

MP5

December 10, 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.

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

```
Requirement already satisfied: gym in /usr/local/lib/python3.6/dist-packages
(0.17.3)
Requirement already satisfied: pyvirtualdisplay in
/usr/local/lib/python3.6/dist-packages (1.3.2)
Requirement already satisfied: scipy in /usr/local/lib/python3.6/dist-packages
(from gym) (1.4.1)
Requirement already satisfied: numpy>=1.10.4 in /usr/local/lib/python3.6/dist-
packages (from gym) (1.18.5)
Requirement already satisfied: pygamelet<=1.5.0,>=1.4.0 in
/usr/local/lib/python3.6/dist-packages (from gym) (1.5.0)
Requirement already satisfied: cloudpickle<1.7.0,>=1.2.0 in
/usr/local/lib/python3.6/dist-packages (from gym) (1.3.0)
Requirement already satisfied: EasyProcess in /usr/local/lib/python3.6/dist-
packages (from pyvirtualdisplay) (0.3)
Requirement already satisfied: future in /usr/local/lib/python3.6/dist-packages
(from pygamelet<=1.5.0,>=1.4.0->gym) (0.16.0)
Reading package lists... Done
Building dependency tree
Reading state information... Done
python-opengl is already the newest version (3.1.0+dfsg-1).
ffmpeg is already the newest version (7:3.4.8-0ubuntu0.2).
xvfb is already the newest version (2:1.19.6-1ubuntu4.8).
0 upgraded, 0 newly installed, 0 to remove and 14 not upgraded.
```

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

```
Requirement already up-to-date: setuptools in /usr/local/lib/python3.6/dist-
packages (51.0.0)
```

Requirement already satisfied: ez_setup in /usr/local/lib/python3.6/dist-packages (0.9)

Requirement already satisfied: gym[atari] in /usr/local/lib/python3.6/dist-packages (0.17.3)

Requirement already satisfied: pygame<=1.5.0,>=1.4.0 in /usr/local/lib/python3.6/dist-packages (from gym[atari]) (1.5.0)

Requirement already satisfied: numpy>=1.10.4 in /usr/local/lib/python3.6/dist-packages (from gym[atari]) (1.18.5)

Requirement already satisfied: scipy in /usr/local/lib/python3.6/dist-packages (from gym[atari]) (1.4.1)

Requirement already satisfied: cloudpickle<1.7.0,>=1.2.0 in /usr/local/lib/python3.6/dist-packages (from gym[atari]) (1.3.0)

Requirement already satisfied: opencv-python; extra == "atari" in /usr/local/lib/python3.6/dist-packages (from gym[atari]) (4.1.2.30)

Requirement already satisfied: atari-py~=0.2.0; extra == "atari" in /usr/local/lib/python3.6/dist-packages (from gym[atari]) (0.2.6)

Requirement already satisfied: Pillow; extra == "atari" in /usr/local/lib/python3.6/dist-packages (from gym[atari]) (7.0.0)

Requirement already satisfied: future in /usr/local/lib/python3.6/dist-packages (from pygame<=1.5.0,>=1.4.0->gym[atari]) (0.16.0)

Requirement already satisfied: six in /usr/local/lib/python3.6/dist-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]: from google.colab import drive
drive.mount('/content/gdrive')
import os
os.chdir("gdrive/My Drive/Colab Notebooks/assignment5_materials")
```

Drive already mounted at /content/gdrive; to attempt to forcibly remount, call drive.mount("/content/gdrive", force_remount=True).

```
[4]: %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.

```

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

```

```

[6]: 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.

```

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

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

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

```

1.2.1 Main Training Loop

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

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

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

    get_init_state(history, state)

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

        # Perform a fire action if ball is no longer on screen to continue onto
        →next life
        if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
            action = 0
        else:
            action = agent.get_action(np.float32(history[: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: 2.0   memory length: 180   epsilon: 1.0   steps: 180   lr:
0.0001      evaluation reward: 2.0
episode: 1   score: 1.0   memory length: 331   epsilon: 1.0   steps: 151   lr:
0.0001      evaluation reward: 1.5
episode: 2   score: 1.0   memory length: 483   epsilon: 1.0   steps: 152   lr:
0.0001      evaluation reward: 1.3333333333333333
episode: 3   score: 3.0   memory length: 750   epsilon: 1.0   steps: 267   lr:
0.0001      evaluation reward: 1.75
episode: 4   score: 0.0   memory length: 872   epsilon: 1.0   steps: 122   lr:
0.0001      evaluation reward: 1.4
episode: 5   score: 1.0   memory length: 1041   epsilon: 1.0   steps: 169
lr: 0.0001   evaluation reward: 1.3333333333333333
episode: 6   score: 0.0   memory length: 1164   epsilon: 1.0   steps: 123
lr: 0.0001   evaluation reward: 1.1428571428571428
episode: 7   score: 0.0   memory length: 1287   epsilon: 1.0   steps: 123
lr: 0.0001   evaluation reward: 1.0
episode: 8   score: 1.0   memory length: 1455   epsilon: 1.0   steps: 168

