296 lr: 0.0001 evaluation reward: 3.64

episode: 2236 score: 4.0 memory length: 467347 epsilon: 0.2726509600094161
 steps: 280 lr: 0.0001 evaluation reward: 3.65
 episode: 2237 score: 4.0 memory length: 467585 epsilon: 0.2721797200094131
 steps: 238 lr: 0.0001 evaluation reward: 3.66
 episode: 2238 score: 4.0 memory length: 467824 epsilon: 0.2717065000094101
 steps: 239 lr: 0.0001 evaluation reward: 3.67
 episode: 2239 score: 4.0 memory length: 468098 epsilon:
 0.27116398000940667 steps: 274 lr: 0.0001 evaluation reward: 3.68
 episode: 2240 score: 1.0 memory length: 468248 epsilon: 0.2708669800094048
 steps: 150 lr: 0.0001 evaluation reward: 3.66
 episode: 2241 score: 3.0 memory length: 468458 epsilon:
 0.27045118000940216 steps: 210 lr: 0.0001 evaluation reward: 3.66
 episode: 2242 score: 3.0 memory length: 468684 epsilon:
 0.27000370000939933 steps: 226 lr: 0.0001 evaluation reward: 3.66
 episode: 2243 score: 5.0 memory length: 469010 epsilon:
 0.26935822000939524 steps: 326 lr: 0.0001 evaluation reward: 3.66
 episode: 2244 score: 3.0 memory length: 469240 epsilon:
 0.26890282000939236 steps: 230 lr: 0.0001 evaluation reward: 3.65
 episode: 2245 score: 5.0 memory length: 469520 epsilon:
 0.26834842000938885 steps: 280 lr: 0.0001 evaluation reward: 3.66
 episode: 2246 score: 3.0 memory length: 469732 epsilon: 0.2679286600093862
 steps: 212 lr: 0.0001 evaluation reward: 3.65
 episode: 2247 score: 4.0 memory length: 469976 epsilon:
 0.26744554000938314 steps: 244 lr: 0.0001 evaluation reward: 3.67
 episode: 2248 score: 2.0 memory length: 470193 epsilon: 0.2670158800093804
 steps: 217 lr: 0.0001 evaluation reward: 3.66
 episode: 2249 score: 3.0 memory length: 470406 epsilon:
 0.26659414000937776 steps: 213 lr: 0.0001 evaluation reward: 3.66
 episode: 2250 score: 4.0 memory length: 470647 epsilon:
 0.26611696000937474 steps: 241 lr: 0.0001 evaluation reward: 3.68
 episode: 2251 score: 3.0 memory length: 470876 epsilon:
 0.26566354000937187 steps: 229 lr: 0.0001 evaluation reward: 3.67
 episode: 2252 score: 3.0 memory length: 471102 epsilon:
 0.26521606000936904 steps: 226 lr: 0.0001 evaluation reward: 3.66
 episode: 2253 score: 4.0 memory length: 471382 epsilon:
 0.26466166000936553 steps: 280 lr: 0.0001 evaluation reward: 3.67
 episode: 2254 score: 3.0 memory length: 471612 epsilon:
 0.26420626000936265 steps: 230 lr: 0.0001 evaluation reward: 3.67
 episode: 2255 score: 5.0 memory length: 471936 epsilon: 0.2635647400093586
 steps: 324 lr: 0.0001 evaluation reward: 3.68
 episode: 2256 score: 3.0 memory length: 472148 epsilon:
 0.26314498000935593 steps: 212 lr: 0.0001 evaluation reward: 3.69
 episode: 2257 score: 5.0 memory length: 472473 epsilon:
 0.26250148000935186 steps: 325 lr: 0.0001 evaluation reward: 3.71
 episode: 2258 score: 3.0 memory length: 472685 epsilon: 0.2620817200093492
 steps: 212 lr: 0.0001 evaluation reward: 3.71
 episode: 2259 score: 3.0 memory length: 472898 epsilon:
 0.26165998000934654 steps: 213 lr: 0.0001 evaluation reward: 3.72

episode: 2260 score: 5.0 memory length: 473226 epsilon:
 0.26101054000934243 steps: 328 lr: 0.0001 evaluation reward: 3.7
 episode: 2261 score: 3.0 memory length: 473472 epsilon:
 0.26052346000933935 steps: 246 lr: 0.0001 evaluation reward: 3.64
 episode: 2262 score: 4.0 memory length: 473731 epsilon: 0.2600106400093361
 steps: 259 lr: 0.0001 evaluation reward: 3.66
 episode: 2263 score: 2.0 memory length: 473912 epsilon:
 0.25965226000933384 steps: 181 lr: 0.0001 evaluation reward: 3.64
 episode: 2264 score: 3.0 memory length: 474125 epsilon:
 0.25923052000933117 steps: 213 lr: 0.0001 evaluation reward: 3.64
 episode: 2265 score: 6.0 memory length: 474482 epsilon: 0.2585236600093267
 steps: 357 lr: 0.0001 evaluation reward: 3.66
 episode: 2266 score: 3.0 memory length: 474711 epsilon: 0.2580702400093238
 steps: 229 lr: 0.0001 evaluation reward: 3.65
 episode: 2267 score: 3.0 memory length: 474943 epsilon: 0.2576108800093209
 steps: 232 lr: 0.0001 evaluation reward: 3.65
 episode: 2268 score: 4.0 memory length: 475205 epsilon:
 0.25709212000931764 steps: 262 lr: 0.0001 evaluation reward: 3.65
 episode: 2269 score: 3.0 memory length: 475423 epsilon: 0.2566604800093149
 steps: 218 lr: 0.0001 evaluation reward: 3.64
 episode: 2270 score: 3.0 memory length: 475650 epsilon:
 0.25621102000931206 steps: 227 lr: 0.0001 evaluation reward: 3.62
 episode: 2271 score: 5.0 memory length: 475958 epsilon: 0.2556011800093082
 steps: 308 lr: 0.0001 evaluation reward: 3.65
 episode: 2272 score: 3.0 memory length: 476187 epsilon:
 0.25514776000930534 steps: 229 lr: 0.0001 evaluation reward: 3.64
 episode: 2273 score: 6.0 memory length: 476519 epsilon: 0.2544904000093012
 steps: 332 lr: 0.0001 evaluation reward: 3.67
 episode: 2274 score: 3.0 memory length: 476732 epsilon: 0.2540686600092985
 steps: 213 lr: 0.0001 evaluation reward: 3.67
 episode: 2275 score: 3.0 memory length: 476942 epsilon: 0.2536528600092959
 steps: 210 lr: 0.0001 evaluation reward: 3.66
 episode: 2276 score: 2.0 memory length: 477122 epsilon: 0.2532964600092936
 steps: 180 lr: 0.0001 evaluation reward: 3.65
 episode: 2277 score: 5.0 memory length: 477453 epsilon: 0.2526410800092895
 steps: 331 lr: 0.0001 evaluation reward: 3.65
 episode: 2278 score: 3.0 memory length: 477666 epsilon: 0.2522193400092868
 steps: 213 lr: 0.0001 evaluation reward: 3.66
 episode: 2279 score: 3.0 memory length: 477876 epsilon: 0.2518035400092842
 steps: 210 lr: 0.0001 evaluation reward: 3.67
 episode: 2280 score: 3.0 memory length: 478102 epsilon:
 0.25135606000928135 steps: 226 lr: 0.0001 evaluation reward: 3.64
 episode: 2281 score: 4.0 memory length: 478377 epsilon: 0.2508115600092779
 steps: 275 lr: 0.0001 evaluation reward: 3.66
 episode: 2282 score: 2.0 memory length: 478559 epsilon: 0.2504512000092756
 steps: 182 lr: 0.0001 evaluation reward: 3.65
 episode: 2283 score: 3.0 memory length: 478786 epsilon: 0.2500017400092728
 steps: 227 lr: 0.0001 evaluation reward: 3.64

episode: 2284 score: 4.0 memory length: 479078 epsilon:
 0.24942358000926912 steps: 292 lr: 0.0001 evaluation reward: 3.66
 episode: 2285 score: 2.0 memory length: 479258 epsilon:
 0.24906718000926686 steps: 180 lr: 0.0001 evaluation reward: 3.65
 episode: 2286 score: 4.0 memory length: 479501 epsilon:
 0.24858604000926382 steps: 243 lr: 0.0001 evaluation reward: 3.66
 episode: 2287 score: 7.0 memory length: 479905 epsilon:
 0.24778612000925876 steps: 404 lr: 0.0001 evaluation reward: 3.69
 episode: 2288 score: 3.0 memory length: 480130 epsilon:
 0.24734062000925594 steps: 225 lr: 0.0001 evaluation reward: 3.7
 episode: 2289 score: 5.0 memory length: 480454 epsilon:
 0.24669910000925188 steps: 324 lr: 0.0001 evaluation reward: 3.7
 episode: 2290 score: 3.0 memory length: 480667 epsilon: 0.2462773600092492
 steps: 213 lr: 0.0001 evaluation reward: 3.69
 episode: 2291 score: 4.0 memory length: 480941 epsilon:
 0.24573484000924578 steps: 274 lr: 0.0001 evaluation reward: 3.69
 episode: 2292 score: 3.0 memory length: 481154 epsilon: 0.2453131000092431
 steps: 213 lr: 0.0001 evaluation reward: 3.69
 episode: 2293 score: 2.0 memory length: 481335 epsilon:
 0.24495472000924084 steps: 181 lr: 0.0001 evaluation reward: 3.66
 episode: 2294 score: 4.0 memory length: 481574 epsilon:
 0.24448150000923785 steps: 239 lr: 0.0001 evaluation reward: 3.6
 episode: 2295 score: 2.0 memory length: 481756 epsilon:
 0.24412114000923557 steps: 182 lr: 0.0001 evaluation reward: 3.58
 episode: 2296 score: 4.0 memory length: 482031 epsilon:
 0.24357664000923213 steps: 275 lr: 0.0001 evaluation reward: 3.57
 episode: 2297 score: 5.0 memory length: 482322 epsilon:
 0.24300046000922848 steps: 291 lr: 0.0001 evaluation reward: 3.57
 episode: 2298 score: 2.0 memory length: 482504 epsilon: 0.2426401000092262
 steps: 182 lr: 0.0001 evaluation reward: 3.57
 episode: 2299 score: 4.0 memory length: 482783 epsilon: 0.2420876800092227
 steps: 279 lr: 0.0001 evaluation reward: 3.59
 episode: 2300 score: 4.0 memory length: 483058 epsilon:
 0.24154318000921926 steps: 275 lr: 0.0001 evaluation reward: 3.59
 episode: 2301 score: 4.0 memory length: 483333 epsilon:
 0.24099868000921582 steps: 275 lr: 0.0001 evaluation reward: 3.6
 episode: 2302 score: 4.0 memory length: 483611 epsilon:
 0.24044824000921233 steps: 278 lr: 0.0001 evaluation reward: 3.6
 episode: 2303 score: 1.0 memory length: 483762 epsilon:
 0.24014926000921044 steps: 151 lr: 0.0001 evaluation reward: 3.59
 episode: 2304 score: 3.0 memory length: 483975 epsilon:
 0.23972752000920777 steps: 213 lr: 0.0001 evaluation reward: 3.53
 episode: 2305 score: 5.0 memory length: 484279 epsilon:
 0.23912560000920396 steps: 304 lr: 0.0001 evaluation reward: 3.56
 episode: 2306 score: 4.0 memory length: 484554 epsilon:
 0.23858110000920052 steps: 275 lr: 0.0001 evaluation reward: 3.54
 episode: 2307 score: 3.0 memory length: 484767 epsilon:
 0.23815936000919785 steps: 213 lr: 0.0001 evaluation reward: 3.51

episode: 2308 score: 5.0 memory length: 485051 epsilon: 0.2375970400091943
 steps: 284 lr: 0.0001 evaluation reward: 3.52
 episode: 2309 score: 3.0 memory length: 485277 epsilon:
 0.23714956000919146 steps: 226 lr: 0.0001 evaluation reward: 3.53
 episode: 2310 score: 3.0 memory length: 485489 epsilon: 0.2367298000091888
 steps: 212 lr: 0.0001 evaluation reward: 3.52
 episode: 2311 score: 4.0 memory length: 485735 epsilon:
 0.23624272000918572 steps: 246 lr: 0.0001 evaluation reward: 3.53
 episode: 2312 score: 3.0 memory length: 485960 epsilon: 0.2357972200091829
 steps: 225 lr: 0.0001 evaluation reward: 3.54
 episode: 2313 score: 3.0 memory length: 486172 epsilon:
 0.23537746000918025 steps: 212 lr: 0.0001 evaluation reward: 3.55
 episode: 2314 score: 4.0 memory length: 486452 epsilon:
 0.23482306000917674 steps: 280 lr: 0.0001 evaluation reward: 3.52
 episode: 2315 score: 3.0 memory length: 486662 epsilon: 0.2344072600091741
 steps: 210 lr: 0.0001 evaluation reward: 3.52
 episode: 2316 score: 5.0 memory length: 486968 epsilon:
 0.23380138000917028 steps: 306 lr: 0.0001 evaluation reward: 3.52
 episode: 2317 score: 4.0 memory length: 487210 epsilon:
 0.23332222000916725 steps: 242 lr: 0.0001 evaluation reward: 3.54
 episode: 2318 score: 3.0 memory length: 487422 epsilon: 0.2329024600091646
 steps: 212 lr: 0.0001 evaluation reward: 3.55
 episode: 2319 score: 5.0 memory length: 487731 epsilon:
 0.23229064000916072 steps: 309 lr: 0.0001 evaluation reward: 3.57
 episode: 2320 score: 4.0 memory length: 488010 epsilon:
 0.23173822000915723 steps: 279 lr: 0.0001 evaluation reward: 3.59
 episode: 2321 score: 3.0 memory length: 488258 epsilon:
 0.23124718000915412 steps: 248 lr: 0.0001 evaluation reward: 3.59
 episode: 2322 score: 3.0 memory length: 488471 epsilon:
 0.23082544000915145 steps: 213 lr: 0.0001 evaluation reward: 3.58
 episode: 2323 score: 8.0 memory length: 488908 epsilon:
 0.22996018000914598 steps: 437 lr: 0.0001 evaluation reward: 3.62
 episode: 2324 score: 4.0 memory length: 489150 epsilon:
 0.22948102000914294 steps: 242 lr: 0.0001 evaluation reward: 3.61
 episode: 2325 score: 3.0 memory length: 489375 epsilon:
 0.22903552000914013 steps: 225 lr: 0.0001 evaluation reward: 3.59
 episode: 2326 score: 6.0 memory length: 489703 epsilon:
 0.22838608000913602 steps: 328 lr: 0.0001 evaluation reward: 3.63
 episode: 2327 score: 4.0 memory length: 489977 epsilon:
 0.22784356000913258 steps: 274 lr: 0.0001 evaluation reward: 3.62
 episode: 2328 score: 3.0 memory length: 490209 epsilon:
 0.22738420000912968 steps: 232 lr: 0.0001 evaluation reward: 3.6
 episode: 2329 score: 3.0 memory length: 490440 epsilon:
 0.22692682000912678 steps: 231 lr: 0.0001 evaluation reward: 3.6
 episode: 2330 score: 3.0 memory length: 490653 epsilon:
 0.22650508000912412 steps: 213 lr: 0.0001 evaluation reward: 3.61
 episode: 2331 score: 3.0 memory length: 490865 epsilon:
 0.22608532000912146 steps: 212 lr: 0.0001 evaluation reward: 3.61

episode: 2332 score: 4.0 memory length: 491145 epsilon:
 0.22553092000911795 steps: 280 lr: 0.0001 evaluation reward: 3.61
 episode: 2333 score: 3.0 memory length: 491355 epsilon:
 0.22511512000911532 steps: 210 lr: 0.0001 evaluation reward: 3.62
 episode: 2334 score: 3.0 memory length: 491568 epsilon:
 0.22469338000911265 steps: 213 lr: 0.0001 evaluation reward: 3.6
 episode: 2335 score: 4.0 memory length: 491824 epsilon:
 0.22418650000910945 steps: 256 lr: 0.0001 evaluation reward: 3.6
 episode: 2336 score: 3.0 memory length: 492069 epsilon:
 0.22370140000910638 steps: 245 lr: 0.0001 evaluation reward: 3.59
 episode: 2337 score: 4.0 memory length: 492343 epsilon:
 0.22315888000910294 steps: 274 lr: 0.0001 evaluation reward: 3.59
 episode: 2338 score: 3.0 memory length: 492553 epsilon: 0.2227430800091003
 steps: 210 lr: 0.0001 evaluation reward: 3.58
 episode: 2339 score: 3.0 memory length: 492780 epsilon:
 0.22229362000909747 steps: 227 lr: 0.0001 evaluation reward: 3.57
 episode: 2340 score: 6.0 memory length: 493153 epsilon: 0.2215550800090928
 steps: 373 lr: 0.0001 evaluation reward: 3.62
 episode: 2341 score: 1.0 memory length: 493303 epsilon:
 0.22125808000909092 steps: 150 lr: 0.0001 evaluation reward: 3.6
 episode: 2342 score: 3.0 memory length: 493516 epsilon:
 0.22083634000908825 steps: 213 lr: 0.0001 evaluation reward: 3.6
 episode: 2343 score: 3.0 memory length: 493725 epsilon:
 0.22042252000908563 steps: 209 lr: 0.0001 evaluation reward: 3.58
 episode: 2344 score: 5.0 memory length: 494040 epsilon:
 0.21979882000908169 steps: 315 lr: 0.0001 evaluation reward: 3.6
 episode: 2345 score: 3.0 memory length: 494253 epsilon:
 0.21937708000907902 steps: 213 lr: 0.0001 evaluation reward: 3.58
 episode: 2346 score: 2.0 memory length: 494435 epsilon:
 0.21901672000907674 steps: 182 lr: 0.0001 evaluation reward: 3.57
 episode: 2347 score: 2.0 memory length: 494616 epsilon:
 0.21865834000907447 steps: 181 lr: 0.0001 evaluation reward: 3.55
 episode: 2348 score: 3.0 memory length: 494828 epsilon:
 0.21823858000907181 steps: 212 lr: 0.0001 evaluation reward: 3.56
 episode: 2349 score: 5.0 memory length: 495151 epsilon:
 0.21759904000906777 steps: 323 lr: 0.0001 evaluation reward: 3.58
 episode: 2350 score: 3.0 memory length: 495364 epsilon: 0.2171773000090651
 steps: 213 lr: 0.0001 evaluation reward: 3.57
 episode: 2351 score: 4.0 memory length: 495639 epsilon:
 0.21663280000906165 steps: 275 lr: 0.0001 evaluation reward: 3.58
 episode: 2352 score: 2.0 memory length: 495821 epsilon:
 0.21627244000905937 steps: 182 lr: 0.0001 evaluation reward: 3.57
 episode: 2353 score: 3.0 memory length: 496031 epsilon:
 0.21585664000905674 steps: 210 lr: 0.0001 evaluation reward: 3.56
 episode: 2354 score: 4.0 memory length: 496308 epsilon:
 0.21530818000905327 steps: 277 lr: 0.0001 evaluation reward: 3.57
 episode: 2355 score: 4.0 memory length: 496582 epsilon:
 0.21476566000904984 steps: 274 lr: 0.0001 evaluation reward: 3.56

episode: 2356 score: 3.0 memory length: 496812 epsilon:
 0.21431026000904696 steps: 230 lr: 0.0001 evaluation reward: 3.56
 episode: 2357 score: 5.0 memory length: 497116 epsilon:
 0.21370834000904315 steps: 304 lr: 0.0001 evaluation reward: 3.56
 episode: 2358 score: 3.0 memory length: 497327 epsilon: 0.2132905600090405
 steps: 211 lr: 0.0001 evaluation reward: 3.56
 episode: 2359 score: 4.0 memory length: 497604 epsilon:
 0.21274210000903704 steps: 277 lr: 0.0001 evaluation reward: 3.57
 episode: 2360 score: 5.0 memory length: 497916 epsilon:
 0.21212434000903313 steps: 312 lr: 0.0001 evaluation reward: 3.57
 episode: 2361 score: 5.0 memory length: 498241 epsilon:
 0.21148084000902906 steps: 325 lr: 0.0001 evaluation reward: 3.59
 episode: 2362 score: 7.0 memory length: 498616 epsilon:
 0.21073834000902436 steps: 375 lr: 0.0001 evaluation reward: 3.62
 episode: 2363 score: 3.0 memory length: 498844 epsilon: 0.2102869000090215
 steps: 228 lr: 0.0001 evaluation reward: 3.63
 episode: 2364 score: 9.0 memory length: 499282 epsilon:
 0.20941966000901602 steps: 438 lr: 0.0001 evaluation reward: 3.69
 episode: 2365 score: 3.0 memory length: 499508 epsilon: 0.2089721800090132
 steps: 226 lr: 0.0001 evaluation reward: 3.66
 episode: 2366 score: 2.0 memory length: 499689 epsilon:
 0.20861380000901092 steps: 181 lr: 0.0001 evaluation reward: 3.65
 episode: 2367 score: 2.0 memory length: 499871 epsilon:
 0.20825344000900864 steps: 182 lr: 0.0001 evaluation reward: 3.64
 episode: 2368 score: 5.0 memory length: 500195 epsilon:
 0.20761192000900458 steps: 324 lr: 0.0001 evaluation reward: 3.65
 episode: 2369 score: 3.0 memory length: 500408 epsilon: 0.2071901800090019
 steps: 213 lr: 0.0001 evaluation reward: 3.65
 episode: 2370 score: 3.0 memory length: 500637 epsilon:
 0.20673676000899904 steps: 229 lr: 0.0001 evaluation reward: 3.65
 episode: 2371 score: 2.0 memory length: 500819 epsilon:
 0.20637640000899676 steps: 182 lr: 0.0001 evaluation reward: 3.62
 episode: 2372 score: 3.0 memory length: 501048 epsilon: 0.2059229800089939
 steps: 229 lr: 0.0001 evaluation reward: 3.62
 episode: 2373 score: 3.0 memory length: 501258 epsilon:
 0.20550718000899126 steps: 210 lr: 0.0001 evaluation reward: 3.59
 episode: 2374 score: 6.0 memory length: 501580 epsilon:
 0.20486962000898723 steps: 322 lr: 0.0001 evaluation reward: 3.62
 episode: 2375 score: 4.0 memory length: 501857 epsilon:
 0.20432116000898376 steps: 277 lr: 0.0001 evaluation reward: 3.63
 episode: 2376 score: 2.0 memory length: 502054 epsilon: 0.2039311000089813
 steps: 197 lr: 0.0001 evaluation reward: 3.63
 episode: 2377 score: 3.0 memory length: 502267 epsilon:
 0.20350936000897862 steps: 213 lr: 0.0001 evaluation reward: 3.61
 episode: 2378 score: 3.0 memory length: 502480 epsilon:
 0.20308762000897596 steps: 213 lr: 0.0001 evaluation reward: 3.61
 episode: 2379 score: 4.0 memory length: 502758 epsilon:
 0.20253718000897247 steps: 278 lr: 0.0001 evaluation reward: 3.62

episode: 2380 score: 4.0 memory length: 503019 epsilon: 0.2020204000089692
 steps: 261 lr: 0.0001 evaluation reward: 3.63
 episode: 2381 score: 4.0 memory length: 503296 epsilon:
 0.20147194000896573 steps: 277 lr: 0.0001 evaluation reward: 3.63
 episode: 2382 score: 3.0 memory length: 503509 epsilon:
 0.20105020000896306 steps: 213 lr: 0.0001 evaluation reward: 3.64
 episode: 2383 score: 3.0 memory length: 503752 epsilon:
 0.20056906000896002 steps: 243 lr: 0.0001 evaluation reward: 3.64
 episode: 2384 score: 3.0 memory length: 503981 epsilon:
 0.20011564000895715 steps: 229 lr: 0.0001 evaluation reward: 3.63
 episode: 2385 score: 2.0 memory length: 504163 epsilon:
 0.19975528000895487 steps: 182 lr: 0.0001 evaluation reward: 3.63
 episode: 2386 score: 4.0 memory length: 504422 epsilon:
 0.19924246000895163 steps: 259 lr: 0.0001 evaluation reward: 3.63
 episode: 2387 score: 3.0 memory length: 504672 epsilon: 0.1987474600089485
 steps: 250 lr: 0.0001 evaluation reward: 3.59
 episode: 2388 score: 3.0 memory length: 504885 epsilon:
 0.19832572000894583 steps: 213 lr: 0.0001 evaluation reward: 3.59
 episode: 2389 score: 5.0 memory length: 505154 epsilon:
 0.19779310000894246 steps: 269 lr: 0.0001 evaluation reward: 3.59
 episode: 2390 score: 2.0 memory length: 505335 epsilon: 0.1974347200089402
 steps: 181 lr: 0.0001 evaluation reward: 3.58
 episode: 2391 score: 2.0 memory length: 505516 epsilon:
 0.19707634000893792 steps: 181 lr: 0.0001 evaluation reward: 3.56
 episode: 2392 score: 3.0 memory length: 505742 epsilon: 0.1966288600089351
 steps: 226 lr: 0.0001 evaluation reward: 3.56
 episode: 2393 score: 3.0 memory length: 505955 epsilon:
 0.19620712000893242 steps: 213 lr: 0.0001 evaluation reward: 3.57
 episode: 2394 score: 3.0 memory length: 506186 epsilon:
 0.19574974000892953 steps: 231 lr: 0.0001 evaluation reward: 3.56
 episode: 2395 score: 3.0 memory length: 506417 epsilon:
 0.19529236000892664 steps: 231 lr: 0.0001 evaluation reward: 3.57
 episode: 2396 score: 3.0 memory length: 506630 epsilon:
 0.19487062000892397 steps: 213 lr: 0.0001 evaluation reward: 3.56
 episode: 2397 score: 3.0 memory length: 506859 epsilon: 0.1944172000089211
 steps: 229 lr: 0.0001 evaluation reward: 3.54
 episode: 2398 score: 3.0 memory length: 507107 epsilon: 0.193926160008918
 steps: 248 lr: 0.0001 evaluation reward: 3.55
 episode: 2399 score: 2.0 memory length: 507289 epsilon: 0.1935658000089157
 steps: 182 lr: 0.0001 evaluation reward: 3.53
 episode: 2400 score: 4.0 memory length: 507551 epsilon:
 0.19304704000891243 steps: 262 lr: 0.0001 evaluation reward: 3.53
 episode: 2401 score: 2.0 memory length: 507733 epsilon:
 0.19268668000891015 steps: 182 lr: 0.0001 evaluation reward: 3.51
 episode: 2402 score: 4.0 memory length: 508013 epsilon:
 0.19213228000890664 steps: 280 lr: 0.0001 evaluation reward: 3.51
 episode: 2403 score: 2.0 memory length: 508213 epsilon:
 0.19173628000890414 steps: 200 lr: 0.0001 evaluation reward: 3.52

episode: 2404 score: 3.0 memory length: 508442 epsilon:
 0.19128286000890127 steps: 229 lr: 0.0001 evaluation reward: 3.52
 episode: 2405 score: 4.0 memory length: 508738 epsilon:
 0.19069678000889756 steps: 296 lr: 0.0001 evaluation reward: 3.51
 episode: 2406 score: 3.0 memory length: 508951 epsilon: 0.1902750400088949
 steps: 213 lr: 0.0001 evaluation reward: 3.5
 episode: 2407 score: 2.0 memory length: 509133 epsilon: 0.1899146800088926
 steps: 182 lr: 0.0001 evaluation reward: 3.49
 episode: 2408 score: 2.0 memory length: 509315 epsilon:
 0.18955432000889033 steps: 182 lr: 0.0001 evaluation reward: 3.46
 episode: 2409 score: 2.0 memory length: 509497 epsilon:
 0.18919396000888805 steps: 182 lr: 0.0001 evaluation reward: 3.45
 episode: 2410 score: 6.0 memory length: 509840 epsilon:
 0.18851482000888375 steps: 343 lr: 0.0001 evaluation reward: 3.48
 episode: 2411 score: 4.0 memory length: 510106 epsilon:
 0.18798814000888042 steps: 266 lr: 0.0001 evaluation reward: 3.48
 episode: 2412 score: 2.0 memory length: 510303 epsilon:
 0.18759808000887795 steps: 197 lr: 0.0001 evaluation reward: 3.47
 episode: 2413 score: 3.0 memory length: 510514 epsilon: 0.1871803000088753
 steps: 211 lr: 0.0001 evaluation reward: 3.47
 episode: 2414 score: 5.0 memory length: 510795 epsilon: 0.1866239200088718
 steps: 281 lr: 0.0001 evaluation reward: 3.48
 episode: 2415 score: 2.0 memory length: 510977 epsilon: 0.1862635600088695
 steps: 182 lr: 0.0001 evaluation reward: 3.47
 episode: 2416 score: 3.0 memory length: 511190 epsilon:
 0.18584182000886684 steps: 213 lr: 0.0001 evaluation reward: 3.45
 episode: 2417 score: 3.0 memory length: 511403 epsilon:
 0.18542008000886417 steps: 213 lr: 0.0001 evaluation reward: 3.44
 episode: 2418 score: 2.0 memory length: 511585 epsilon: 0.1850597200088619
 steps: 182 lr: 0.0001 evaluation reward: 3.43
 episode: 2419 score: 3.0 memory length: 511798 epsilon:
 0.18463798000885923 steps: 213 lr: 0.0001 evaluation reward: 3.41
 episode: 2420 score: 5.0 memory length: 512122 epsilon:
 0.18399646000885517 steps: 324 lr: 0.0001 evaluation reward: 3.42
 episode: 2421 score: 4.0 memory length: 512364 epsilon:
 0.18351730000885214 steps: 242 lr: 0.0001 evaluation reward: 3.43
 episode: 2422 score: 6.0 memory length: 512740 epsilon:
 0.18277282000884743 steps: 376 lr: 0.0001 evaluation reward: 3.46
 episode: 2423 score: 3.0 memory length: 512951 epsilon:
 0.18235504000884478 steps: 211 lr: 0.0001 evaluation reward: 3.41
 episode: 2424 score: 5.0 memory length: 513242 epsilon:
 0.18177886000884114 steps: 291 lr: 0.0001 evaluation reward: 3.42
 episode: 2425 score: 4.0 memory length: 513501 epsilon: 0.1812660400088379
 steps: 259 lr: 0.0001 evaluation reward: 3.43
 episode: 2426 score: 3.0 memory length: 513714 epsilon:
 0.18084430000883522 steps: 213 lr: 0.0001 evaluation reward: 3.4
 episode: 2427 score: 4.0 memory length: 513972 epsilon: 0.180333460008832
 steps: 258 lr: 0.0001 evaluation reward: 3.4

episode: 2428 score: 3.0 memory length: 514185 epsilon:
 0.17991172000882932 steps: 213 lr: 0.0001 evaluation reward: 3.4
 episode: 2429 score: 4.0 memory length: 514461 epsilon:
 0.17936524000882587 steps: 276 lr: 0.0001 evaluation reward: 3.41
 episode: 2430 score: 4.0 memory length: 514721 epsilon: 0.1788504400088226
 steps: 260 lr: 0.0001 evaluation reward: 3.42
 episode: 2431 score: 3.0 memory length: 514934 epsilon:
 0.17842870000881994 steps: 213 lr: 0.0001 evaluation reward: 3.42
 episode: 2432 score: 3.0 memory length: 515144 epsilon: 0.1780129000088173
 steps: 210 lr: 0.0001 evaluation reward: 3.41
 episode: 2433 score: 3.0 memory length: 515357 epsilon:
 0.17759116000881464 steps: 213 lr: 0.0001 evaluation reward: 3.41
 episode: 2434 score: 4.0 memory length: 515596 epsilon:
 0.17711794000881165 steps: 239 lr: 0.0001 evaluation reward: 3.42
 episode: 2435 score: 3.0 memory length: 515807 epsilon: 0.176700160008809
 steps: 211 lr: 0.0001 evaluation reward: 3.41
 episode: 2436 score: 2.0 memory length: 515988 epsilon:
 0.17634178000880674 steps: 181 lr: 0.0001 evaluation reward: 3.4
 episode: 2437 score: 3.0 memory length: 516197 epsilon:
 0.17592796000880412 steps: 209 lr: 0.0001 evaluation reward: 3.39
 episode: 2438 score: 2.0 memory length: 516379 epsilon:
 0.17556760000880184 steps: 182 lr: 0.0001 evaluation reward: 3.38
 episode: 2439 score: 4.0 memory length: 516639 epsilon:
 0.17505280000879858 steps: 260 lr: 0.0001 evaluation reward: 3.39
 episode: 2440 score: 3.0 memory length: 516868 epsilon: 0.1745993800087957
 steps: 229 lr: 0.0001 evaluation reward: 3.36
 episode: 2441 score: 5.0 memory length: 517192 epsilon:
 0.17395786000879165 steps: 324 lr: 0.0001 evaluation reward: 3.4
 episode: 2442 score: 2.0 memory length: 517373 epsilon: 0.1735994800087894
 steps: 181 lr: 0.0001 evaluation reward: 3.39
 episode: 2443 score: 3.0 memory length: 517602 epsilon:
 0.17314606000878652 steps: 229 lr: 0.0001 evaluation reward: 3.39
 episode: 2444 score: 5.0 memory length: 517929 epsilon:
 0.17249860000878242 steps: 327 lr: 0.0001 evaluation reward: 3.39
 episode: 2445 score: 4.0 memory length: 518192 epsilon:
 0.17197786000877913 steps: 263 lr: 0.0001 evaluation reward: 3.4
 episode: 2446 score: 4.0 memory length: 518434 epsilon: 0.1714987000087761
 steps: 242 lr: 0.0001 evaluation reward: 3.42
 episode: 2447 score: 4.0 memory length: 518695 epsilon:
 0.17098192000877283 steps: 261 lr: 0.0001 evaluation reward: 3.44
 episode: 2448 score: 5.0 memory length: 518983 epsilon:
 0.17041168000876922 steps: 288 lr: 0.0001 evaluation reward: 3.46
 episode: 2449 score: 3.0 memory length: 519213 epsilon:
 0.16995628000876634 steps: 230 lr: 0.0001 evaluation reward: 3.44
 episode: 2450 score: 7.0 memory length: 519588 epsilon:
 0.16921378000876164 steps: 375 lr: 0.0001 evaluation reward: 3.48
 episode: 2451 score: 3.0 memory length: 519801 epsilon:
 0.16879204000875897 steps: 213 lr: 0.0001 evaluation reward: 3.47

episode: 2452 score: 5.0 memory length: 520130 epsilon:
 0.16814062000875485 steps: 329 lr: 0.0001 evaluation reward: 3.5
 episode: 2453 score: 4.0 memory length: 520390 epsilon: 0.1676258200087516
 steps: 260 lr: 0.0001 evaluation reward: 3.51
 episode: 2454 score: 5.0 memory length: 520716 epsilon: 0.1669803400087475
 steps: 326 lr: 0.0001 evaluation reward: 3.52
 episode: 2455 score: 7.0 memory length: 521109 epsilon:
 0.16620220000874258 steps: 393 lr: 0.0001 evaluation reward: 3.55
 episode: 2456 score: 5.0 memory length: 521395 epsilon: 0.165635920008739
 steps: 286 lr: 0.0001 evaluation reward: 3.57
 episode: 2457 score: 3.0 memory length: 521626 epsilon: 0.1651785400087361
 steps: 231 lr: 0.0001 evaluation reward: 3.55
 episode: 2458 score: 2.0 memory length: 521808 epsilon:
 0.16481818000873383 steps: 182 lr: 0.0001 evaluation reward: 3.54
 episode: 2459 score: 4.0 memory length: 522068 epsilon:
 0.16430338000873057 steps: 260 lr: 0.0001 evaluation reward: 3.54
 episode: 2460 score: 3.0 memory length: 522281 epsilon: 0.1638816400087279
 steps: 213 lr: 0.0001 evaluation reward: 3.52
 episode: 2461 score: 2.0 memory length: 522463 epsilon:
 0.16352128000872562 steps: 182 lr: 0.0001 evaluation reward: 3.49
 episode: 2462 score: 4.0 memory length: 522704 epsilon: 0.1630441000087226
 steps: 241 lr: 0.0001 evaluation reward: 3.46
 episode: 2463 score: 4.0 memory length: 522979 epsilon:
 0.16249960000871916 steps: 275 lr: 0.0001 evaluation reward: 3.47
 episode: 2464 score: 4.0 memory length: 523252 epsilon:
 0.16195906000871574 steps: 273 lr: 0.0001 evaluation reward: 3.42
 episode: 2465 score: 2.0 memory length: 523433 epsilon:
 0.16160068000871347 steps: 181 lr: 0.0001 evaluation reward: 3.41
 episode: 2466 score: 4.0 memory length: 523693 epsilon: 0.1610858800087102
 steps: 260 lr: 0.0001 evaluation reward: 3.43
 episode: 2467 score: 5.0 memory length: 524019 epsilon:
 0.16044040000870613 steps: 326 lr: 0.0001 evaluation reward: 3.46
 episode: 2468 score: 4.0 memory length: 524275 epsilon:
 0.15993352000870292 steps: 256 lr: 0.0001 evaluation reward: 3.45
 episode: 2469 score: 4.0 memory length: 524550 epsilon:
 0.15938902000869948 steps: 275 lr: 0.0001 evaluation reward: 3.46
 episode: 2470 score: 6.0 memory length: 524899 epsilon: 0.1586980000086951
 steps: 349 lr: 0.0001 evaluation reward: 3.49
 episode: 2471 score: 3.0 memory length: 525112 epsilon:
 0.15827626000869244 steps: 213 lr: 0.0001 evaluation reward: 3.5
 episode: 2472 score: 6.0 memory length: 525467 epsilon: 0.157573360008688
 steps: 355 lr: 0.0001 evaluation reward: 3.53
 episode: 2473 score: 3.0 memory length: 525695 epsilon:
 0.15712192000868513 steps: 228 lr: 0.0001 evaluation reward: 3.53
 episode: 2474 score: 6.0 memory length: 526038 epsilon:
 0.15644278000868084 steps: 343 lr: 0.0001 evaluation reward: 3.53
 episode: 2475 score: 2.0 memory length: 526220 epsilon:
 0.15608242000867856 steps: 182 lr: 0.0001 evaluation reward: 3.51

episode: 2476 score: 4.0 memory length: 526482 epsilon:
 0.15556366000867528 steps: 262 lr: 0.0001 evaluation reward: 3.53
 episode: 2477 score: 3.0 memory length: 526733 epsilon:
 0.15506668000867213 steps: 251 lr: 0.0001 evaluation reward: 3.53
 episode: 2478 score: 4.0 memory length: 526990 epsilon: 0.1545578200086689
 steps: 257 lr: 0.0001 evaluation reward: 3.54
 episode: 2479 score: 3.0 memory length: 527203 epsilon:
 0.15413608000866624 steps: 213 lr: 0.0001 evaluation reward: 3.53
 episode: 2480 score: 3.0 memory length: 527416 epsilon:
 0.15371434000866357 steps: 213 lr: 0.0001 evaluation reward: 3.52
 episode: 2481 score: 4.0 memory length: 527658 epsilon:
 0.15323518000866054 steps: 242 lr: 0.0001 evaluation reward: 3.52
 episode: 2482 score: 4.0 memory length: 527932 epsilon: 0.1526926600086571
 steps: 274 lr: 0.0001 evaluation reward: 3.53
 episode: 2483 score: 3.0 memory length: 528158 epsilon:
 0.15224518000865428 steps: 226 lr: 0.0001 evaluation reward: 3.53
 episode: 2484 score: 3.0 memory length: 528390 epsilon:
 0.15178582000865137 steps: 232 lr: 0.0001 evaluation reward: 3.53
 episode: 2485 score: 3.0 memory length: 528603 epsilon: 0.1513640800086487
 steps: 213 lr: 0.0001 evaluation reward: 3.54
 episode: 2486 score: 5.0 memory length: 528899 epsilon: 0.150778000008645
 steps: 296 lr: 0.0001 evaluation reward: 3.55
 episode: 2487 score: 2.0 memory length: 529080 epsilon:
 0.15041962000864273 steps: 181 lr: 0.0001 evaluation reward: 3.54
 episode: 2488 score: 5.0 memory length: 529402 epsilon: 0.1497820600086387
 steps: 322 lr: 0.0001 evaluation reward: 3.56
 episode: 2489 score: 2.0 memory length: 529584 epsilon:
 0.14942170000863642 steps: 182 lr: 0.0001 evaluation reward: 3.53
 episode: 2490 score: 3.0 memory length: 529813 epsilon:
 0.14896828000863355 steps: 229 lr: 0.0001 evaluation reward: 3.54
 episode: 2491 score: 4.0 memory length: 530054 epsilon:
 0.14849110000863053 steps: 241 lr: 0.0001 evaluation reward: 3.56
 episode: 2492 score: 2.0 memory length: 530236 epsilon:
 0.14813074000862825 steps: 182 lr: 0.0001 evaluation reward: 3.55
 episode: 2493 score: 3.0 memory length: 530449 epsilon:
 0.14770900000862558 steps: 213 lr: 0.0001 evaluation reward: 3.55
 episode: 2494 score: 2.0 memory length: 530631 epsilon: 0.1473486400086233
 steps: 182 lr: 0.0001 evaluation reward: 3.54
 episode: 2495 score: 3.0 memory length: 530860 epsilon:
 0.14689522000862043 steps: 229 lr: 0.0001 evaluation reward: 3.54
 episode: 2496 score: 5.0 memory length: 531168 epsilon:
 0.14628538000861657 steps: 308 lr: 0.0001 evaluation reward: 3.56
 episode: 2497 score: 4.0 memory length: 531426 epsilon:
 0.14577454000861334 steps: 258 lr: 0.0001 evaluation reward: 3.57
 episode: 2498 score: 3.0 memory length: 531657 epsilon:
 0.14531716000861045 steps: 231 lr: 0.0001 evaluation reward: 3.57
 episode: 2499 score: 3.0 memory length: 531870 epsilon:
 0.14489542000860778 steps: 213 lr: 0.0001 evaluation reward: 3.58

episode: 2500 score: 3.0 memory length: 532098 epsilon:
 0.14444398000860492 steps: 228 lr: 0.0001 evaluation reward: 3.57
 episode: 2501 score: 3.0 memory length: 532324 epsilon: 0.1439965000086021
 steps: 226 lr: 0.0001 evaluation reward: 3.58
 episode: 2502 score: 3.0 memory length: 532537 epsilon:
 0.14357476000859942 steps: 213 lr: 0.0001 evaluation reward: 3.57
 episode: 2503 score: 3.0 memory length: 532766 epsilon:
 0.14312134000859655 steps: 229 lr: 0.0001 evaluation reward: 3.58
 episode: 2504 score: 3.0 memory length: 532995 epsilon:
 0.14266792000859368 steps: 229 lr: 0.0001 evaluation reward: 3.58
 episode: 2505 score: 3.0 memory length: 533206 epsilon:
 0.14225014000859104 steps: 211 lr: 0.0001 evaluation reward: 3.57
 episode: 2506 score: 4.0 memory length: 533445 epsilon:
 0.14177692000858805 steps: 239 lr: 0.0001 evaluation reward: 3.58
 episode: 2507 score: 3.0 memory length: 533693 epsilon:
 0.14128588000858494 steps: 248 lr: 0.0001 evaluation reward: 3.59
 episode: 2508 score: 4.0 memory length: 533952 epsilon: 0.1407730600085817
 steps: 259 lr: 0.0001 evaluation reward: 3.61
 episode: 2509 score: 7.0 memory length: 534342 epsilon: 0.1400008600085768
 steps: 390 lr: 0.0001 evaluation reward: 3.66
 episode: 2510 score: 3.0 memory length: 534555 epsilon:
 0.13957912000857414 steps: 213 lr: 0.0001 evaluation reward: 3.63
 episode: 2511 score: 6.0 memory length: 534908 epsilon:
 0.13888018000856972 steps: 353 lr: 0.0001 evaluation reward: 3.65
 episode: 2512 score: 5.0 memory length: 535220 epsilon: 0.1382624200085658
 steps: 312 lr: 0.0001 evaluation reward: 3.68
 episode: 2513 score: 3.0 memory length: 535433 epsilon:
 0.13784068000856314 steps: 213 lr: 0.0001 evaluation reward: 3.68
 episode: 2514 score: 3.0 memory length: 535646 epsilon:
 0.13741894000856048 steps: 213 lr: 0.0001 evaluation reward: 3.66
 episode: 2515 score: 3.0 memory length: 535876 epsilon: 0.1369635400085576
 steps: 230 lr: 0.0001 evaluation reward: 3.67
 episode: 2516 score: 5.0 memory length: 536198 epsilon:
 0.13632598000855356 steps: 322 lr: 0.0001 evaluation reward: 3.69
 episode: 2517 score: 3.0 memory length: 536411 epsilon: 0.1359042400085509
 steps: 213 lr: 0.0001 evaluation reward: 3.69
 episode: 2518 score: 5.0 memory length: 536719 epsilon:
 0.13529440000854703 steps: 308 lr: 0.0001 evaluation reward: 3.72
 episode: 2519 score: 5.0 memory length: 537018 epsilon: 0.1347023800085433
 steps: 299 lr: 0.0001 evaluation reward: 3.74
 episode: 2520 score: 5.0 memory length: 537307 epsilon:
 0.13413016000853967 steps: 289 lr: 0.0001 evaluation reward: 3.74
 episode: 2521 score: 5.0 memory length: 537598 epsilon:
 0.13355398000853602 steps: 291 lr: 0.0001 evaluation reward: 3.75
 episode: 2522 score: 6.0 memory length: 537916 epsilon:
 0.13292434000853204 steps: 318 lr: 0.0001 evaluation reward: 3.75
 episode: 2523 score: 4.0 memory length: 538178 epsilon:
 0.13240558000852876 steps: 262 lr: 0.0001 evaluation reward: 3.76

episode: 2524 score: 3.0 memory length: 538391 epsilon: 0.1319838400085261
 steps: 213 lr: 0.0001 evaluation reward: 3.74
 episode: 2525 score: 4.0 memory length: 538672 epsilon:
 0.13142746000852257 steps: 281 lr: 0.0001 evaluation reward: 3.74
 episode: 2526 score: 3.0 memory length: 538883 epsilon:
 0.13100968000851992 steps: 211 lr: 0.0001 evaluation reward: 3.74
 episode: 2527 score: 2.0 memory length: 539064 epsilon:
 0.13065130000851766 steps: 181 lr: 0.0001 evaluation reward: 3.72
 episode: 2528 score: 3.0 memory length: 539294 epsilon:
 0.13019590000851478 steps: 230 lr: 0.0001 evaluation reward: 3.72
 episode: 2529 score: 3.0 memory length: 539506 epsilon:
 0.12977614000851212 steps: 212 lr: 0.0001 evaluation reward: 3.71
 episode: 2530 score: 3.0 memory length: 539736 epsilon:
 0.12932074000850924 steps: 230 lr: 0.0001 evaluation reward: 3.7
 episode: 2531 score: 3.0 memory length: 539965 epsilon:
 0.12886732000850637 steps: 229 lr: 0.0001 evaluation reward: 3.7
 episode: 2532 score: 3.0 memory length: 540178 epsilon: 0.1284455800085037
 steps: 213 lr: 0.0001 evaluation reward: 3.7
 episode: 2533 score: 3.0 memory length: 540406 epsilon:
 0.12799414000850085 steps: 228 lr: 0.0001 evaluation reward: 3.7
 episode: 2534 score: 4.0 memory length: 540663 epsilon:
 0.12748528000849763 steps: 257 lr: 0.0001 evaluation reward: 3.7
 episode: 2535 score: 3.0 memory length: 540876 epsilon:
 0.12706354000849496 steps: 213 lr: 0.0001 evaluation reward: 3.7
 episode: 2536 score: 4.0 memory length: 541136 epsilon: 0.1265487400084917
 steps: 260 lr: 0.0001 evaluation reward: 3.72
 episode: 2537 score: 7.0 memory length: 541524 epsilon:
 0.12578050000848684 steps: 388 lr: 0.0001 evaluation reward: 3.76
 episode: 2538 score: 4.0 memory length: 541782 epsilon: 0.1252696600084836
 steps: 258 lr: 0.0001 evaluation reward: 3.78
 episode: 2539 score: 2.0 memory length: 541964 epsilon:
 0.12490930000848197 steps: 182 lr: 0.0001 evaluation reward: 3.76
 episode: 2540 score: 3.0 memory length: 542195 epsilon:
 0.12445192000848228 steps: 231 lr: 0.0001 evaluation reward: 3.76
 episode: 2541 score: 3.0 memory length: 542425 epsilon:
 0.12399652000848259 steps: 230 lr: 0.0001 evaluation reward: 3.74
 episode: 2542 score: 5.0 memory length: 542730 epsilon: 0.123392620008483
 steps: 305 lr: 0.0001 evaluation reward: 3.77
 episode: 2543 score: 4.0 memory length: 543009 epsilon:
 0.12284020000848338 steps: 279 lr: 0.0001 evaluation reward: 3.78
 episode: 2544 score: 4.0 memory length: 543268 epsilon:
 0.12232738000848373 steps: 259 lr: 0.0001 evaluation reward: 3.77
 episode: 2545 score: 3.0 memory length: 543499 epsilon:
 0.12187000000848404 steps: 231 lr: 0.0001 evaluation reward: 3.76
 episode: 2546 score: 3.0 memory length: 543709 epsilon:
 0.12145420000848432 steps: 210 lr: 0.0001 evaluation reward: 3.75
 episode: 2547 score: 3.0 memory length: 543921 epsilon:
 0.12103444000848461 steps: 212 lr: 0.0001 evaluation reward: 3.74

episode: 2548 score: 3.0 memory length: 544134 epsilon: 0.1206127000084849
 steps: 213 lr: 0.0001 evaluation reward: 3.72
 episode: 2549 score: 5.0 memory length: 544444 epsilon:
 0.11999890000848532 steps: 310 lr: 0.0001 evaluation reward: 3.74
 episode: 2550 score: 3.0 memory length: 544672 epsilon:
 0.11954746000848562 steps: 228 lr: 0.0001 evaluation reward: 3.7
 episode: 2551 score: 3.0 memory length: 544898 epsilon:
 0.11909998000848593 steps: 226 lr: 0.0001 evaluation reward: 3.7
 episode: 2552 score: 3.0 memory length: 545144 epsilon:
 0.11861290000848626 steps: 246 lr: 0.0001 evaluation reward: 3.68
 episode: 2553 score: 6.0 memory length: 545453 epsilon:
 0.11800108000848668 steps: 309 lr: 0.0001 evaluation reward: 3.7
 episode: 2554 score: 5.0 memory length: 545776 epsilon:
 0.11736154000848711 steps: 323 lr: 0.0001 evaluation reward: 3.7
 episode: 2555 score: 4.0 memory length: 546038 epsilon:
 0.11684278000848747 steps: 262 lr: 0.0001 evaluation reward: 3.67
 episode: 2556 score: 5.0 memory length: 546328 epsilon:
 0.11626858000848786 steps: 290 lr: 0.0001 evaluation reward: 3.67
 episode: 2557 score: 4.0 memory length: 546607 epsilon:
 0.11571616000848824 steps: 279 lr: 0.0001 evaluation reward: 3.68
 episode: 2558 score: 2.0 memory length: 546789 epsilon:
 0.11535580000848848 steps: 182 lr: 0.0001 evaluation reward: 3.68
 episode: 2559 score: 7.0 memory length: 547143 epsilon:
 0.11465488000848896 steps: 354 lr: 0.0001 evaluation reward: 3.71
 episode: 2560 score: 7.0 memory length: 547519 epsilon:
 0.11391040000848947 steps: 376 lr: 0.0001 evaluation reward: 3.75
 episode: 2561 score: 5.0 memory length: 547823 epsilon:
 0.11330848000848988 steps: 304 lr: 0.0001 evaluation reward: 3.78
 episode: 2562 score: 5.0 memory length: 548128 epsilon:
 0.11270458000849029 steps: 305 lr: 0.0001 evaluation reward: 3.79
 episode: 2563 score: 3.0 memory length: 548341 epsilon:
 0.11228284000849058 steps: 213 lr: 0.0001 evaluation reward: 3.78
 episode: 2564 score: 3.0 memory length: 548554 epsilon:
 0.11186110000849087 steps: 213 lr: 0.0001 evaluation reward: 3.77
 episode: 2565 score: 3.0 memory length: 548766 epsilon:
 0.11144134000849115 steps: 212 lr: 0.0001 evaluation reward: 3.78
 episode: 2566 score: 4.0 memory length: 549025 epsilon: 0.1109285200084915
 steps: 259 lr: 0.0001 evaluation reward: 3.78
 episode: 2567 score: 8.0 memory length: 549448 epsilon:
 0.11009098000849207 steps: 423 lr: 0.0001 evaluation reward: 3.81
 episode: 2568 score: 6.0 memory length: 549776 epsilon:
 0.10944154000849252 steps: 328 lr: 0.0001 evaluation reward: 3.83
 episode: 2569 score: 3.0 memory length: 550006 epsilon:
 0.10898614000849283 steps: 230 lr: 0.0001 evaluation reward: 3.82
 episode: 2570 score: 3.0 memory length: 550237 epsilon:
 0.10852876000849314 steps: 231 lr: 0.0001 evaluation reward: 3.79
 episode: 2571 score: 2.0 memory length: 550417 epsilon:
 0.10817236000849338 steps: 180 lr: 0.0001 evaluation reward: 3.78

episode: 2572 score: 3.0 memory length: 550629 epsilon:
 0.1077526000849367 steps: 212 lr: 0.0001 evaluation reward: 3.75
 episode: 2573 score: 4.0 memory length: 550885 epsilon:
 0.10724572000849401 steps: 256 lr: 0.0001 evaluation reward: 3.76
 episode: 2574 score: 3.0 memory length: 551094 epsilon: 0.1068319000084943
 steps: 209 lr: 0.0001 evaluation reward: 3.73
 episode: 2575 score: 3.0 memory length: 551323 epsilon: 0.1063784800084946
 steps: 229 lr: 0.0001 evaluation reward: 3.74
 episode: 2576 score: 3.0 memory length: 551550 epsilon:
 0.10592902000849491 steps: 227 lr: 0.0001 evaluation reward: 3.73
 episode: 2577 score: 4.0 memory length: 551827 epsilon:
 0.10538056000849529 steps: 277 lr: 0.0001 evaluation reward: 3.74
 episode: 2578 score: 5.0 memory length: 552097 epsilon:
 0.10484596000849565 steps: 270 lr: 0.0001 evaluation reward: 3.75
 episode: 2579 score: 3.0 memory length: 552309 epsilon:
 0.10442620000849594 steps: 212 lr: 0.0001 evaluation reward: 3.75
 episode: 2580 score: 2.0 memory length: 552491 epsilon:
 0.10406584000849618 steps: 182 lr: 0.0001 evaluation reward: 3.74
 episode: 2581 score: 8.0 memory length: 552879 epsilon: 0.1032976000084967
 steps: 388 lr: 0.0001 evaluation reward: 3.78
 episode: 2582 score: 3.0 memory length: 553107 epsilon:
 0.10284616000849701 steps: 228 lr: 0.0001 evaluation reward: 3.77
 episode: 2583 score: 4.0 memory length: 553365 epsilon:
 0.10233532000849736 steps: 258 lr: 0.0001 evaluation reward: 3.78
 episode: 2584 score: 3.0 memory length: 553596 epsilon:
 0.10187794000849767 steps: 231 lr: 0.0001 evaluation reward: 3.78
 episode: 2585 score: 3.0 memory length: 553824 epsilon:
 0.10142650000849798 steps: 228 lr: 0.0001 evaluation reward: 3.78
 episode: 2586 score: 8.0 memory length: 554263 epsilon:
 0.10055728000849858 steps: 439 lr: 0.0001 evaluation reward: 3.81
 episode: 2587 score: 8.0 memory length: 554540 epsilon:
 0.10000882000849895 steps: 277 lr: 0.0001 evaluation reward: 3.87
 episode: 2588 score: 5.0 memory length: 554847 epsilon:
 0.09940096000849936 steps: 307 lr: 0.0001 evaluation reward: 3.87
 episode: 2589 score: 3.0 memory length: 555057 epsilon:
 0.09898516000849965 steps: 210 lr: 0.0001 evaluation reward: 3.88
 episode: 2590 score: 2.0 memory length: 555238 epsilon:
 0.09862678000849989 steps: 181 lr: 0.0001 evaluation reward: 3.87
 episode: 2591 score: 3.0 memory length: 555451 epsilon:
 0.09820504000850018 steps: 213 lr: 0.0001 evaluation reward: 3.86
 episode: 2592 score: 4.0 memory length: 555714 epsilon:
 0.09768430000850054 steps: 263 lr: 0.0001 evaluation reward: 3.88
 episode: 2593 score: 6.0 memory length: 556089 epsilon:
 0.09694180000850104 steps: 375 lr: 0.0001 evaluation reward: 3.91
 episode: 2594 score: 3.0 memory length: 556298 epsilon:
 0.09652798000850132 steps: 209 lr: 0.0001 evaluation reward: 3.92
 episode: 2595 score: 3.0 memory length: 556507 epsilon: 0.0961141600085016
 steps: 209 lr: 0.0001 evaluation reward: 3.92

episode: 2596 score: 3.0 memory length: 556716 epsilon:
 0.09570034000850189 steps: 209 lr: 0.0001 evaluation reward: 3.9
 episode: 2597 score: 3.0 memory length: 556947 epsilon: 0.0952429600085022
 steps: 231 lr: 0.0001 evaluation reward: 3.89
 episode: 2598 score: 3.0 memory length: 557159 epsilon:
 0.09482320000850249 steps: 212 lr: 0.0001 evaluation reward: 3.89
 episode: 2599 score: 3.0 memory length: 557371 epsilon:
 0.09440344000850277 steps: 212 lr: 0.0001 evaluation reward: 3.89
 episode: 2600 score: 4.0 memory length: 557647 epsilon:
 0.09385696000850315 steps: 276 lr: 0.0001 evaluation reward: 3.9
 episode: 2601 score: 3.0 memory length: 557859 epsilon:
 0.09343720000850343 steps: 212 lr: 0.0001 evaluation reward: 3.9
 episode: 2602 score: 3.0 memory length: 558071 epsilon:
 0.09301744000850372 steps: 212 lr: 0.0001 evaluation reward: 3.9
 episode: 2603 score: 6.0 memory length: 558404 epsilon:
 0.09235810000850417 steps: 333 lr: 0.0001 evaluation reward: 3.93
 episode: 2604 score: 5.0 memory length: 558708 epsilon:
 0.09175618000850458 steps: 304 lr: 0.0001 evaluation reward: 3.95
 episode: 2605 score: 3.0 memory length: 558936 epsilon:
 0.09130474000850489 steps: 228 lr: 0.0001 evaluation reward: 3.95
 episode: 2606 score: 4.0 memory length: 559216 epsilon:
 0.09075034000850526 steps: 280 lr: 0.0001 evaluation reward: 3.95
 episode: 2607 score: 3.0 memory length: 559425 epsilon:
 0.09033652000850555 steps: 209 lr: 0.0001 evaluation reward: 3.95
 episode: 2608 score: 3.0 memory length: 559634 epsilon:
 0.08992270000850583 steps: 209 lr: 0.0001 evaluation reward: 3.94
 episode: 2609 score: 4.0 memory length: 559890 epsilon:
 0.08941582000850617 steps: 256 lr: 0.0001 evaluation reward: 3.91
 episode: 2610 score: 2.0 memory length: 560072 epsilon:
 0.08905546000850642 steps: 182 lr: 0.0001 evaluation reward: 3.9
 episode: 2611 score: 3.0 memory length: 560284 epsilon:
 0.08863570000850671 steps: 212 lr: 0.0001 evaluation reward: 3.87
 episode: 2612 score: 3.0 memory length: 560496 epsilon: 0.088215940008507
 steps: 212 lr: 0.0001 evaluation reward: 3.85
 episode: 2613 score: 4.0 memory length: 560740 epsilon:
 0.08773282000850732 steps: 244 lr: 0.0001 evaluation reward: 3.86
 episode: 2614 score: 3.0 memory length: 560952 epsilon:
 0.08731306000850761 steps: 212 lr: 0.0001 evaluation reward: 3.86
 episode: 2615 score: 3.0 memory length: 561162 epsilon:
 0.08689726000850789 steps: 210 lr: 0.0001 evaluation reward: 3.86
 episode: 2616 score: 4.0 memory length: 561437 epsilon:
 0.08635276000850826 steps: 275 lr: 0.0001 evaluation reward: 3.85
 episode: 2617 score: 4.0 memory length: 561711 epsilon:
 0.08581024000850863 steps: 274 lr: 0.0001 evaluation reward: 3.86
 episode: 2618 score: 4.0 memory length: 561972 epsilon:
 0.08529346000850899 steps: 261 lr: 0.0001 evaluation reward: 3.85
 episode: 2619 score: 6.0 memory length: 562306 epsilon:
 0.08463214000850944 steps: 334 lr: 0.0001 evaluation reward: 3.86

episode: 2620	score: 4.0	memory length: 562585	epsilon:
0.08407972000850981	steps: 279	lr: 0.0001	evaluation reward: 3.85
episode: 2621	score: 5.0	memory length: 562932	epsilon:
0.08339266000851028	steps: 347	lr: 0.0001	evaluation reward: 3.85
episode: 2622	score: 3.0	memory length: 563144	epsilon:
0.08297290000851057	steps: 212	lr: 0.0001	evaluation reward: 3.82
episode: 2623	score: 4.0	memory length: 563401	epsilon:
0.08246404000851092	steps: 257	lr: 0.0001	evaluation reward: 3.82
episode: 2624	score: 3.0	memory length: 563630	epsilon:
0.08201062000851123	steps: 229	lr: 0.0001	evaluation reward: 3.82
episode: 2625	score: 7.0	memory length: 564034	epsilon:
0.08121070000851177	steps: 404	lr: 0.0001	evaluation reward: 3.85
episode: 2626	score: 3.0	memory length: 564245	epsilon:
0.08079292000851206	steps: 211	lr: 0.0001	evaluation reward: 3.85
episode: 2627	score: 2.0	memory length: 564427	epsilon: 0.0804325600085123
steps: 182	lr: 0.0001	evaluation reward: 3.85	
episode: 2628	score: 3.0	memory length: 564640	epsilon:
0.08001082000851259	steps: 213	lr: 0.0001	evaluation reward: 3.85
episode: 2629	score: 4.0	memory length: 564902	epsilon:
0.07949206000851294	steps: 262	lr: 0.0001	evaluation reward: 3.86
episode: 2630	score: 3.0	memory length: 565132	epsilon:
0.07903666000851325	steps: 230	lr: 0.0001	evaluation reward: 3.86
episode: 2631	score: 5.0	memory length: 565419	epsilon:
0.07846840000851364	steps: 287	lr: 0.0001	evaluation reward: 3.88
episode: 2632	score: 7.0	memory length: 565845	epsilon:
0.07762492000851422	steps: 426	lr: 0.0001	evaluation reward: 3.92
episode: 2633	score: 4.0	memory length: 566105	epsilon:
0.07711012000851457	steps: 260	lr: 0.0001	evaluation reward: 3.93
episode: 2634	score: 2.0	memory length: 566287	epsilon:
0.07674976000851481	steps: 182	lr: 0.0001	evaluation reward: 3.91
episode: 2635	score: 4.0	memory length: 566549	epsilon:
0.07623100000851517	steps: 262	lr: 0.0001	evaluation reward: 3.92
episode: 2636	score: 4.0	memory length: 566808	epsilon:
0.07571818000851552	steps: 259	lr: 0.0001	evaluation reward: 3.92
episode: 2637	score: 4.0	memory length: 567050	epsilon:
0.07523902000851584	steps: 242	lr: 0.0001	evaluation reward: 3.89
episode: 2638	score: 3.0	memory length: 567260	epsilon:
0.07482322000851613	steps: 210	lr: 0.0001	evaluation reward: 3.88
episode: 2639	score: 3.0	memory length: 567486	epsilon:
0.07437574000851643	steps: 226	lr: 0.0001	evaluation reward: 3.89
episode: 2640	score: 3.0	memory length: 567698	epsilon:
0.07395598000851672	steps: 212	lr: 0.0001	evaluation reward: 3.89
episode: 2641	score: 3.0	memory length: 567927	epsilon:
0.07350256000851703	steps: 229	lr: 0.0001	evaluation reward: 3.89
episode: 2642	score: 4.0	memory length: 568168	epsilon:
0.07302538000851735	steps: 241	lr: 0.0001	evaluation reward: 3.88
episode: 2643	score: 1.0	memory length: 568319	epsilon:
0.07272640000851756	steps: 151	lr: 0.0001	evaluation reward: 3.85

episode: 2644 score: 3.0 memory length: 568550 epsilon:
 0.07226902000851787 steps: 231 lr: 0.0001 evaluation reward: 3.84
 episode: 2645 score: 4.0 memory length: 568827 epsilon:
 0.07172056000851824 steps: 277 lr: 0.0001 evaluation reward: 3.85
 episode: 2646 score: 3.0 memory length: 569040 epsilon:
 0.07129882000851853 steps: 213 lr: 0.0001 evaluation reward: 3.85
 episode: 2647 score: 3.0 memory length: 569253 epsilon:
 0.07087708000851882 steps: 213 lr: 0.0001 evaluation reward: 3.85
 episode: 2648 score: 5.0 memory length: 569523 epsilon:
 0.07034248000851918 steps: 270 lr: 0.0001 evaluation reward: 3.87
 episode: 2649 score: 9.0 memory length: 569866 epsilon:
 0.06966334000851965 steps: 343 lr: 0.0001 evaluation reward: 3.91
 episode: 2650 score: 2.0 memory length: 570048 epsilon: 0.0693029800085199
 steps: 182 lr: 0.0001 evaluation reward: 3.9
 episode: 2651 score: 3.0 memory length: 570274 epsilon: 0.0688555000085202
 steps: 226 lr: 0.0001 evaluation reward: 3.9
 episode: 2652 score: 5.0 memory length: 570565 epsilon:
 0.06827932000852059 steps: 291 lr: 0.0001 evaluation reward: 3.92
 episode: 2653 score: 4.0 memory length: 570823 epsilon:
 0.06776848000852094 steps: 258 lr: 0.0001 evaluation reward: 3.9
 episode: 2654 score: 5.0 memory length: 571120 epsilon:
 0.06718042000852134 steps: 297 lr: 0.0001 evaluation reward: 3.9
 episode: 2655 score: 7.0 memory length: 571461 epsilon: 0.0665052400085218
 steps: 341 lr: 0.0001 evaluation reward: 3.93
 episode: 2656 score: 3.0 memory length: 571672 epsilon:
 0.06608746000852209 steps: 211 lr: 0.0001 evaluation reward: 3.91
 episode: 2657 score: 8.0 memory length: 572094 epsilon:
 0.06525190000852266 steps: 422 lr: 0.0001 evaluation reward: 3.95
 episode: 2658 score: 4.0 memory length: 572336 epsilon:
 0.06477274000852298 steps: 242 lr: 0.0001 evaluation reward: 3.97
 episode: 2659 score: 3.0 memory length: 572549 epsilon:
 0.06435100000852327 steps: 213 lr: 0.0001 evaluation reward: 3.93
 episode: 2660 score: 3.0 memory length: 572762 epsilon:
 0.06392926000852356 steps: 213 lr: 0.0001 evaluation reward: 3.89
 episode: 2661 score: 3.0 memory length: 572974 epsilon:
 0.06350950000852384 steps: 212 lr: 0.0001 evaluation reward: 3.87
 episode: 2662 score: 3.0 memory length: 573184 epsilon:
 0.06309370000852413 steps: 210 lr: 0.0001 evaluation reward: 3.85
 episode: 2663 score: 2.0 memory length: 573365 epsilon:
 0.06273532000852437 steps: 181 lr: 0.0001 evaluation reward: 3.84
 episode: 2664 score: 3.0 memory length: 573575 epsilon:
 0.062319520008524656 steps: 210 lr: 0.0001 evaluation reward: 3.84
 episode: 2665 score: 7.0 memory length: 573953 epsilon:
 0.06157108000852517 steps: 378 lr: 0.0001 evaluation reward: 3.88
 episode: 2666 score: 4.0 memory length: 574229 epsilon:
 0.06102460000852554 steps: 276 lr: 0.0001 evaluation reward: 3.88
 episode: 2667 score: 5.0 memory length: 574520 epsilon:
 0.06044842000852593 steps: 291 lr: 0.0001 evaluation reward: 3.85

episode: 2668 score: 3.0 memory length: 574749 epsilon:
 0.05999500000852624 steps: 229 lr: 0.0001 evaluation reward: 3.82
 episode: 2669 score: 6.0 memory length: 575088 epsilon: 0.0593237800085267
 steps: 339 lr: 0.0001 evaluation reward: 3.85
 episode: 2670 score: 3.0 memory length: 575301 epsilon:
 0.05890204000852699 steps: 213 lr: 0.0001 evaluation reward: 3.85
 episode: 2671 score: 3.0 memory length: 575514 epsilon:
 0.058480300008527275 steps: 213 lr: 0.0001 evaluation reward: 3.86
 episode: 2672 score: 5.0 memory length: 575818 epsilon:
 0.057878380008527686 steps: 304 lr: 0.0001 evaluation reward: 3.88
 episode: 2673 score: 3.0 memory length: 576046 epsilon:
 0.057426940008527994 steps: 228 lr: 0.0001 evaluation reward: 3.87
 episode: 2674 score: 4.0 memory length: 576305 epsilon:
 0.05691412000852834 steps: 259 lr: 0.0001 evaluation reward: 3.88
 episode: 2675 score: 9.0 memory length: 576736 epsilon:
 0.056060740008528925 steps: 431 lr: 0.0001 evaluation reward: 3.94
 episode: 2676 score: 5.0 memory length: 577063 epsilon:
 0.05541328000852937 steps: 327 lr: 0.0001 evaluation reward: 3.96
 episode: 2677 score: 5.0 memory length: 577371 epsilon:
 0.05480344000852978 steps: 308 lr: 0.0001 evaluation reward: 3.97
 episode: 2678 score: 4.0 memory length: 577646 epsilon:
 0.054258940008530154 steps: 275 lr: 0.0001 evaluation reward: 3.96
 episode: 2679 score: 4.0 memory length: 577906 epsilon:
 0.053744140008530505 steps: 260 lr: 0.0001 evaluation reward: 3.97
 episode: 2680 score: 4.0 memory length: 578168 epsilon:
 0.05322538000853086 steps: 262 lr: 0.0001 evaluation reward: 3.99
 episode: 2681 score: 4.0 memory length: 578465 epsilon:
 0.05263732000853126 steps: 297 lr: 0.0001 evaluation reward: 3.95
 episode: 2682 score: 2.0 memory length: 578647 epsilon:
 0.052276960008531506 steps: 182 lr: 0.0001 evaluation reward: 3.94
 episode: 2683 score: 6.0 memory length: 578999 epsilon:
 0.05158000000853198 steps: 352 lr: 0.0001 evaluation reward: 3.96
 episode: 2684 score: 7.0 memory length: 579375 epsilon:
 0.05083552000853249 steps: 376 lr: 0.0001 evaluation reward: 4.0
 episode: 2685 score: 4.0 memory length: 579655 epsilon:
 0.05028112000853287 steps: 280 lr: 0.0001 evaluation reward: 4.01
 episode: 2686 score: 4.0 memory length: 579913 epsilon:
 0.049770280008533216 steps: 258 lr: 0.0001 evaluation reward: 3.97
 episode: 2687 score: 3.0 memory length: 580123 epsilon: 0.0493544800085335
 steps: 210 lr: 0.0001 evaluation reward: 3.92
 episode: 2688 score: 3.0 memory length: 580336 epsilon:
 0.04893274000853379 steps: 213 lr: 0.0001 evaluation reward: 3.9
 episode: 2689 score: 3.0 memory length: 580546 epsilon:
 0.04851694000853407 steps: 210 lr: 0.0001 evaluation reward: 3.9
 episode: 2690 score: 4.0 memory length: 580824 epsilon:
 0.047966500008534446 steps: 278 lr: 0.0001 evaluation reward: 3.92
 episode: 2691 score: 3.0 memory length: 581054 epsilon:
 0.04751110000853476 steps: 230 lr: 0.0001 evaluation reward: 3.92

episode: 2692 score: 3.0 memory length: 581265 epsilon:
 0.04709332000853504 steps: 211 lr: 0.0001 evaluation reward: 3.91
 episode: 2693 score: 3.0 memory length: 581478 epsilon:
 0.04667158000853533 steps: 213 lr: 0.0001 evaluation reward: 3.88
 episode: 2694 score: 3.0 memory length: 581705 epsilon:
 0.046222120008535636 steps: 227 lr: 0.0001 evaluation reward: 3.88
 episode: 2695 score: 4.0 memory length: 582003 epsilon:
 0.04563208000853604 steps: 298 lr: 0.0001 evaluation reward: 3.89
 episode: 2696 score: 3.0 memory length: 582230 epsilon:
 0.045182620008536345 steps: 227 lr: 0.0001 evaluation reward: 3.89
 episode: 2697 score: 5.0 memory length: 582556 epsilon:
 0.044537140008536785 steps: 326 lr: 0.0001 evaluation reward: 3.91
 episode: 2698 score: 3.0 memory length: 582769 epsilon:
 0.04411540000853707 steps: 213 lr: 0.0001 evaluation reward: 3.91
 episode: 2699 score: 5.0 memory length: 583122 epsilon:
 0.04341646000853755 steps: 353 lr: 0.0001 evaluation reward: 3.93
 episode: 2700 score: 4.0 memory length: 583379 epsilon: 0.0429076000085379
 steps: 257 lr: 0.0001 evaluation reward: 3.93
 episode: 2701 score: 4.0 memory length: 583641 epsilon:
 0.04238884000853825 steps: 262 lr: 0.0001 evaluation reward: 3.94
 episode: 2702 score: 3.0 memory length: 583871 epsilon:
 0.04193344000853856 steps: 230 lr: 0.0001 evaluation reward: 3.94
 episode: 2703 score: 4.0 memory length: 584113 epsilon:
 0.04145428000853889 steps: 242 lr: 0.0001 evaluation reward: 3.92
 episode: 2704 score: 3.0 memory length: 584343 epsilon: 0.0409988800085392
 steps: 230 lr: 0.0001 evaluation reward: 3.9
 episode: 2705 score: 4.0 memory length: 584603 epsilon:
 0.04048408000853955 steps: 260 lr: 0.0001 evaluation reward: 3.91
 episode: 2706 score: 3.0 memory length: 584834 epsilon:
 0.04002670000853986 steps: 231 lr: 0.0001 evaluation reward: 3.9
 episode: 2707 score: 3.0 memory length: 585063 epsilon:
 0.03957328000854017 steps: 229 lr: 0.0001 evaluation reward: 3.9
 episode: 2708 score: 7.0 memory length: 585420 epsilon:
 0.03886642000854065 steps: 357 lr: 0.0001 evaluation reward: 3.94
 episode: 2709 score: 5.0 memory length: 585747 epsilon:
 0.038218960008541095 steps: 327 lr: 0.0001 evaluation reward: 3.95
 episode: 2710 score: 4.0 memory length: 586007 epsilon:
 0.037704160008541446 steps: 260 lr: 0.0001 evaluation reward: 3.97
 episode: 2711 score: 4.0 memory length: 586266 epsilon:
 0.037191340008541796 steps: 259 lr: 0.0001 evaluation reward: 3.98
 episode: 2712 score: 3.0 memory length: 586494 epsilon:
 0.036739900008542103 steps: 228 lr: 0.0001 evaluation reward: 3.98
 episode: 2713 score: 6.0 memory length: 586847 epsilon:
 0.03604096000854258 steps: 353 lr: 0.0001 evaluation reward: 4.0
 episode: 2714 score: 4.0 memory length: 587120 epsilon:
 0.03550042000854295 steps: 273 lr: 0.0001 evaluation reward: 4.01
 episode: 2715 score: 4.0 memory length: 587379 epsilon: 0.0349876000085433
 steps: 259 lr: 0.0001 evaluation reward: 4.02

episode: 2716	score: 7.0	memory length: 587762	epsilon:
0.034229260008543816	steps: 383	lr: 0.0001	evaluation reward: 4.05
episode: 2717	score: 3.0	memory length: 587975	epsilon:
0.033807520008544104	steps: 213	lr: 0.0001	evaluation reward: 4.04
episode: 2718	score: 3.0	memory length: 588204	epsilon:
0.03335410000854441	steps: 229	lr: 0.0001	evaluation reward: 4.03
episode: 2719	score: 4.0	memory length: 588444	epsilon:
0.03287890000854474	steps: 240	lr: 0.0001	evaluation reward: 4.01
episode: 2720	score: 3.0	memory length: 588654	epsilon:
0.03246310000854502	steps: 210	lr: 0.0001	evaluation reward: 4.0
episode: 2721	score: 6.0	memory length: 588992	epsilon:
0.03179386000854548	steps: 338	lr: 0.0001	evaluation reward: 4.01
episode: 2722	score: 3.0	memory length: 589205	epsilon:
0.031372120008545765	steps: 213	lr: 0.0001	evaluation reward: 4.01
episode: 2723	score: 3.0	memory length: 589418	epsilon:
0.030950380008546052	steps: 213	lr: 0.0001	evaluation reward: 4.0
episode: 2724	score: 5.0	memory length: 589707	epsilon:
0.030378160008546443	steps: 289	lr: 0.0001	evaluation reward: 4.02
episode: 2725	score: 3.0	memory length: 589918	epsilon:
0.029960380008546728	steps: 211	lr: 0.0001	evaluation reward: 3.98
episode: 2726	score: 3.0	memory length: 590128	epsilon:
0.02954458000854701	steps: 210	lr: 0.0001	evaluation reward: 3.98
episode: 2727	score: 3.0	memory length: 590341	epsilon:
steps: 213	lr: 0.0001	evaluation reward: 3.99	epsilon: 0.0291228400085473
episode: 2728	score: 4.0	memory length: 590618	epsilon:
0.028574380008547673	steps: 277	lr: 0.0001	evaluation reward: 4.0
episode: 2729	score: 4.0	memory length: 590878	epsilon:
0.028059580008548024	steps: 260	lr: 0.0001	evaluation reward: 4.0
episode: 2730	score: 3.0	memory length: 591091	epsilon:
0.02763784000854831	steps: 213	lr: 0.0001	evaluation reward: 4.0
episode: 2731	score: 3.0	memory length: 591301	epsilon:
0.027222040008548595	steps: 210	lr: 0.0001	evaluation reward: 3.98
episode: 2732	score: 3.0	memory length: 591530	epsilon:
0.026768620008548905	steps: 229	lr: 0.0001	evaluation reward: 3.94
episode: 2733	score: 7.0	memory length: 591933	epsilon:
0.02597068000854945	steps: 403	lr: 0.0001	evaluation reward: 3.97
episode: 2734	score: 6.0	memory length: 592261	epsilon:
0.025321240008549892	steps: 328	lr: 0.0001	evaluation reward: 4.01
episode: 2735	score: 4.0	memory length: 592518	epsilon:
0.02481238000855024	steps: 257	lr: 0.0001	evaluation reward: 4.01
episode: 2736	score: 3.0	memory length: 592767	epsilon:
0.024319360008550575	steps: 249	lr: 0.0001	evaluation reward: 4.0
episode: 2737	score: 4.0	memory length: 593043	epsilon:
0.023772880008550948	steps: 276	lr: 0.0001	evaluation reward: 4.0
episode: 2738	score: 3.0	memory length: 593254	epsilon:
0.023355100008551233	steps: 211	lr: 0.0001	evaluation reward: 4.0
episode: 2739	score: 2.0	memory length: 593436	epsilon:
0.02299474000855148	steps: 182	lr: 0.0001	evaluation reward: 3.99

episode: 2740	score: 3.0	memory length: 593666	epsilon:
0.02253934000855179	steps: 230	lr: 0.0001	evaluation reward: 3.99
episode: 2741	score: 7.0	memory length: 594037	epsilon:
0.02180476000855229	steps: 371	lr: 0.0001	evaluation reward: 4.03
episode: 2742	score: 7.0	memory length: 594403	epsilon:
0.021080080008552785	steps: 366	lr: 0.0001	evaluation reward: 4.06
episode: 2743	score: 2.0	memory length: 594600	epsilon:
0.02069002000855305	steps: 197	lr: 0.0001	evaluation reward: 4.07
episode: 2744	score: 4.0	memory length: 594860	epsilon: 0.0201752200085534
steps: 260	lr: 0.0001	evaluation reward: 4.08	
episode: 2745	score: 5.0	memory length: 595149	epsilon:
0.01960300008553792	steps: 289	lr: 0.0001	evaluation reward: 4.09
episode: 2746	score: 3.0	memory length: 595358	epsilon:
0.019189180008554074	steps: 209	lr: 0.0001	evaluation reward: 4.09
episode: 2747	score: 4.0	memory length: 595618	epsilon:
0.018674380008554425	steps: 260	lr: 0.0001	evaluation reward: 4.1
episode: 2748	score: 3.0	memory length: 595848	epsilon:
0.018218980008554736	steps: 230	lr: 0.0001	evaluation reward: 4.08
episode: 2749	score: 3.0	memory length: 596061	epsilon:
0.017797240008555024	steps: 213	lr: 0.0001	evaluation reward: 4.02
episode: 2750	score: 4.0	memory length: 596338	epsilon:
0.017248780008555398	steps: 277	lr: 0.0001	evaluation reward: 4.04
episode: 2751	score: 2.0	memory length: 596520	epsilon:
0.016888420008555644	steps: 182	lr: 0.0001	evaluation reward: 4.03
episode: 2752	score: 5.0	memory length: 596823	epsilon:
0.016288480008556053	steps: 303	lr: 0.0001	evaluation reward: 4.03
episode: 2753	score: 4.0	memory length: 597061	epsilon:
0.015817240008556374	steps: 238	lr: 0.0001	evaluation reward: 4.03
episode: 2754	score: 5.0	memory length: 597355	epsilon:
0.01523512000855643	steps: 294	lr: 0.0001	evaluation reward: 4.03
episode: 2755	score: 3.0	memory length: 597585	epsilon:
0.014779720008556341	steps: 230	lr: 0.0001	evaluation reward: 3.99
episode: 2756	score: 3.0	memory length: 597814	epsilon:
0.014326300008556253	steps: 229	lr: 0.0001	evaluation reward: 3.99
episode: 2757	score: 6.0	memory length: 598189	epsilon:
0.01358380008556109	steps: 375	lr: 0.0001	evaluation reward: 3.97
episode: 2758	score: 3.0	memory length: 598399	epsilon:
0.01316800008556028	steps: 210	lr: 0.0001	evaluation reward: 3.96
episode: 2759	score: 3.0	memory length: 598609	epsilon:
0.01275220008555948	steps: 210	lr: 0.0001	evaluation reward: 3.96
episode: 2760	score: 6.0	memory length: 598952	epsilon:
0.012073060008555816	steps: 343	lr: 0.0001	evaluation reward: 3.99
episode: 2761	score: 8.0	memory length: 599426	epsilon:
0.011134540008555634	steps: 474	lr: 0.0001	evaluation reward: 4.04
episode: 2762	score: 5.0	memory length: 599722	epsilon:
0.01054846000855552	steps: 296	lr: 0.0001	evaluation reward: 4.06
episode: 2763	score: 4.0	memory length: 599982	epsilon:
0.01003366000855542	steps: 260	lr: 0.0001	evaluation reward: 4.08

episode: 2764	score: 6.0	memory length: 600344	epsilon:
0.009998020008555413	steps: 362	lr: 0.0001	evaluation reward: 4.11
episode: 2765	score: 2.0	memory length: 600526	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 4.06
episode: 2766	score: 4.0	memory length: 600784	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 4.06
episode: 2767	score: 4.0	memory length: 601043	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.05
episode: 2768	score: 4.0	memory length: 601287	epsilon:
0.009998020008555413	steps: 244	lr: 0.0001	evaluation reward: 4.06
episode: 2769	score: 6.0	memory length: 601660	epsilon:
0.009998020008555413	steps: 373	lr: 0.0001	evaluation reward: 4.06
episode: 2770	score: 3.0	memory length: 601906	epsilon:
0.009998020008555413	steps: 246	lr: 0.0001	evaluation reward: 4.06
episode: 2771	score: 4.0	memory length: 602164	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 4.07
episode: 2772	score: 6.0	memory length: 602504	epsilon:
0.009998020008555413	steps: 340	lr: 0.0001	evaluation reward: 4.08
episode: 2773	score: 5.0	memory length: 602815	epsilon:
0.009998020008555413	steps: 311	lr: 0.0001	evaluation reward: 4.1
episode: 2774	score: 7.0	memory length: 603192	epsilon:
0.009998020008555413	steps: 377	lr: 0.0001	evaluation reward: 4.13
episode: 2775	score: 4.0	memory length: 603449	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 4.08
episode: 2776	score: 5.0	memory length: 603758	epsilon:
0.009998020008555413	steps: 309	lr: 0.0001	evaluation reward: 4.08
episode: 2777	score: 4.0	memory length: 604018	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.07
episode: 2778	score: 5.0	memory length: 604328	epsilon:
0.009998020008555413	steps: 310	lr: 0.0001	evaluation reward: 4.08
episode: 2779	score: 4.0	memory length: 604603	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 4.08
episode: 2780	score: 4.0	memory length: 604863	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.08
episode: 2781	score: 8.0	memory length: 605320	epsilon:
0.009998020008555413	steps: 457	lr: 0.0001	evaluation reward: 4.12
episode: 2782	score: 5.0	memory length: 605648	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 4.15
episode: 2783	score: 4.0	memory length: 605908	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.13
episode: 2784	score: 4.0	memory length: 606184	epsilon:
0.009998020008555413	steps: 276	lr: 0.0001	evaluation reward: 4.1
episode: 2785	score: 3.0	memory length: 606411	epsilon:
0.009998020008555413	steps: 227	lr: 0.0001	evaluation reward: 4.09
episode: 2786	score: 3.0	memory length: 606641	epsilon:
0.009998020008555413	steps: 230	lr: 0.0001	evaluation reward: 4.08
episode: 2787	score: 3.0	memory length: 606871	epsilon:
0.009998020008555413	steps: 230	lr: 0.0001	evaluation reward: 4.08

episode: 2788	score: 5.0	memory length: 607196	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 4.1
episode: 2789	score: 6.0	memory length: 607552	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 4.13
episode: 2790	score: 9.0	memory length: 608040	epsilon:
0.009998020008555413	steps: 488	lr: 0.0001	evaluation reward: 4.18
episode: 2791	score: 5.0	memory length: 608332	epsilon:
0.009998020008555413	steps: 292	lr: 0.0001	evaluation reward: 4.2
episode: 2792	score: 4.0	memory length: 608592	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.21
episode: 2793	score: 4.0	memory length: 608867	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 4.22
episode: 2794	score: 5.0	memory length: 609175	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 4.24
episode: 2795	score: 4.0	memory length: 609437	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 4.24
episode: 2796	score: 3.0	memory length: 609650	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 4.24
episode: 2797	score: 3.0	memory length: 609863	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 4.22
episode: 2798	score: 5.0	memory length: 610154	epsilon:
0.009998020008555413	steps: 291	lr: 0.0001	evaluation reward: 4.24
episode: 2799	score: 4.0	memory length: 610413	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.23
episode: 2800	score: 7.0	memory length: 610783	epsilon:
0.009998020008555413	steps: 370	lr: 0.0001	evaluation reward: 4.26
episode: 2801	score: 4.0	memory length: 611045	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 4.26
episode: 2802	score: 5.0	memory length: 611371	epsilon:
0.009998020008555413	steps: 326	lr: 0.0001	evaluation reward: 4.28
episode: 2803	score: 3.0	memory length: 611600	epsilon:
0.009998020008555413	steps: 229	lr: 0.0001	evaluation reward: 4.27
episode: 2804	score: 3.0	memory length: 611812	epsilon:
0.009998020008555413	steps: 212	lr: 0.0001	evaluation reward: 4.27
episode: 2805	score: 2.0	memory length: 611994	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 4.25
episode: 2806	score: 3.0	memory length: 612207	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 4.25
episode: 2807	score: 4.0	memory length: 612485	epsilon:
0.009998020008555413	steps: 278	lr: 0.0001	evaluation reward: 4.26
episode: 2808	score: 7.0	memory length: 612843	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 4.26
episode: 2809	score: 7.0	memory length: 613213	epsilon:
0.009998020008555413	steps: 370	lr: 0.0001	evaluation reward: 4.28
episode: 2810	score: 4.0	memory length: 613471	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 4.28
episode: 2811	score: 6.0	memory length: 613823	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 4.3

episode: 2812	score: 4.0	memory length: 614082	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.31
episode: 2813	score: 5.0	memory length: 614390	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 4.3
episode: 2814	score: 3.0	memory length: 614616	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 4.29
episode: 2815	score: 3.0	memory length: 614844	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 4.28
episode: 2816	score: 8.0	memory length: 615269	epsilon:
0.009998020008555413	steps: 425	lr: 0.0001	evaluation reward: 4.29
episode: 2817	score: 8.0	memory length: 615678	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 4.34
episode: 2818	score: 5.0	memory length: 616026	epsilon:
0.009998020008555413	steps: 348	lr: 0.0001	evaluation reward: 4.36
episode: 2819	score: 4.0	memory length: 616301	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 4.36
episode: 2820	score: 6.0	memory length: 616628	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 4.39
episode: 2821	score: 4.0	memory length: 616904	epsilon:
0.009998020008555413	steps: 276	lr: 0.0001	evaluation reward: 4.37
episode: 2822	score: 3.0	memory length: 617132	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 4.37
episode: 2823	score: 3.0	memory length: 617360	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 4.37
episode: 2824	score: 6.0	memory length: 617699	epsilon:
0.009998020008555413	steps: 339	lr: 0.0001	evaluation reward: 4.38
episode: 2825	score: 4.0	memory length: 617956	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 4.39
episode: 2826	score: 4.0	memory length: 618215	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.4
episode: 2827	score: 4.0	memory length: 618493	epsilon:
0.009998020008555413	steps: 278	lr: 0.0001	evaluation reward: 4.41
episode: 2828	score: 6.0	memory length: 618857	epsilon:
0.009998020008555413	steps: 364	lr: 0.0001	evaluation reward: 4.43
episode: 2829	score: 5.0	memory length: 619138	epsilon:
0.009998020008555413	steps: 281	lr: 0.0001	evaluation reward: 4.44
episode: 2830	score: 3.0	memory length: 619366	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 4.44
episode: 2831	score: 4.0	memory length: 619641	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 4.45
episode: 2832	score: 4.0	memory length: 619916	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 4.46
episode: 2833	score: 4.0	memory length: 620175	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.43
episode: 2834	score: 5.0	memory length: 620483	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 4.42
episode: 2835	score: 3.0	memory length: 620693	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 4.41

episode: 2836	score: 3.0	memory length: 620941	epsilon:
0.009998020008555413	steps: 248	lr: 0.0001	evaluation reward: 4.41
episode: 2837	score: 3.0	memory length: 621151	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 4.4
episode: 2838	score: 3.0	memory length: 621381	epsilon:
0.009998020008555413	steps: 230	lr: 0.0001	evaluation reward: 4.4
episode: 2839	score: 5.0	memory length: 621693	epsilon:
0.009998020008555413	steps: 312	lr: 0.0001	evaluation reward: 4.43
episode: 2840	score: 3.0	memory length: 621906	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 4.43
episode: 2841	score: 9.0	memory length: 622355	epsilon:
0.009998020008555413	steps: 449	lr: 0.0001	evaluation reward: 4.45
episode: 2842	score: 3.0	memory length: 622581	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 4.41
episode: 2843	score: 5.0	memory length: 622909	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 4.44
episode: 2844	score: 7.0	memory length: 623282	epsilon:
0.009998020008555413	steps: 373	lr: 0.0001	evaluation reward: 4.47
episode: 2845	score: 5.0	memory length: 623591	epsilon:
0.009998020008555413	steps: 309	lr: 0.0001	evaluation reward: 4.47
episode: 2846	score: 5.0	memory length: 623867	epsilon:
0.009998020008555413	steps: 276	lr: 0.0001	evaluation reward: 4.49
episode: 2847	score: 4.0	memory length: 624144	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 4.49
episode: 2848	score: 5.0	memory length: 624468	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 4.51
episode: 2849	score: 4.0	memory length: 624727	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.52
episode: 2850	score: 4.0	memory length: 625004	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 4.52
episode: 2851	score: 4.0	memory length: 625264	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.54
episode: 2852	score: 5.0	memory length: 625572	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 4.54
episode: 2853	score: 5.0	memory length: 625893	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 4.55
episode: 2854	score: 6.0	memory length: 626227	epsilon:
0.009998020008555413	steps: 334	lr: 0.0001	evaluation reward: 4.56
episode: 2855	score: 4.0	memory length: 626507	epsilon:
0.009998020008555413	steps: 280	lr: 0.0001	evaluation reward: 4.57
episode: 2856	score: 3.0	memory length: 626717	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 4.57
episode: 2857	score: 4.0	memory length: 626974	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 4.55
episode: 2858	score: 3.0	memory length: 627187	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 4.55
episode: 2859	score: 4.0	memory length: 627444	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 4.56

episode: 2860	score: 3.0	memory length: 627672	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 4.53
episode: 2861	score: 5.0	memory length: 627993	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 4.5
episode: 2862	score: 6.0	memory length: 628349	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 4.51
episode: 2863	score: 4.0	memory length: 628627	epsilon:
0.009998020008555413	steps: 278	lr: 0.0001	evaluation reward: 4.51
episode: 2864	score: 7.0	memory length: 629032	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 4.52
episode: 2865	score: 3.0	memory length: 629245	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 4.53
episode: 2866	score: 3.0	memory length: 629470	epsilon:
0.009998020008555413	steps: 225	lr: 0.0001	evaluation reward: 4.52
episode: 2867	score: 5.0	memory length: 629813	epsilon:
0.009998020008555413	steps: 343	lr: 0.0001	evaluation reward: 4.53
episode: 2868	score: 8.0	memory length: 630236	epsilon:
0.009998020008555413	steps: 423	lr: 0.0001	evaluation reward: 4.57
episode: 2869	score: 4.0	memory length: 630494	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 4.55
episode: 2870	score: 4.0	memory length: 630754	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.56
episode: 2871	score: 5.0	memory length: 631079	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 4.57
episode: 2872	score: 4.0	memory length: 631356	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 4.55
episode: 2873	score: 4.0	memory length: 631595	epsilon:
0.009998020008555413	steps: 239	lr: 0.0001	evaluation reward: 4.54
episode: 2874	score: 4.0	memory length: 631835	epsilon:
0.009998020008555413	steps: 240	lr: 0.0001	evaluation reward: 4.51
episode: 2875	score: 8.0	memory length: 632243	epsilon:
0.009998020008555413	steps: 408	lr: 0.0001	evaluation reward: 4.55
episode: 2876	score: 4.0	memory length: 632524	epsilon:
0.009998020008555413	steps: 281	lr: 0.0001	evaluation reward: 4.54
episode: 2877	score: 4.0	memory length: 632763	epsilon:
0.009998020008555413	steps: 239	lr: 0.0001	evaluation reward: 4.54
episode: 2878	score: 4.0	memory length: 633022	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.53
episode: 2879	score: 6.0	memory length: 633350	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 4.55
episode: 2880	score: 4.0	memory length: 633610	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.55
episode: 2881	score: 5.0	memory length: 633899	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 4.52
episode: 2882	score: 5.0	memory length: 634188	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 4.52
episode: 2883	score: 5.0	memory length: 634492	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 4.53

episode: 2884	score: 8.0	memory length: 634878	epsilon:
0.009998020008555413	steps: 386	lr: 0.0001	evaluation reward: 4.57
episode: 2885	score: 5.0	memory length: 635150	epsilon:
0.009998020008555413	steps: 272	lr: 0.0001	evaluation reward: 4.59
episode: 2886	score: 3.0	memory length: 635380	epsilon:
0.009998020008555413	steps: 230	lr: 0.0001	evaluation reward: 4.59
episode: 2887	score: 6.0	memory length: 635717	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 4.62
episode: 2888	score: 9.0	memory length: 636220	epsilon:
0.009998020008555413	steps: 503	lr: 0.0001	evaluation reward: 4.66
episode: 2889	score: 3.0	memory length: 636465	epsilon:
0.009998020008555413	steps: 245	lr: 0.0001	evaluation reward: 4.63
episode: 2890	score: 6.0	memory length: 636820	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 4.6
episode: 2891	score: 4.0	memory length: 637058	epsilon:
0.009998020008555413	steps: 238	lr: 0.0001	evaluation reward: 4.59
episode: 2892	score: 5.0	memory length: 637344	epsilon:
0.009998020008555413	steps: 286	lr: 0.0001	evaluation reward: 4.6
episode: 2893	score: 3.0	memory length: 637553	epsilon:
0.009998020008555413	steps: 209	lr: 0.0001	evaluation reward: 4.59
episode: 2894	score: 5.0	memory length: 637841	epsilon:
0.009998020008555413	steps: 288	lr: 0.0001	evaluation reward: 4.59
episode: 2895	score: 4.0	memory length: 638098	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 4.59
episode: 2896	score: 8.0	memory length: 638498	epsilon:
0.009998020008555413	steps: 400	lr: 0.0001	evaluation reward: 4.64
episode: 2897	score: 3.0	memory length: 638725	epsilon:
0.009998020008555413	steps: 227	lr: 0.0001	evaluation reward: 4.64
episode: 2898	score: 3.0	memory length: 638938	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 4.62
episode: 2899	score: 9.0	memory length: 639267	epsilon:
0.009998020008555413	steps: 329	lr: 0.0001	evaluation reward: 4.67
episode: 2900	score: 4.0	memory length: 639507	epsilon:
0.009998020008555413	steps: 240	lr: 0.0001	evaluation reward: 4.64
episode: 2901	score: 6.0	memory length: 639860	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 4.66
episode: 2902	score: 6.0	memory length: 640235	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 4.67
episode: 2903	score: 6.0	memory length: 640610	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 4.7
episode: 2904	score: 5.0	memory length: 640957	epsilon:
0.009998020008555413	steps: 347	lr: 0.0001	evaluation reward: 4.72
episode: 2905	score: 8.0	memory length: 641393	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 4.78
episode: 2906	score: 5.0	memory length: 641683	epsilon:
0.009998020008555413	steps: 290	lr: 0.0001	evaluation reward: 4.8
episode: 2907	score: 4.0	memory length: 641962	epsilon:
0.009998020008555413	steps: 279	lr: 0.0001	evaluation reward: 4.8

episode: 2908	score: 7.0	memory length: 642369	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 4.8
episode: 2909	score: 5.0	memory length: 642672	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 4.78
episode: 2910	score: 4.0	memory length: 642947	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 4.78
episode: 2911	score: 5.0	memory length: 643271	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 4.77
episode: 2912	score: 7.0	memory length: 643656	epsilon:
0.009998020008555413	steps: 385	lr: 0.0001	evaluation reward: 4.8
episode: 2913	score: 10.0	memory length: 644032	epsilon:
0.009998020008555413	steps: 376	lr: 0.0001	evaluation reward: 4.85
episode: 2914	score: 4.0	memory length: 644291	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 4.86
episode: 2915	score: 2.0	memory length: 644489	epsilon:
0.009998020008555413	steps: 198	lr: 0.0001	evaluation reward: 4.85
episode: 2916	score: 7.0	memory length: 644895	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 4.84
episode: 2917	score: 3.0	memory length: 645124	epsilon:
0.009998020008555413	steps: 229	lr: 0.0001	evaluation reward: 4.79
episode: 2918	score: 7.0	memory length: 645517	epsilon:
0.009998020008555413	steps: 393	lr: 0.0001	evaluation reward: 4.81
episode: 2919	score: 8.0	memory length: 645907	epsilon:
0.009998020008555413	steps: 390	lr: 0.0001	evaluation reward: 4.85
episode: 2920	score: 4.0	memory length: 646167	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.83
episode: 2921	score: 5.0	memory length: 646474	epsilon:
0.009998020008555413	steps: 307	lr: 0.0001	evaluation reward: 4.84
episode: 2922	score: 6.0	memory length: 646811	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 4.87
episode: 2923	score: 7.0	memory length: 647234	epsilon:
0.009998020008555413	steps: 423	lr: 0.0001	evaluation reward: 4.91
episode: 2924	score: 6.0	memory length: 647602	epsilon:
0.009998020008555413	steps: 368	lr: 0.0001	evaluation reward: 4.91
episode: 2925	score: 3.0	memory length: 647831	epsilon:
0.009998020008555413	steps: 229	lr: 0.0001	evaluation reward: 4.9
episode: 2926	score: 3.0	memory length: 648059	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 4.89
episode: 2927	score: 5.0	memory length: 648332	epsilon:
0.009998020008555413	steps: 273	lr: 0.0001	evaluation reward: 4.9
episode: 2928	score: 3.0	memory length: 648564	epsilon:
0.009998020008555413	steps: 232	lr: 0.0001	evaluation reward: 4.87
episode: 2929	score: 5.0	memory length: 648845	epsilon:
0.009998020008555413	steps: 281	lr: 0.0001	evaluation reward: 4.87
episode: 2930	score: 6.0	memory length: 649205	epsilon:
0.009998020008555413	steps: 360	lr: 0.0001	evaluation reward: 4.9
episode: 2931	score: 4.0	memory length: 649465	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 4.9

episode: 2932	score: 8.0	memory length: 649860	epsilon:
0.009998020008555413	steps: 395	lr: 0.0001	evaluation reward: 4.94
episode: 2933	score: 4.0	memory length: 650140	epsilon:
0.009998020008555413	steps: 280	lr: 0.0001	evaluation reward: 4.94
episode: 2934	score: 6.0	memory length: 650498	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 4.95
episode: 2935	score: 3.0	memory length: 650727	epsilon:
0.009998020008555413	steps: 229	lr: 0.0001	evaluation reward: 4.95
episode: 2936	score: 6.0	memory length: 651061	epsilon:
0.009998020008555413	steps: 334	lr: 0.0001	evaluation reward: 4.98
episode: 2937	score: 8.0	memory length: 651483	epsilon:
0.009998020008555413	steps: 422	lr: 0.0001	evaluation reward: 5.03
episode: 2938	score: 4.0	memory length: 651759	epsilon:
0.009998020008555413	steps: 276	lr: 0.0001	evaluation reward: 5.04
episode: 2939	score: 4.0	memory length: 652018	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 5.03
episode: 2940	score: 7.0	memory length: 652419	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 5.07
episode: 2941	score: 3.0	memory length: 652664	epsilon:
0.009998020008555413	steps: 245	lr: 0.0001	evaluation reward: 5.01
episode: 2942	score: 7.0	memory length: 653022	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 5.05
episode: 2943	score: 5.0	memory length: 653344	epsilon:
0.009998020008555413	steps: 322	lr: 0.0001	evaluation reward: 5.05
episode: 2944	score: 8.0	memory length: 653800	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 5.06
episode: 2945	score: 4.0	memory length: 654078	epsilon:
0.009998020008555413	steps: 278	lr: 0.0001	evaluation reward: 5.05
episode: 2946	score: 6.0	memory length: 654406	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 5.06
episode: 2947	score: 3.0	memory length: 654653	epsilon:
0.009998020008555413	steps: 247	lr: 0.0001	evaluation reward: 5.05
episode: 2948	score: 4.0	memory length: 654930	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 5.04
episode: 2949	score: 4.0	memory length: 655212	epsilon:
0.009998020008555413	steps: 282	lr: 0.0001	evaluation reward: 5.04
episode: 2950	score: 5.0	memory length: 655517	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 5.05
episode: 2951	score: 6.0	memory length: 655892	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 5.07
episode: 2952	score: 8.0	memory length: 656327	epsilon:
0.009998020008555413	steps: 435	lr: 0.0001	evaluation reward: 5.1
episode: 2953	score: 6.0	memory length: 656641	epsilon:
0.009998020008555413	steps: 314	lr: 0.0001	evaluation reward: 5.11
episode: 2954	score: 3.0	memory length: 656871	epsilon:
0.009998020008555413	steps: 230	lr: 0.0001	evaluation reward: 5.08
episode: 2955	score: 5.0	memory length: 657178	epsilon:
0.009998020008555413	steps: 307	lr: 0.0001	evaluation reward: 5.09

episode: 2956	score: 2.0	memory length: 657359	epsilon:
0.009998020008555413	steps: 181	lr: 0.0001	evaluation reward: 5.08
episode: 2957	score: 7.0	memory length: 657737	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 5.11
episode: 2958	score: 4.0	memory length: 658013	epsilon:
0.009998020008555413	steps: 276	lr: 0.0001	evaluation reward: 5.12
episode: 2959	score: 3.0	memory length: 658241	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 5.11
episode: 2960	score: 5.0	memory length: 658544	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 5.13
episode: 2961	score: 6.0	memory length: 658884	epsilon:
0.009998020008555413	steps: 340	lr: 0.0001	evaluation reward: 5.14
episode: 2962	score: 7.0	memory length: 659294	epsilon:
0.009998020008555413	steps: 410	lr: 0.0001	evaluation reward: 5.15
episode: 2963	score: 5.0	memory length: 659583	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 5.16
episode: 2964	score: 5.0	memory length: 659887	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 5.14
episode: 2965	score: 14.0	memory length: 660560	epsilon:
0.009998020008555413	steps: 673	lr: 0.0001	evaluation reward: 5.25
episode: 2966	score: 8.0	memory length: 660984	epsilon:
0.009998020008555413	steps: 424	lr: 0.0001	evaluation reward: 5.3
episode: 2967	score: 4.0	memory length: 661240	epsilon:
0.009998020008555413	steps: 256	lr: 0.0001	evaluation reward: 5.29
episode: 2968	score: 5.0	memory length: 661566	epsilon:
0.009998020008555413	steps: 326	lr: 0.0001	evaluation reward: 5.26
episode: 2969	score: 10.0	memory length: 662112	epsilon:
0.009998020008555413	steps: 546	lr: 0.0001	evaluation reward: 5.32
episode: 2970	score: 5.0	memory length: 662437	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 5.33
episode: 2971	score: 3.0	memory length: 662685	epsilon:
0.009998020008555413	steps: 248	lr: 0.0001	evaluation reward: 5.31
episode: 2972	score: 7.0	memory length: 663102	epsilon:
0.009998020008555413	steps: 417	lr: 0.0001	evaluation reward: 5.34
episode: 2973	score: 4.0	memory length: 663359	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 5.34
episode: 2974	score: 3.0	memory length: 663569	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 5.33
episode: 2975	score: 3.0	memory length: 663798	epsilon:
0.009998020008555413	steps: 229	lr: 0.0001	evaluation reward: 5.28
episode: 2976	score: 6.0	memory length: 664119	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 5.3
episode: 2977	score: 11.0	memory length: 664690	epsilon:
0.009998020008555413	steps: 571	lr: 0.0001	evaluation reward: 5.37
episode: 2978	score: 4.0	memory length: 664929	epsilon:
0.009998020008555413	steps: 239	lr: 0.0001	evaluation reward: 5.37
episode: 2979	score: 4.0	memory length: 665207	epsilon:
0.009998020008555413	steps: 278	lr: 0.0001	evaluation reward: 5.35

episode: 2980	score: 6.0	memory length: 665560	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 5.37
episode: 2981	score: 4.0	memory length: 665860	epsilon:
0.009998020008555413	steps: 300	lr: 0.0001	evaluation reward: 5.36
episode: 2982	score: 7.0	memory length: 666223	epsilon:
0.009998020008555413	steps: 363	lr: 0.0001	evaluation reward: 5.38
episode: 2983	score: 6.0	memory length: 666579	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 5.39
episode: 2984	score: 6.0	memory length: 666959	epsilon:
0.009998020008555413	steps: 380	lr: 0.0001	evaluation reward: 5.37
episode: 2985	score: 7.0	memory length: 667312	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 5.39
episode: 2986	score: 3.0	memory length: 667542	epsilon:
0.009998020008555413	steps: 230	lr: 0.0001	evaluation reward: 5.39
episode: 2987	score: 6.0	memory length: 667894	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 5.39
episode: 2988	score: 13.0	memory length: 668375	epsilon:
0.009998020008555413	steps: 481	lr: 0.0001	evaluation reward: 5.43
episode: 2989	score: 3.0	memory length: 668585	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 5.43
episode: 2990	score: 7.0	memory length: 668963	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 5.44
episode: 2991	score: 7.0	memory length: 669339	epsilon:
0.009998020008555413	steps: 376	lr: 0.0001	evaluation reward: 5.47
episode: 2992	score: 4.0	memory length: 669616	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 5.46
episode: 2993	score: 4.0	memory length: 669876	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 5.47
episode: 2994	score: 8.0	memory length: 670329	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 5.5
episode: 2995	score: 4.0	memory length: 670604	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 5.5
episode: 2996	score: 4.0	memory length: 670862	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 5.46
episode: 2997	score: 6.0	memory length: 671237	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 5.49
episode: 2998	score: 4.0	memory length: 671512	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 5.5
episode: 2999	score: 6.0	memory length: 671831	epsilon:
0.009998020008555413	steps: 319	lr: 0.0001	evaluation reward: 5.47
episode: 3000	score: 5.0	memory length: 672137	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 5.48
episode: 3001	score: 4.0	memory length: 672433	epsilon:
0.009998020008555413	steps: 296	lr: 0.0001	evaluation reward: 5.46
episode: 3002	score: 8.0	memory length: 672873	epsilon:
0.009998020008555413	steps: 440	lr: 0.0001	evaluation reward: 5.48
episode: 3003	score: 5.0	memory length: 673181	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 5.47

episode: 3004	score: 4.0	memory length: 673458	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 5.46
episode: 3005	score: 4.0	memory length: 673718	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 5.42
episode: 3006	score: 4.0	memory length: 673995	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 5.41
episode: 3007	score: 3.0	memory length: 674223	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 5.4
episode: 3008	score: 6.0	memory length: 674582	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 5.39
episode: 3009	score: 9.0	memory length: 674995	epsilon:
0.009998020008555413	steps: 413	lr: 0.0001	evaluation reward: 5.43
episode: 3010	score: 4.0	memory length: 675237	epsilon:
0.009998020008555413	steps: 242	lr: 0.0001	evaluation reward: 5.43
episode: 3011	score: 5.0	memory length: 675507	epsilon:
0.009998020008555413	steps: 270	lr: 0.0001	evaluation reward: 5.43
episode: 3012	score: 11.0	memory length: 675993	epsilon:
0.009998020008555413	steps: 486	lr: 0.0001	evaluation reward: 5.47
episode: 3013	score: 9.0	memory length: 676401	epsilon:
0.009998020008555413	steps: 408	lr: 0.0001	evaluation reward: 5.46
episode: 3014	score: 7.0	memory length: 676809	epsilon:
0.009998020008555413	steps: 408	lr: 0.0001	evaluation reward: 5.49
episode: 3015	score: 4.0	memory length: 677071	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 5.51
episode: 3016	score: 4.0	memory length: 677353	epsilon:
0.009998020008555413	steps: 282	lr: 0.0001	evaluation reward: 5.48
episode: 3017	score: 7.0	memory length: 677724	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 5.52
episode: 3018	score: 6.0	memory length: 678059	epsilon:
0.009998020008555413	steps: 335	lr: 0.0001	evaluation reward: 5.51
episode: 3019	score: 5.0	memory length: 678365	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 5.48
episode: 3020	score: 6.0	memory length: 678685	epsilon:
0.009998020008555413	steps: 320	lr: 0.0001	evaluation reward: 5.5
episode: 3021	score: 6.0	memory length: 679059	epsilon:
0.009998020008555413	steps: 374	lr: 0.0001	evaluation reward: 5.51
episode: 3022	score: 8.0	memory length: 679502	epsilon:
0.009998020008555413	steps: 443	lr: 0.0001	evaluation reward: 5.53
episode: 3023	score: 11.0	memory length: 680027	epsilon:
0.009998020008555413	steps: 525	lr: 0.0001	evaluation reward: 5.57
episode: 3024	score: 7.0	memory length: 680402	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 5.58
episode: 3025	score: 4.0	memory length: 680702	epsilon:
0.009998020008555413	steps: 300	lr: 0.0001	evaluation reward: 5.59
episode: 3026	score: 6.0	memory length: 681060	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 5.62
episode: 3027	score: 9.0	memory length: 681516	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 5.66

episode: 3028	score: 5.0	memory length: 681804	epsilon:
0.009998020008555413	steps: 288	lr: 0.0001	evaluation reward: 5.68
episode: 3029	score: 6.0	memory length: 682114	epsilon:
0.009998020008555413	steps: 310	lr: 0.0001	evaluation reward: 5.69
episode: 3030	score: 8.0	memory length: 682552	epsilon:
0.009998020008555413	steps: 438	lr: 0.0001	evaluation reward: 5.71
episode: 3031	score: 8.0	memory length: 682988	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 5.75
episode: 3032	score: 8.0	memory length: 683424	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 5.75
episode: 3033	score: 3.0	memory length: 683650	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 5.74
episode: 3034	score: 11.0	memory length: 684081	epsilon:
0.009998020008555413	steps: 431	lr: 0.0001	evaluation reward: 5.79
episode: 3035	score: 6.0	memory length: 684459	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 5.82
episode: 3036	score: 5.0	memory length: 684785	epsilon:
0.009998020008555413	steps: 326	lr: 0.0001	evaluation reward: 5.81
episode: 3037	score: 6.0	memory length: 685119	epsilon:
0.009998020008555413	steps: 334	lr: 0.0001	evaluation reward: 5.79
episode: 3038	score: 4.0	memory length: 685363	epsilon:
0.009998020008555413	steps: 244	lr: 0.0001	evaluation reward: 5.79
episode: 3039	score: 3.0	memory length: 685573	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 5.78
episode: 3040	score: 6.0	memory length: 685949	epsilon:
0.009998020008555413	steps: 376	lr: 0.0001	evaluation reward: 5.77
episode: 3041	score: 5.0	memory length: 686240	epsilon:
0.009998020008555413	steps: 291	lr: 0.0001	evaluation reward: 5.79
episode: 3042	score: 3.0	memory length: 686471	epsilon:
0.009998020008555413	steps: 231	lr: 0.0001	evaluation reward: 5.75
episode: 3043	score: 9.0	memory length: 686954	epsilon:
0.009998020008555413	steps: 483	lr: 0.0001	evaluation reward: 5.79
episode: 3044	score: 4.0	memory length: 687214	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 5.75
episode: 3045	score: 6.0	memory length: 687565	epsilon:
0.009998020008555413	steps: 351	lr: 0.0001	evaluation reward: 5.77
episode: 3046	score: 7.0	memory length: 687954	epsilon:
0.009998020008555413	steps: 389	lr: 0.0001	evaluation reward: 5.78
episode: 3047	score: 7.0	memory length: 688355	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 5.82
episode: 3048	score: 9.0	memory length: 688813	epsilon:
0.009998020008555413	steps: 458	lr: 0.0001	evaluation reward: 5.87
episode: 3049	score: 5.0	memory length: 689137	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 5.88
episode: 3050	score: 4.0	memory length: 689412	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 5.87
episode: 3051	score: 7.0	memory length: 689817	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 5.88

episode: 3052	score: 5.0	memory length: 690121	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 5.85
episode: 3053	score: 7.0	memory length: 690547	epsilon:
0.009998020008555413	steps: 426	lr: 0.0001	evaluation reward: 5.86
episode: 3054	score: 7.0	memory length: 690971	epsilon:
0.009998020008555413	steps: 424	lr: 0.0001	evaluation reward: 5.9
episode: 3055	score: 6.0	memory length: 691318	epsilon:
0.009998020008555413	steps: 347	lr: 0.0001	evaluation reward: 5.91
episode: 3056	score: 8.0	memory length: 691726	epsilon:
0.009998020008555413	steps: 408	lr: 0.0001	evaluation reward: 5.97
episode: 3057	score: 6.0	memory length: 692080	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 5.96
episode: 3058	score: 7.0	memory length: 692444	epsilon:
0.009998020008555413	steps: 364	lr: 0.0001	evaluation reward: 5.99
episode: 3059	score: 9.0	memory length: 692885	epsilon:
0.009998020008555413	steps: 441	lr: 0.0001	evaluation reward: 6.05
episode: 3060	score: 7.0	memory length: 693288	epsilon:
0.009998020008555413	steps: 403	lr: 0.0001	evaluation reward: 6.07
episode: 3061	score: 7.0	memory length: 693629	epsilon:
0.009998020008555413	steps: 341	lr: 0.0001	evaluation reward: 6.08
episode: 3062	score: 7.0	memory length: 693998	epsilon:
0.009998020008555413	steps: 369	lr: 0.0001	evaluation reward: 6.08
episode: 3063	score: 3.0	memory length: 694211	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 6.06
episode: 3064	score: 5.0	memory length: 694517	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 6.06
episode: 3065	score: 5.0	memory length: 694845	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 5.97
episode: 3066	score: 3.0	memory length: 695055	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 5.92
episode: 3067	score: 6.0	memory length: 695392	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 5.94
episode: 3068	score: 6.0	memory length: 695750	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 5.95
episode: 3069	score: 5.0	memory length: 696080	epsilon:
0.009998020008555413	steps: 330	lr: 0.0001	evaluation reward: 5.9
episode: 3070	score: 6.0	memory length: 696421	epsilon:
0.009998020008555413	steps: 341	lr: 0.0001	evaluation reward: 5.91
episode: 3071	score: 4.0	memory length: 696698	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 5.92
episode: 3072	score: 6.0	memory length: 697056	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 5.91
episode: 3073	score: 6.0	memory length: 697405	epsilon:
0.009998020008555413	steps: 349	lr: 0.0001	evaluation reward: 5.93
episode: 3074	score: 6.0	memory length: 697761	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 5.96
episode: 3075	score: 7.0	memory length: 698166	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 6.0

episode: 3076	score: 6.0	memory length: 698504	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 6.0
episode: 3077	score: 4.0	memory length: 698743	epsilon:
0.009998020008555413	steps: 239	lr: 0.0001	evaluation reward: 5.93
episode: 3078	score: 7.0	memory length: 699127	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 5.96
episode: 3079	score: 6.0	memory length: 699448	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 5.98
episode: 3080	score: 6.0	memory length: 699804	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 5.98
episode: 3081	score: 5.0	memory length: 700118	epsilon:
0.009998020008555413	steps: 314	lr: 0.0001	evaluation reward: 5.99
episode: 3082	score: 5.0	memory length: 700420	epsilon:
0.009998020008555413	steps: 302	lr: 0.0001	evaluation reward: 5.97
episode: 3083	score: 6.0	memory length: 700779	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 5.97
episode: 3084	score: 6.0	memory length: 701157	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 5.97
episode: 3085	score: 3.0	memory length: 701383	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 5.93
episode: 3086	score: 6.0	memory length: 701702	epsilon:
0.009998020008555413	steps: 319	lr: 0.0001	evaluation reward: 5.96
episode: 3087	score: 4.0	memory length: 701979	epsilon:
0.009998020008555413	steps: 277	lr: 0.0001	evaluation reward: 5.94
episode: 3088	score: 8.0	memory length: 702413	epsilon:
0.009998020008555413	steps: 434	lr: 0.0001	evaluation reward: 5.89
episode: 3089	score: 6.0	memory length: 702754	epsilon:
0.009998020008555413	steps: 341	lr: 0.0001	evaluation reward: 5.92
episode: 3090	score: 11.0	memory length: 703269	epsilon:
0.009998020008555413	steps: 515	lr: 0.0001	evaluation reward: 5.96
episode: 3091	score: 3.0	memory length: 703497	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 5.92
episode: 3092	score: 3.0	memory length: 703710	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 5.91
episode: 3093	score: 14.0	memory length: 704286	epsilon:
0.009998020008555413	steps: 576	lr: 0.0001	evaluation reward: 6.01
episode: 3094	score: 6.0	memory length: 704620	epsilon:
0.009998020008555413	steps: 334	lr: 0.0001	evaluation reward: 5.99
episode: 3095	score: 4.0	memory length: 704880	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 5.99
episode: 3096	score: 10.0	memory length: 705350	epsilon:
0.009998020008555413	steps: 470	lr: 0.0001	evaluation reward: 6.05
episode: 3097	score: 5.0	memory length: 705672	epsilon:
0.009998020008555413	steps: 322	lr: 0.0001	evaluation reward: 6.04
episode: 3098	score: 5.0	memory length: 705956	epsilon:
0.009998020008555413	steps: 284	lr: 0.0001	evaluation reward: 6.05
episode: 3099	score: 6.0	memory length: 706288	epsilon:
0.009998020008555413	steps: 332	lr: 0.0001	evaluation reward: 6.05

episode: 3100	score: 4.0	memory length: 706548	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 6.04
episode: 3101	score: 5.0	memory length: 706853	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 6.05
episode: 3102	score: 7.0	memory length: 707241	epsilon:
0.009998020008555413	steps: 388	lr: 0.0001	evaluation reward: 6.04
episode: 3103	score: 6.0	memory length: 707599	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 6.05
episode: 3104	score: 4.0	memory length: 707856	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 6.05
episode: 3105	score: 6.0	memory length: 708212	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 6.07
episode: 3106	score: 7.0	memory length: 708590	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 6.1
episode: 3107	score: 5.0	memory length: 708895	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 6.12
episode: 3108	score: 7.0	memory length: 709281	epsilon:
0.009998020008555413	steps: 386	lr: 0.0001	evaluation reward: 6.13
episode: 3109	score: 10.0	memory length: 709763	epsilon:
0.009998020008555413	steps: 482	lr: 0.0001	evaluation reward: 6.14
episode: 3110	score: 5.0	memory length: 710066	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 6.15
episode: 3111	score: 8.0	memory length: 710431	epsilon:
0.009998020008555413	steps: 365	lr: 0.0001	evaluation reward: 6.18
episode: 3112	score: 5.0	memory length: 710725	epsilon:
0.009998020008555413	steps: 294	lr: 0.0001	evaluation reward: 6.12
episode: 3113	score: 6.0	memory length: 711062	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 6.09
episode: 3114	score: 6.0	memory length: 711372	epsilon:
0.009998020008555413	steps: 310	lr: 0.0001	evaluation reward: 6.08
episode: 3115	score: 5.0	memory length: 711646	epsilon:
0.009998020008555413	steps: 274	lr: 0.0001	evaluation reward: 6.09
episode: 3116	score: 4.0	memory length: 711921	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 6.09
episode: 3117	score: 3.0	memory length: 712148	epsilon:
0.009998020008555413	steps: 227	lr: 0.0001	evaluation reward: 6.05
episode: 3118	score: 7.0	memory length: 712535	epsilon:
0.009998020008555413	steps: 387	lr: 0.0001	evaluation reward: 6.06
episode: 3119	score: 6.0	memory length: 712873	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 6.07
episode: 3120	score: 5.0	memory length: 713176	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 6.06
episode: 3121	score: 6.0	memory length: 713528	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 6.06
episode: 3122	score: 5.0	memory length: 713855	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 6.03
episode: 3123	score: 4.0	memory length: 714133	epsilon:
0.009998020008555413	steps: 278	lr: 0.0001	evaluation reward: 5.96

episode: 3124	score: 5.0	memory length: 714459	epsilon:
0.009998020008555413	steps: 326	lr: 0.0001	evaluation reward: 5.94
episode: 3125	score: 4.0	memory length: 714717	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 5.94
episode: 3126	score: 4.0	memory length: 714996	epsilon:
0.009998020008555413	steps: 279	lr: 0.0001	evaluation reward: 5.92
episode: 3127	score: 8.0	memory length: 715371	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 5.91
episode: 3128	score: 5.0	memory length: 715659	epsilon:
0.009998020008555413	steps: 288	lr: 0.0001	evaluation reward: 5.91
episode: 3129	score: 5.0	memory length: 715987	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 5.9
episode: 3130	score: 9.0	memory length: 716442	epsilon:
0.009998020008555413	steps: 455	lr: 0.0001	evaluation reward: 5.91
episode: 3131	score: 5.0	memory length: 716765	epsilon:
0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 5.88
episode: 3132	score: 12.0	memory length: 717367	epsilon:
0.009998020008555413	steps: 602	lr: 0.0001	evaluation reward: 5.92
episode: 3133	score: 10.0	memory length: 717869	epsilon:
0.009998020008555413	steps: 502	lr: 0.0001	evaluation reward: 5.99
episode: 3134	score: 6.0	memory length: 718219	epsilon:
0.009998020008555413	steps: 350	lr: 0.0001	evaluation reward: 5.94
episode: 3135	score: 4.0	memory length: 718519	epsilon:
0.009998020008555413	steps: 300	lr: 0.0001	evaluation reward: 5.92
episode: 3136	score: 2.0	memory length: 718701	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 5.89
episode: 3137	score: 5.0	memory length: 718991	epsilon:
0.009998020008555413	steps: 290	lr: 0.0001	evaluation reward: 5.88
episode: 3138	score: 4.0	memory length: 719233	epsilon:
0.009998020008555413	steps: 242	lr: 0.0001	evaluation reward: 5.88
episode: 3139	score: 8.0	memory length: 719666	epsilon:
0.009998020008555413	steps: 433	lr: 0.0001	evaluation reward: 5.93
episode: 3140	score: 8.0	memory length: 720097	epsilon:
0.009998020008555413	steps: 431	lr: 0.0001	evaluation reward: 5.95
episode: 3141	score: 4.0	memory length: 720372	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 5.94
episode: 3142	score: 3.0	memory length: 720603	epsilon:
0.009998020008555413	steps: 231	lr: 0.0001	evaluation reward: 5.94
episode: 3143	score: 10.0	memory length: 721086	epsilon:
0.009998020008555413	steps: 483	lr: 0.0001	evaluation reward: 5.95
episode: 3144	score: 6.0	memory length: 721445	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 5.97
episode: 3145	score: 4.0	memory length: 721705	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 5.95
episode: 3146	score: 3.0	memory length: 721918	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 5.91
episode: 3147	score: 9.0	memory length: 722369	epsilon:
0.009998020008555413	steps: 451	lr: 0.0001	evaluation reward: 5.93

episode: 3148	score: 6.0	memory length: 722692	epsilon:
0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 5.9
episode: 3149	score: 8.0	memory length: 723118	epsilon:
0.009998020008555413	steps: 426	lr: 0.0001	evaluation reward: 5.93
episode: 3150	score: 6.0	memory length: 723421	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 5.95
episode: 3151	score: 7.0	memory length: 723765	epsilon:
0.009998020008555413	steps: 344	lr: 0.0001	evaluation reward: 5.95
episode: 3152	score: 11.0	memory length: 724299	epsilon:
0.009998020008555413	steps: 534	lr: 0.0001	evaluation reward: 6.01
episode: 3153	score: 11.0	memory length: 724790	epsilon:
0.009998020008555413	steps: 491	lr: 0.0001	evaluation reward: 6.05
episode: 3154	score: 5.0	memory length: 725117	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 6.03
episode: 3155	score: 3.0	memory length: 725326	epsilon:
0.009998020008555413	steps: 209	lr: 0.0001	evaluation reward: 6.0
episode: 3156	score: 5.0	memory length: 725650	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 5.97
episode: 3157	score: 5.0	memory length: 725977	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 5.96
episode: 3158	score: 4.0	memory length: 726238	epsilon:
0.009998020008555413	steps: 261	lr: 0.0001	evaluation reward: 5.93
episode: 3159	score: 9.0	memory length: 726759	epsilon:
0.009998020008555413	steps: 521	lr: 0.0001	evaluation reward: 5.93
episode: 3160	score: 5.0	memory length: 727067	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 5.91
episode: 3161	score: 9.0	memory length: 727467	epsilon:
0.009998020008555413	steps: 400	lr: 0.0001	evaluation reward: 5.93
episode: 3162	score: 3.0	memory length: 727678	epsilon:
0.009998020008555413	steps: 211	lr: 0.0001	evaluation reward: 5.89
episode: 3163	score: 5.0	memory length: 727984	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 5.91
episode: 3164	score: 10.0	memory length: 728492	epsilon:
0.009998020008555413	steps: 508	lr: 0.0001	evaluation reward: 5.96
episode: 3165	score: 9.0	memory length: 728963	epsilon:
0.009998020008555413	steps: 471	lr: 0.0001	evaluation reward: 6.0
episode: 3166	score: 11.0	memory length: 729506	epsilon:
0.009998020008555413	steps: 543	lr: 0.0001	evaluation reward: 6.08
episode: 3167	score: 8.0	memory length: 729895	epsilon:
0.009998020008555413	steps: 389	lr: 0.0001	evaluation reward: 6.1
episode: 3168	score: 8.0	memory length: 730314	epsilon:
0.009998020008555413	steps: 419	lr: 0.0001	evaluation reward: 6.12
episode: 3169	score: 8.0	memory length: 730718	epsilon:
0.009998020008555413	steps: 404	lr: 0.0001	evaluation reward: 6.15
episode: 3170	score: 5.0	memory length: 731004	epsilon:
0.009998020008555413	steps: 286	lr: 0.0001	evaluation reward: 6.14
episode: 3171	score: 7.0	memory length: 731435	epsilon:
0.009998020008555413	steps: 431	lr: 0.0001	evaluation reward: 6.17

episode: 3172	score: 4.0	memory length: 731691	epsilon:
0.009998020008555413	steps: 256	lr: 0.0001	evaluation reward: 6.15
episode: 3173	score: 3.0	memory length: 731920	epsilon:
0.009998020008555413	steps: 229	lr: 0.0001	evaluation reward: 6.12
episode: 3174	score: 9.0	memory length: 732374	epsilon:
0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 6.15
episode: 3175	score: 2.0	memory length: 732556	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 6.1
episode: 3176	score: 9.0	memory length: 733029	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 6.13
episode: 3177	score: 5.0	memory length: 733335	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 6.14
episode: 3178	score: 6.0	memory length: 733650	epsilon:
0.009998020008555413	steps: 315	lr: 0.0001	evaluation reward: 6.13
episode: 3179	score: 5.0	memory length: 733956	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 6.12
episode: 3180	score: 7.0	memory length: 734305	epsilon:
0.009998020008555413	steps: 349	lr: 0.0001	evaluation reward: 6.13
episode: 3181	score: 9.0	memory length: 734758	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 6.17
episode: 3182	score: 5.0	memory length: 735083	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 6.17
episode: 3183	score: 3.0	memory length: 735292	epsilon:
0.009998020008555413	steps: 209	lr: 0.0001	evaluation reward: 6.14
episode: 3184	score: 3.0	memory length: 735518	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 6.11
episode: 3185	score: 10.0	memory length: 735977	epsilon:
0.009998020008555413	steps: 459	lr: 0.0001	evaluation reward: 6.18
episode: 3186	score: 7.0	memory length: 736327	epsilon:
0.009998020008555413	steps: 350	lr: 0.0001	evaluation reward: 6.19
episode: 3187	score: 9.0	memory length: 736792	epsilon:
0.009998020008555413	steps: 465	lr: 0.0001	evaluation reward: 6.24
episode: 3188	score: 4.0	memory length: 737085	epsilon:
0.009998020008555413	steps: 293	lr: 0.0001	evaluation reward: 6.2
episode: 3189	score: 5.0	memory length: 737393	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 6.19
episode: 3190	score: 7.0	memory length: 737801	epsilon:
0.009998020008555413	steps: 408	lr: 0.0001	evaluation reward: 6.15
episode: 3191	score: 3.0	memory length: 738010	epsilon:
0.009998020008555413	steps: 209	lr: 0.0001	evaluation reward: 6.15
episode: 3192	score: 4.0	memory length: 738285	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 6.16
episode: 3193	score: 4.0	memory length: 738544	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.06
episode: 3194	score: 5.0	memory length: 738851	epsilon:
0.009998020008555413	steps: 307	lr: 0.0001	evaluation reward: 6.05
episode: 3195	score: 6.0	memory length: 739207	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 6.07

episode: 3196	score: 9.0	memory length: 739658	epsilon:
0.009998020008555413	steps: 451	lr: 0.0001	evaluation reward: 6.06
episode: 3197	score: 6.0	memory length: 739994	epsilon:
0.009998020008555413	steps: 336	lr: 0.0001	evaluation reward: 6.07
episode: 3198	score: 9.0	memory length: 740448	epsilon:
0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 6.11
episode: 3199	score: 9.0	memory length: 740904	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 6.14
episode: 3200	score: 3.0	memory length: 741117	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 6.13
episode: 3201	score: 7.0	memory length: 741529	epsilon:
0.009998020008555413	steps: 412	lr: 0.0001	evaluation reward: 6.15
episode: 3202	score: 9.0	memory length: 741938	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 6.17
episode: 3203	score: 6.0	memory length: 742261	epsilon:
0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 6.17
episode: 3204	score: 3.0	memory length: 742474	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 6.16
episode: 3205	score: 8.0	memory length: 742887	epsilon:
0.009998020008555413	steps: 413	lr: 0.0001	evaluation reward: 6.18
episode: 3206	score: 5.0	memory length: 743209	epsilon:
0.009998020008555413	steps: 322	lr: 0.0001	evaluation reward: 6.16
episode: 3207	score: 5.0	memory length: 743512	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 6.16
episode: 3208	score: 8.0	memory length: 743928	epsilon:
0.009998020008555413	steps: 416	lr: 0.0001	evaluation reward: 6.17
episode: 3209	score: 7.0	memory length: 744307	epsilon:
0.009998020008555413	steps: 379	lr: 0.0001	evaluation reward: 6.14
episode: 3210	score: 5.0	memory length: 744615	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 6.14
episode: 3211	score: 4.0	memory length: 744878	epsilon:
0.009998020008555413	steps: 263	lr: 0.0001	evaluation reward: 6.1
episode: 3212	score: 3.0	memory length: 745087	epsilon:
0.009998020008555413	steps: 209	lr: 0.0001	evaluation reward: 6.08
episode: 3213	score: 9.0	memory length: 745501	epsilon:
0.009998020008555413	steps: 414	lr: 0.0001	evaluation reward: 6.11
episode: 3214	score: 8.0	memory length: 745940	epsilon:
0.009998020008555413	steps: 439	lr: 0.0001	evaluation reward: 6.13
episode: 3215	score: 3.0	memory length: 746168	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 6.11
episode: 3216	score: 6.0	memory length: 746504	epsilon:
0.009998020008555413	steps: 336	lr: 0.0001	evaluation reward: 6.13
episode: 3217	score: 10.0	memory length: 747020	epsilon:
0.009998020008555413	steps: 516	lr: 0.0001	evaluation reward: 6.2
episode: 3218	score: 6.0	memory length: 747391	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 6.19
episode: 3219	score: 4.0	memory length: 747665	epsilon:
0.009998020008555413	steps: 274	lr: 0.0001	evaluation reward: 6.17

episode: 3220	score: 11.0	memory length: 748206	epsilon:
0.009998020008555413	steps: 541	lr: 0.0001	evaluation reward: 6.23
episode: 3221	score: 3.0	memory length: 748416	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 6.2
episode: 3222	score: 7.0	memory length: 748844	epsilon:
0.009998020008555413	steps: 428	lr: 0.0001	evaluation reward: 6.22
episode: 3223	score: 8.0	memory length: 749292	epsilon:
0.009998020008555413	steps: 448	lr: 0.0001	evaluation reward: 6.26
episode: 3224	score: 7.0	memory length: 749662	epsilon:
0.009998020008555413	steps: 370	lr: 0.0001	evaluation reward: 6.28
episode: 3225	score: 6.0	memory length: 750031	epsilon:
0.009998020008555413	steps: 369	lr: 0.0001	evaluation reward: 6.3
episode: 3226	score: 4.0	memory length: 750293	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 6.3
episode: 3227	score: 4.0	memory length: 750551	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 6.26
episode: 3228	score: 9.0	memory length: 751004	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 6.3
episode: 3229	score: 11.0	memory length: 751500	epsilon:
0.009998020008555413	steps: 496	lr: 0.0001	evaluation reward: 6.36
episode: 3230	score: 2.0	memory length: 751701	epsilon:
0.009998020008555413	steps: 201	lr: 0.0001	evaluation reward: 6.29
episode: 3231	score: 2.0	memory length: 751883	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 6.26
episode: 3232	score: 6.0	memory length: 752238	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 6.2
episode: 3233	score: 4.0	memory length: 752498	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 6.14
episode: 3234	score: 15.0	memory length: 753042	epsilon:
0.009998020008555413	steps: 544	lr: 0.0001	evaluation reward: 6.23
episode: 3235	score: 10.0	memory length: 753568	epsilon:
0.009998020008555413	steps: 526	lr: 0.0001	evaluation reward: 6.29
episode: 3236	score: 9.0	memory length: 754036	epsilon:
0.009998020008555413	steps: 468	lr: 0.0001	evaluation reward: 6.36
episode: 3237	score: 6.0	memory length: 754356	epsilon:
0.009998020008555413	steps: 320	lr: 0.0001	evaluation reward: 6.37
episode: 3238	score: 7.0	memory length: 754733	epsilon:
0.009998020008555413	steps: 377	lr: 0.0001	evaluation reward: 6.4
episode: 3239	score: 8.0	memory length: 755158	epsilon:
0.009998020008555413	steps: 425	lr: 0.0001	evaluation reward: 6.4
episode: 3240	score: 6.0	memory length: 755512	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 6.38
episode: 3241	score: 6.0	memory length: 755867	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 6.4
episode: 3242	score: 10.0	memory length: 756258	epsilon:
0.009998020008555413	steps: 391	lr: 0.0001	evaluation reward: 6.47
episode: 3243	score: 6.0	memory length: 756568	epsilon:
0.009998020008555413	steps: 310	lr: 0.0001	evaluation reward: 6.43

episode: 3244	score: 8.0	memory length: 756997	epsilon:
0.009998020008555413	steps: 429	lr: 0.0001	evaluation reward: 6.45
episode: 3245	score: 4.0	memory length: 757278	epsilon:
0.009998020008555413	steps: 281	lr: 0.0001	evaluation reward: 6.45
episode: 3246	score: 8.0	memory length: 757748	epsilon:
0.009998020008555413	steps: 470	lr: 0.0001	evaluation reward: 6.5
episode: 3247	score: 4.0	memory length: 758024	epsilon:
0.009998020008555413	steps: 276	lr: 0.0001	evaluation reward: 6.45
episode: 3248	score: 7.0	memory length: 758411	epsilon:
0.009998020008555413	steps: 387	lr: 0.0001	evaluation reward: 6.46
episode: 3249	score: 9.0	memory length: 758882	epsilon:
0.009998020008555413	steps: 471	lr: 0.0001	evaluation reward: 6.47
episode: 3250	score: 10.0	memory length: 759388	epsilon:
0.009998020008555413	steps: 506	lr: 0.0001	evaluation reward: 6.51
episode: 3251	score: 7.0	memory length: 759766	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 6.51
episode: 3252	score: 2.0	memory length: 759948	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 6.42
episode: 3253	score: 9.0	memory length: 760412	epsilon:
0.009998020008555413	steps: 464	lr: 0.0001	evaluation reward: 6.4
episode: 3254	score: 5.0	memory length: 760741	epsilon:
0.009998020008555413	steps: 329	lr: 0.0001	evaluation reward: 6.4
episode: 3255	score: 5.0	memory length: 761064	epsilon:
0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 6.42
episode: 3256	score: 8.0	memory length: 761539	epsilon:
0.009998020008555413	steps: 475	lr: 0.0001	evaluation reward: 6.45
episode: 3257	score: 6.0	memory length: 761877	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 6.46
episode: 3258	score: 4.0	memory length: 762151	epsilon:
0.009998020008555413	steps: 274	lr: 0.0001	evaluation reward: 6.46
episode: 3259	score: 9.0	memory length: 762658	epsilon:
0.009998020008555413	steps: 507	lr: 0.0001	evaluation reward: 6.46
episode: 3260	score: 10.0	memory length: 763178	epsilon:
0.009998020008555413	steps: 520	lr: 0.0001	evaluation reward: 6.51
episode: 3261	score: 5.0	memory length: 763501	epsilon:
0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 6.47
episode: 3262	score: 7.0	memory length: 763880	epsilon:
0.009998020008555413	steps: 379	lr: 0.0001	evaluation reward: 6.51
episode: 3263	score: 3.0	memory length: 764125	epsilon:
0.009998020008555413	steps: 245	lr: 0.0001	evaluation reward: 6.49
episode: 3264	score: 6.0	memory length: 764496	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 6.45
episode: 3265	score: 6.0	memory length: 764849	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 6.42
episode: 3266	score: 7.0	memory length: 765232	epsilon:
0.009998020008555413	steps: 383	lr: 0.0001	evaluation reward: 6.38
episode: 3267	score: 7.0	memory length: 765632	epsilon:
0.009998020008555413	steps: 400	lr: 0.0001	evaluation reward: 6.37

episode: 3268	score: 8.0	memory length: 766080	epsilon:
0.009998020008555413	steps: 448	lr: 0.0001	evaluation reward: 6.37
episode: 3269	score: 8.0	memory length: 766523	epsilon:
0.009998020008555413	steps: 443	lr: 0.0001	evaluation reward: 6.37
episode: 3270	score: 6.0	memory length: 766879	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 6.38
episode: 3271	score: 10.0	memory length: 767395	epsilon:
0.009998020008555413	steps: 516	lr: 0.0001	evaluation reward: 6.41
episode: 3272	score: 7.0	memory length: 767821	epsilon:
0.009998020008555413	steps: 426	lr: 0.0001	evaluation reward: 6.44
episode: 3273	score: 5.0	memory length: 768129	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 6.46
episode: 3274	score: 7.0	memory length: 768530	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 6.44
episode: 3275	score: 5.0	memory length: 768842	epsilon:
0.009998020008555413	steps: 312	lr: 0.0001	evaluation reward: 6.47
episode: 3276	score: 6.0	memory length: 769199	epsilon:
0.009998020008555413	steps: 357	lr: 0.0001	evaluation reward: 6.44
episode: 3277	score: 5.0	memory length: 769523	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 6.44
episode: 3278	score: 4.0	memory length: 769782	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.42
episode: 3279	score: 7.0	memory length: 770156	epsilon:
0.009998020008555413	steps: 374	lr: 0.0001	evaluation reward: 6.44
episode: 3280	score: 3.0	memory length: 770368	epsilon:
0.009998020008555413	steps: 212	lr: 0.0001	evaluation reward: 6.4
episode: 3281	score: 10.0	memory length: 770878	epsilon:
0.009998020008555413	steps: 510	lr: 0.0001	evaluation reward: 6.41
episode: 3282	score: 5.0	memory length: 771201	epsilon:
0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 6.41
episode: 3283	score: 7.0	memory length: 771560	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 6.45
episode: 3284	score: 6.0	memory length: 771878	epsilon:
0.009998020008555413	steps: 318	lr: 0.0001	evaluation reward: 6.48
episode: 3285	score: 5.0	memory length: 772205	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 6.43
episode: 3286	score: 7.0	memory length: 772601	epsilon:
0.009998020008555413	steps: 396	lr: 0.0001	evaluation reward: 6.43
episode: 3287	score: 13.0	memory length: 773062	epsilon:
0.009998020008555413	steps: 461	lr: 0.0001	evaluation reward: 6.47
episode: 3288	score: 6.0	memory length: 773390	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 6.49
episode: 3289	score: 8.0	memory length: 773812	epsilon:
0.009998020008555413	steps: 422	lr: 0.0001	evaluation reward: 6.52
episode: 3290	score: 3.0	memory length: 774025	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 6.48
episode: 3291	score: 4.0	memory length: 774282	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 6.49

episode: 3292	score: 4.0	memory length: 774541	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.49
episode: 3293	score: 7.0	memory length: 774911	epsilon:
0.009998020008555413	steps: 370	lr: 0.0001	evaluation reward: 6.52
episode: 3294	score: 10.0	memory length: 775443	epsilon:
0.009998020008555413	steps: 532	lr: 0.0001	evaluation reward: 6.57
episode: 3295	score: 12.0	memory length: 775919	epsilon:
0.009998020008555413	steps: 476	lr: 0.0001	evaluation reward: 6.63
episode: 3296	score: 7.0	memory length: 776259	epsilon:
0.009998020008555413	steps: 340	lr: 0.0001	evaluation reward: 6.61
episode: 3297	score: 3.0	memory length: 776467	epsilon:
0.009998020008555413	steps: 208	lr: 0.0001	evaluation reward: 6.58
episode: 3298	score: 5.0	memory length: 776791	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 6.54
episode: 3299	score: 12.0	memory length: 777261	epsilon:
0.009998020008555413	steps: 470	lr: 0.0001	evaluation reward: 6.57
episode: 3300	score: 5.0	memory length: 777588	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 6.59
episode: 3301	score: 4.0	memory length: 777847	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.56
episode: 3302	score: 5.0	memory length: 778153	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 6.52
episode: 3303	score: 7.0	memory length: 778539	epsilon:
0.009998020008555413	steps: 386	lr: 0.0001	evaluation reward: 6.53
episode: 3304	score: 4.0	memory length: 778782	epsilon:
0.009998020008555413	steps: 243	lr: 0.0001	evaluation reward: 6.54
episode: 3305	score: 11.0	memory length: 779280	epsilon:
0.009998020008555413	steps: 498	lr: 0.0001	evaluation reward: 6.57
episode: 3306	score: 6.0	memory length: 779652	epsilon:
0.009998020008555413	steps: 372	lr: 0.0001	evaluation reward: 6.58
episode: 3307	score: 3.0	memory length: 779878	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 6.56
episode: 3308	score: 3.0	memory length: 780104	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 6.51
episode: 3309	score: 11.0	memory length: 780537	epsilon:
0.009998020008555413	steps: 433	lr: 0.0001	evaluation reward: 6.55
episode: 3310	score: 6.0	memory length: 780900	epsilon:
0.009998020008555413	steps: 363	lr: 0.0001	evaluation reward: 6.56
episode: 3311	score: 6.0	memory length: 781257	epsilon:
0.009998020008555413	steps: 357	lr: 0.0001	evaluation reward: 6.58
episode: 3312	score: 7.0	memory length: 781641	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 6.62
episode: 3313	score: 6.0	memory length: 781976	epsilon:
0.009998020008555413	steps: 335	lr: 0.0001	evaluation reward: 6.59
episode: 3314	score: 13.0	memory length: 782559	epsilon:
0.009998020008555413	steps: 583	lr: 0.0001	evaluation reward: 6.64
episode: 3315	score: 2.0	memory length: 782741	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 6.63

episode: 3316	score: 5.0	memory length: 783030	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 6.62
episode: 3317	score: 4.0	memory length: 783289	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.56
episode: 3318	score: 4.0	memory length: 783548	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.54
episode: 3319	score: 7.0	memory length: 783954	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 6.57
episode: 3320	score: 8.0	memory length: 784397	epsilon:
0.009998020008555413	steps: 443	lr: 0.0001	evaluation reward: 6.54
episode: 3321	score: 4.0	memory length: 784670	epsilon:
0.009998020008555413	steps: 273	lr: 0.0001	evaluation reward: 6.55
episode: 3322	score: 7.0	memory length: 785056	epsilon:
0.009998020008555413	steps: 386	lr: 0.0001	evaluation reward: 6.55
episode: 3323	score: 9.0	memory length: 785495	epsilon:
0.009998020008555413	steps: 439	lr: 0.0001	evaluation reward: 6.56
episode: 3324	score: 6.0	memory length: 785866	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 6.55
episode: 3325	score: 4.0	memory length: 786105	epsilon:
0.009998020008555413	steps: 239	lr: 0.0001	evaluation reward: 6.53
episode: 3326	score: 4.0	memory length: 786362	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 6.53
episode: 3327	score: 6.0	memory length: 786715	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 6.55
episode: 3328	score: 3.0	memory length: 786925	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 6.49
episode: 3329	score: 6.0	memory length: 787284	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 6.44
episode: 3330	score: 8.0	memory length: 787698	epsilon:
0.009998020008555413	steps: 414	lr: 0.0001	evaluation reward: 6.5
episode: 3331	score: 10.0	memory length: 788175	epsilon:
0.009998020008555413	steps: 477	lr: 0.0001	evaluation reward: 6.58
episode: 3332	score: 4.0	memory length: 788432	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 6.56
episode: 3333	score: 6.0	memory length: 788768	epsilon:
0.009998020008555413	steps: 336	lr: 0.0001	evaluation reward: 6.58
episode: 3334	score: 6.0	memory length: 789119	epsilon:
0.009998020008555413	steps: 351	lr: 0.0001	evaluation reward: 6.49
episode: 3335	score: 5.0	memory length: 789437	epsilon:
0.009998020008555413	steps: 318	lr: 0.0001	evaluation reward: 6.44
episode: 3336	score: 11.0	memory length: 789909	epsilon:
0.009998020008555413	steps: 472	lr: 0.0001	evaluation reward: 6.46
episode: 3337	score: 5.0	memory length: 790237	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 6.45
episode: 3338	score: 5.0	memory length: 790526	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 6.43
episode: 3339	score: 9.0	memory length: 790997	epsilon:
0.009998020008555413	steps: 471	lr: 0.0001	evaluation reward: 6.44

episode: 3340	score: 8.0	memory length: 791449	epsilon:
0.009998020008555413	steps: 452	lr: 0.0001	evaluation reward: 6.46
episode: 3341	score: 6.0	memory length: 791794	epsilon:
0.009998020008555413	steps: 345	lr: 0.0001	evaluation reward: 6.46
episode: 3342	score: 4.0	memory length: 792056	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 6.4
episode: 3343	score: 7.0	memory length: 792448	epsilon:
0.009998020008555413	steps: 392	lr: 0.0001	evaluation reward: 6.41
episode: 3344	score: 5.0	memory length: 792756	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 6.38
episode: 3345	score: 8.0	memory length: 793190	epsilon:
0.009998020008555413	steps: 434	lr: 0.0001	evaluation reward: 6.42
episode: 3346	score: 5.0	memory length: 793494	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 6.39
episode: 3347	score: 17.0	memory length: 794130	epsilon:
0.009998020008555413	steps: 636	lr: 0.0001	evaluation reward: 6.52
episode: 3348	score: 10.0	memory length: 794634	epsilon:
0.009998020008555413	steps: 504	lr: 0.0001	evaluation reward: 6.55
episode: 3349	score: 10.0	memory length: 795130	epsilon:
0.009998020008555413	steps: 496	lr: 0.0001	evaluation reward: 6.56
episode: 3350	score: 5.0	memory length: 795438	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 6.51
episode: 3351	score: 8.0	memory length: 795896	epsilon:
0.009998020008555413	steps: 458	lr: 0.0001	evaluation reward: 6.52
episode: 3352	score: 6.0	memory length: 796250	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 6.56
episode: 3353	score: 3.0	memory length: 796459	epsilon:
0.009998020008555413	steps: 209	lr: 0.0001	evaluation reward: 6.5
episode: 3354	score: 5.0	memory length: 796762	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 6.5
episode: 3355	score: 4.0	memory length: 797018	epsilon:
0.009998020008555413	steps: 256	lr: 0.0001	evaluation reward: 6.49
episode: 3356	score: 5.0	memory length: 797323	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 6.46
episode: 3357	score: 3.0	memory length: 797548	epsilon:
0.009998020008555413	steps: 225	lr: 0.0001	evaluation reward: 6.43
episode: 3358	score: 2.0	memory length: 797729	epsilon:
0.009998020008555413	steps: 181	lr: 0.0001	evaluation reward: 6.41
episode: 3359	score: 4.0	memory length: 797987	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 6.36
episode: 3360	score: 4.0	memory length: 798244	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 6.3
episode: 3361	score: 3.0	memory length: 798469	epsilon:
0.009998020008555413	steps: 225	lr: 0.0001	evaluation reward: 6.28
episode: 3362	score: 6.0	memory length: 798807	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 6.27
episode: 3363	score: 10.0	memory length: 799313	epsilon:
0.009998020008555413	steps: 506	lr: 0.0001	evaluation reward: 6.34

episode: 3364	score: 6.0	memory length: 799647	epsilon:
0.009998020008555413	steps: 334	lr: 0.0001	evaluation reward: 6.34
episode: 3365	score: 6.0	memory length: 800003	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 6.34
episode: 3366	score: 4.0	memory length: 800261	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 6.31
episode: 3367	score: 9.0	memory length: 800704	epsilon:
0.009998020008555413	steps: 443	lr: 0.0001	evaluation reward: 6.33
episode: 3368	score: 3.0	memory length: 800917	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 6.28
episode: 3369	score: 4.0	memory length: 801176	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.24
episode: 3370	score: 6.0	memory length: 801532	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 6.24
episode: 3371	score: 19.0	memory length: 802099	epsilon:
0.009998020008555413	steps: 567	lr: 0.0001	evaluation reward: 6.33
episode: 3372	score: 8.0	memory length: 802550	epsilon:
0.009998020008555413	steps: 451	lr: 0.0001	evaluation reward: 6.34
episode: 3373	score: 9.0	memory length: 803017	epsilon:
0.009998020008555413	steps: 467	lr: 0.0001	evaluation reward: 6.38
episode: 3374	score: 9.0	memory length: 803460	epsilon:
0.009998020008555413	steps: 443	lr: 0.0001	evaluation reward: 6.4
episode: 3375	score: 4.0	memory length: 803718	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 6.39
episode: 3376	score: 8.0	memory length: 804113	epsilon:
0.009998020008555413	steps: 395	lr: 0.0001	evaluation reward: 6.41
episode: 3377	score: 7.0	memory length: 804489	epsilon:
0.009998020008555413	steps: 376	lr: 0.0001	evaluation reward: 6.43
episode: 3378	score: 3.0	memory length: 804717	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 6.42
episode: 3379	score: 13.0	memory length: 805334	epsilon:
0.009998020008555413	steps: 617	lr: 0.0001	evaluation reward: 6.48
episode: 3380	score: 6.0	memory length: 805688	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 6.51
episode: 3381	score: 5.0	memory length: 806034	epsilon:
0.009998020008555413	steps: 346	lr: 0.0001	evaluation reward: 6.46
episode: 3382	score: 4.0	memory length: 806306	epsilon:
0.009998020008555413	steps: 272	lr: 0.0001	evaluation reward: 6.45
episode: 3383	score: 4.0	memory length: 806579	epsilon:
0.009998020008555413	steps: 273	lr: 0.0001	evaluation reward: 6.42
episode: 3384	score: 8.0	memory length: 807051	epsilon:
0.009998020008555413	steps: 472	lr: 0.0001	evaluation reward: 6.44
episode: 3385	score: 7.0	memory length: 807459	epsilon:
0.009998020008555413	steps: 408	lr: 0.0001	evaluation reward: 6.46
episode: 3386	score: 3.0	memory length: 807670	epsilon:
0.009998020008555413	steps: 211	lr: 0.0001	evaluation reward: 6.42
episode: 3387	score: 4.0	memory length: 807948	epsilon:
0.009998020008555413	steps: 278	lr: 0.0001	evaluation reward: 6.33

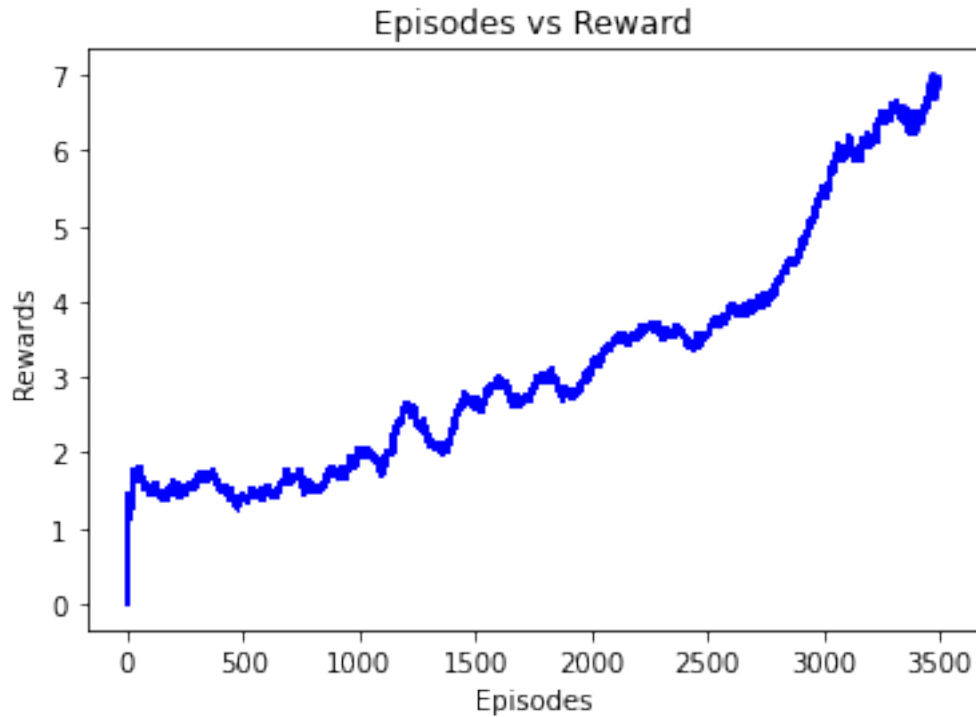
episode: 3388	score: 14.0	memory length: 808515	epsilon:
0.009998020008555413	steps: 567	lr: 0.0001	evaluation reward: 6.41
episode: 3389	score: 4.0	memory length: 808757	epsilon:
0.009998020008555413	steps: 242	lr: 0.0001	evaluation reward: 6.37
episode: 3390	score: 8.0	memory length: 809205	epsilon:
0.009998020008555413	steps: 448	lr: 0.0001	evaluation reward: 6.42
episode: 3391	score: 4.0	memory length: 809462	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 6.42
episode: 3392	score: 4.0	memory length: 809721	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.42
episode: 3393	score: 3.0	memory length: 809949	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 6.38
episode: 3394	score: 4.0	memory length: 810208	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.32
episode: 3395	score: 6.0	memory length: 810564	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 6.26
episode: 3396	score: 5.0	memory length: 810869	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 6.24
episode: 3397	score: 7.0	memory length: 811271	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 6.28
episode: 3398	score: 6.0	memory length: 811630	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 6.29
episode: 3399	score: 11.0	memory length: 812031	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 6.28
episode: 3400	score: 8.0	memory length: 812421	epsilon:
0.009998020008555413	steps: 390	lr: 0.0001	evaluation reward: 6.31
episode: 3401	score: 4.0	memory length: 812683	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 6.31
episode: 3402	score: 8.0	memory length: 813140	epsilon:
0.009998020008555413	steps: 457	lr: 0.0001	evaluation reward: 6.34
episode: 3403	score: 7.0	memory length: 813527	epsilon:
0.009998020008555413	steps: 387	lr: 0.0001	evaluation reward: 6.34
episode: 3404	score: 10.0	memory length: 814034	epsilon:
0.009998020008555413	steps: 507	lr: 0.0001	evaluation reward: 6.4
episode: 3405	score: 7.0	memory length: 814436	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 6.36
episode: 3406	score: 4.0	memory length: 814696	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 6.34
episode: 3407	score: 14.0	memory length: 815355	epsilon:
0.009998020008555413	steps: 659	lr: 0.0001	evaluation reward: 6.45
episode: 3408	score: 9.0	memory length: 815815	epsilon:
0.009998020008555413	steps: 460	lr: 0.0001	evaluation reward: 6.51
episode: 3409	score: 7.0	memory length: 816218	epsilon:
0.009998020008555413	steps: 403	lr: 0.0001	evaluation reward: 6.47
episode: 3410	score: 5.0	memory length: 816525	epsilon:
0.009998020008555413	steps: 307	lr: 0.0001	evaluation reward: 6.46
episode: 3411	score: 3.0	memory length: 816737	epsilon:
0.009998020008555413	steps: 212	lr: 0.0001	evaluation reward: 6.43

episode: 3412	score: 7.0	memory length: 817104	epsilon:
0.009998020008555413	steps: 367	lr: 0.0001	evaluation reward: 6.43
episode: 3413	score: 10.0	memory length: 817575	epsilon:
0.009998020008555413	steps: 471	lr: 0.0001	evaluation reward: 6.47
episode: 3414	score: 3.0	memory length: 817800	epsilon:
0.009998020008555413	steps: 225	lr: 0.0001	evaluation reward: 6.37
episode: 3415	score: 7.0	memory length: 818218	epsilon:
0.009998020008555413	steps: 418	lr: 0.0001	evaluation reward: 6.42
episode: 3416	score: 4.0	memory length: 818493	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 6.41
episode: 3417	score: 3.0	memory length: 818723	epsilon:
0.009998020008555413	steps: 230	lr: 0.0001	evaluation reward: 6.4
episode: 3418	score: 3.0	memory length: 818933	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 6.39
episode: 3419	score: 9.0	memory length: 819398	epsilon:
0.009998020008555413	steps: 465	lr: 0.0001	evaluation reward: 6.41
episode: 3420	score: 6.0	memory length: 819749	epsilon:
0.009998020008555413	steps: 351	lr: 0.0001	evaluation reward: 6.39
episode: 3421	score: 8.0	memory length: 820154	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 6.43
episode: 3422	score: 6.0	memory length: 820510	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 6.42
episode: 3423	score: 7.0	memory length: 820895	epsilon:
0.009998020008555413	steps: 385	lr: 0.0001	evaluation reward: 6.4
episode: 3424	score: 5.0	memory length: 821189	epsilon:
0.009998020008555413	steps: 294	lr: 0.0001	evaluation reward: 6.39
episode: 3425	score: 9.0	memory length: 821646	epsilon:
0.009998020008555413	steps: 457	lr: 0.0001	evaluation reward: 6.44
episode: 3426	score: 5.0	memory length: 821971	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 6.45
episode: 3427	score: 5.0	memory length: 822262	epsilon:
0.009998020008555413	steps: 291	lr: 0.0001	evaluation reward: 6.44
episode: 3428	score: 9.0	memory length: 822723	epsilon:
0.009998020008555413	steps: 461	lr: 0.0001	evaluation reward: 6.5
episode: 3429	score: 6.0	memory length: 823087	epsilon:
0.009998020008555413	steps: 364	lr: 0.0001	evaluation reward: 6.5
episode: 3430	score: 9.0	memory length: 823542	epsilon:
0.009998020008555413	steps: 455	lr: 0.0001	evaluation reward: 6.51
episode: 3431	score: 8.0	memory length: 823995	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 6.49
episode: 3432	score: 11.0	memory length: 824522	epsilon:
0.009998020008555413	steps: 527	lr: 0.0001	evaluation reward: 6.56
episode: 3433	score: 7.0	memory length: 824884	epsilon:
0.009998020008555413	steps: 362	lr: 0.0001	evaluation reward: 6.57
episode: 3434	score: 8.0	memory length: 825350	epsilon:
0.009998020008555413	steps: 466	lr: 0.0001	evaluation reward: 6.59
episode: 3435	score: 5.0	memory length: 825675	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 6.59

episode: 3436	score: 8.0	memory length: 826127	epsilon:
0.009998020008555413	steps: 452	lr: 0.0001	evaluation reward: 6.56
episode: 3437	score: 5.0	memory length: 826436	epsilon:
0.009998020008555413	steps: 309	lr: 0.0001	evaluation reward: 6.56
episode: 3438	score: 14.0	memory length: 827012	epsilon:
0.009998020008555413	steps: 576	lr: 0.0001	evaluation reward: 6.65
episode: 3439	score: 4.0	memory length: 827269	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 6.6
episode: 3440	score: 8.0	memory length: 827711	epsilon:
0.009998020008555413	steps: 442	lr: 0.0001	evaluation reward: 6.6
episode: 3441	score: 11.0	memory length: 828131	epsilon:
0.009998020008555413	steps: 420	lr: 0.0001	evaluation reward: 6.65
episode: 3442	score: 5.0	memory length: 828457	epsilon:
0.009998020008555413	steps: 326	lr: 0.0001	evaluation reward: 6.66
episode: 3443	score: 5.0	memory length: 828764	epsilon:
0.009998020008555413	steps: 307	lr: 0.0001	evaluation reward: 6.64
episode: 3444	score: 7.0	memory length: 829136	epsilon:
0.009998020008555413	steps: 372	lr: 0.0001	evaluation reward: 6.66
episode: 3445	score: 12.0	memory length: 829703	epsilon:
0.009998020008555413	steps: 567	lr: 0.0001	evaluation reward: 6.7
episode: 3446	score: 5.0	memory length: 830012	epsilon:
0.009998020008555413	steps: 309	lr: 0.0001	evaluation reward: 6.7
episode: 3447	score: 9.0	memory length: 830471	epsilon:
0.009998020008555413	steps: 459	lr: 0.0001	evaluation reward: 6.62
episode: 3448	score: 7.0	memory length: 830838	epsilon:
0.009998020008555413	steps: 367	lr: 0.0001	evaluation reward: 6.59
episode: 3449	score: 11.0	memory length: 831298	epsilon:
0.009998020008555413	steps: 460	lr: 0.0001	evaluation reward: 6.6
episode: 3450	score: 7.0	memory length: 831669	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 6.62
episode: 3451	score: 8.0	memory length: 832074	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 6.62
episode: 3452	score: 10.0	memory length: 832558	epsilon:
0.009998020008555413	steps: 484	lr: 0.0001	evaluation reward: 6.66
episode: 3453	score: 7.0	memory length: 832942	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 6.7
episode: 3454	score: 8.0	memory length: 833368	epsilon:
0.009998020008555413	steps: 426	lr: 0.0001	evaluation reward: 6.73
episode: 3455	score: 3.0	memory length: 833578	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 6.72
episode: 3456	score: 6.0	memory length: 833974	epsilon:
0.009998020008555413	steps: 396	lr: 0.0001	evaluation reward: 6.73
episode: 3457	score: 6.0	memory length: 834328	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 6.76
episode: 3458	score: 10.0	memory length: 834729	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 6.84
episode: 3459	score: 9.0	memory length: 835212	epsilon:
0.009998020008555413	steps: 483	lr: 0.0001	evaluation reward: 6.89

episode: 3460	score: 6.0	memory length: 835522	epsilon:
0.009998020008555413	steps: 310	lr: 0.0001	evaluation reward: 6.91
episode: 3461	score: 7.0	memory length: 835917	epsilon:
0.009998020008555413	steps: 395	lr: 0.0001	evaluation reward: 6.95
episode: 3462	score: 2.0	memory length: 836099	epsilon:
0.009998020008555413	steps: 182	lr: 0.0001	evaluation reward: 6.91
episode: 3463	score: 10.0	memory length: 836607	epsilon:
0.009998020008555413	steps: 508	lr: 0.0001	evaluation reward: 6.91
episode: 3464	score: 7.0	memory length: 836975	epsilon:
0.009998020008555413	steps: 368	lr: 0.0001	evaluation reward: 6.92
episode: 3465	score: 3.0	memory length: 837184	epsilon:
0.009998020008555413	steps: 209	lr: 0.0001	evaluation reward: 6.89
episode: 3466	score: 4.0	memory length: 837440	epsilon:
0.009998020008555413	steps: 256	lr: 0.0001	evaluation reward: 6.89
episode: 3467	score: 8.0	memory length: 837876	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 6.88
episode: 3468	score: 8.0	memory length: 838315	epsilon:
0.009998020008555413	steps: 439	lr: 0.0001	evaluation reward: 6.93
episode: 3469	score: 11.0	memory length: 838866	epsilon:
0.009998020008555413	steps: 551	lr: 0.0001	evaluation reward: 7.0
episode: 3470	score: 5.0	memory length: 839159	epsilon:
0.009998020008555413	steps: 293	lr: 0.0001	evaluation reward: 6.99
episode: 3471	score: 6.0	memory length: 839499	epsilon:
0.009998020008555413	steps: 340	lr: 0.0001	evaluation reward: 6.86
episode: 3472	score: 4.0	memory length: 839761	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 6.82
episode: 3473	score: 4.0	memory length: 840034	epsilon:
0.009998020008555413	steps: 273	lr: 0.0001	evaluation reward: 6.77
episode: 3474	score: 5.0	memory length: 840304	epsilon:
0.009998020008555413	steps: 270	lr: 0.0001	evaluation reward: 6.73
episode: 3475	score: 11.0	memory length: 840705	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 6.8
episode: 3476	score: 4.0	memory length: 840964	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.76
episode: 3477	score: 6.0	memory length: 841318	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 6.75
episode: 3478	score: 8.0	memory length: 841754	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 6.8
episode: 3479	score: 2.0	memory length: 841935	epsilon:
0.009998020008555413	steps: 181	lr: 0.0001	evaluation reward: 6.69
episode: 3480	score: 11.0	memory length: 842481	epsilon:
0.009998020008555413	steps: 546	lr: 0.0001	evaluation reward: 6.74
episode: 3481	score: 9.0	memory length: 842839	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 6.78
episode: 3482	score: 8.0	memory length: 843248	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 6.82
episode: 3483	score: 11.0	memory length: 843788	epsilon:
0.009998020008555413	steps: 540	lr: 0.0001	evaluation reward: 6.89

episode: 3484	score: 4.0	memory length: 844049	epsilon:
0.009998020008555413	steps: 261	lr: 0.0001	evaluation reward: 6.85
episode: 3485	score: 5.0	memory length: 844341	epsilon:
0.009998020008555413	steps: 292	lr: 0.0001	evaluation reward: 6.83
episode: 3486	score: 7.0	memory length: 844743	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 6.87
episode: 3487	score: 6.0	memory length: 845098	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 6.89
episode: 3488	score: 8.0	memory length: 845521	epsilon:
0.009998020008555413	steps: 423	lr: 0.0001	evaluation reward: 6.83
episode: 3489	score: 4.0	memory length: 845780	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 6.83
episode: 3490	score: 7.0	memory length: 846154	epsilon:
0.009998020008555413	steps: 374	lr: 0.0001	evaluation reward: 6.82
episode: 3491	score: 7.0	memory length: 846514	epsilon:
0.009998020008555413	steps: 360	lr: 0.0001	evaluation reward: 6.85
episode: 3492	score: 7.0	memory length: 846882	epsilon:
0.009998020008555413	steps: 368	lr: 0.0001	evaluation reward: 6.88
episode: 3493	score: 10.0	memory length: 847340	epsilon:
0.009998020008555413	steps: 458	lr: 0.0001	evaluation reward: 6.95
episode: 3494	score: 6.0	memory length: 847679	epsilon:
0.009998020008555413	steps: 339	lr: 0.0001	evaluation reward: 6.97
episode: 3495	score: 4.0	memory length: 847937	epsilon:
0.009998020008555413	steps: 258	lr: 0.0001	evaluation reward: 6.95
episode: 3496	score: 5.0	memory length: 848261	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 6.95
episode: 3497	score: 5.0	memory length: 848589	epsilon:
0.009998020008555413	steps: 328	lr: 0.0001	evaluation reward: 6.93
episode: 3498	score: 6.0	memory length: 848942	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 6.93
episode: 3499	score: 8.0	memory length: 849413	epsilon:
0.009998020008555413	steps: 471	lr: 0.0001	evaluation reward: 6.9



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_dqn_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()

```