```

lr: 0.0001	evaluation reward: 1.0			
episode: 9	score: 2.0	memory length: 1654	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.1			
episode: 10	score: 2.0	memory length: 1852	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.1818181818181819			
episode: 11	score: 1.0	memory length: 2022	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.1666666666666667			
episode: 12	score: 1.0	memory length: 2193	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.1538461538461537			
episode: 13	score: 3.0	memory length: 2442	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.2857142857142858			
episode: 14	score: 6.0	memory length: 2837	epsilon: 1.0	steps: 395
lr: 0.0001	evaluation reward: 1.6			
episode: 15	score: 3.0	memory length: 3084	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.6875			
episode: 16	score: 4.0	memory length: 3381	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.8235294117647058			
episode: 17	score: 0.0	memory length: 3504	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7222222222222223			
episode: 18	score: 4.0	memory length: 3795	epsilon: 1.0	steps: 291
lr: 0.0001	evaluation reward: 1.8421052631578947			
episode: 19	score: 5.0	memory length: 4103	epsilon: 1.0	steps: 308
lr: 0.0001	evaluation reward: 2.0			
episode: 20	score: 0.0	memory length: 4226	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.9047619047619047			
episode: 21	score: 4.0	memory length: 4503	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 2.0			
episode: 22	score: 1.0	memory length: 4673	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.9565217391304348			
episode: 23	score: 4.0	memory length: 4947	epsilon: 1.0	steps: 274
lr: 0.0001	evaluation reward: 2.0416666666666665			
episode: 24	score: 0.0	memory length: 5070	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.96			
episode: 25	score: 1.0	memory length: 5221	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.9230769230769231			
episode: 26	score: 0.0	memory length: 5344	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.8518518518518519			
episode: 27	score: 2.0	memory length: 5542	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.8571428571428572			
episode: 28	score: 4.0	memory length: 5809	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.9310344827586208			
episode: 29	score: 1.0	memory length: 5977	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.9			
episode: 30	score: 1.0	memory length: 6149	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.8709677419354838			
episode: 31	score: 0.0	memory length: 6272	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.8125			
episode: 32	score: 4.0	memory length: 6568	epsilon: 1.0	steps: 296

lr: 0.0001	evaluation reward: 1.878787878787879		
episode: 33	score: 1.0	memory length: 6718	epsilon: 1.0 steps: 150
lr: 0.0001	evaluation reward: 1.8529411764705883		
episode: 34	score: 1.0	memory length: 6889	epsilon: 1.0 steps: 171
lr: 0.0001	evaluation reward: 1.8285714285714285		
episode: 35	score: 1.0	memory length: 7058	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.8055555555555556		
episode: 36	score: 3.0	memory length: 7287	epsilon: 1.0 steps: 229
lr: 0.0001	evaluation reward: 1.837837837837838		
episode: 37	score: 2.0	memory length: 7485	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.8421052631578947		
episode: 38	score: 2.0	memory length: 7683	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.8461538461538463		
episode: 39	score: 1.0	memory length: 7852	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.825		
episode: 40	score: 1.0	memory length: 8004	epsilon: 1.0 steps: 152
lr: 0.0001	evaluation reward: 1.8048780487804879		
episode: 41	score: 4.0	memory length: 8279	epsilon: 1.0 steps: 275
lr: 0.0001	evaluation reward: 1.8571428571428572		
episode: 42	score: 4.0	memory length: 8567	epsilon: 1.0 steps: 288
lr: 0.0001	evaluation reward: 1.9069767441860466		
episode: 43	score: 1.0	memory length: 8737	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.8863636363636365		
episode: 44	score: 2.0	memory length: 8935	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.8888888888888888		
episode: 45	score: 1.0	memory length: 9086	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.8695652173913044		
episode: 46	score: 1.0	memory length: 9237	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.851063829787234		
episode: 47	score: 2.0	memory length: 9452	epsilon: 1.0 steps: 215
lr: 0.0001	evaluation reward: 1.8541666666666667		
episode: 48	score: 1.0	memory length: 9602	epsilon: 1.0 steps: 150
lr: 0.0001	evaluation reward: 1.836734693877551		
episode: 49	score: 1.0	memory length: 9772	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.82		
episode: 50	score: 1.0	memory length: 9923	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.803921568627451		
episode: 51	score: 1.0	memory length: 10074	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.7884615384615385		
episode: 52	score: 0.0	memory length: 10197	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7547169811320755		
episode: 53	score: 3.0	memory length: 10443	epsilon: 1.0 steps: 246
lr: 0.0001	evaluation reward: 1.7777777777777777		
episode: 54	score: 4.0	memory length: 10742	epsilon: 1.0 steps: 299
lr: 0.0001	evaluation reward: 1.8181818181818181		
episode: 55	score: 2.0	memory length: 10940	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.8214285714285714		
episode: 56	score: 2.0	memory length: 11137	epsilon: 1.0 steps: 197

lr: 0.0001	evaluation reward: 1.8245614035087718		
episode: 57	score: 1.0 memory length: 11289	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.8103448275862069		
episode: 58	score: 1.0 memory length: 11458	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.7966101694915255		
episode: 59	score: 2.0 memory length: 11658	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.8		
episode: 60	score: 3.0 memory length: 11904	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.819672131147541		
episode: 61	score: 3.0 memory length: 12150	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.8387096774193548		
episode: 62	score: 4.0 memory length: 12441	epsilon: 1.0	steps: 291
lr: 0.0001	evaluation reward: 1.873015873015873		
episode: 63	score: 3.0 memory length: 12667	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.890625		
episode: 64	score: 2.0 memory length: 12887	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.8923076923076922		
episode: 65	score: 0.0 memory length: 13010	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.8636363636363635		
episode: 66	score: 2.0 memory length: 13231	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.8656716417910448		
episode: 67	score: 0.0 memory length: 13354	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.838235294117647		
episode: 68	score: 0.0 memory length: 13477	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.8115942028985508		
episode: 69	score: 2.0 memory length: 13693	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.8142857142857143		
episode: 70	score: 0.0 memory length: 13815	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.7887323943661972		
episode: 71	score: 0.0 memory length: 13938	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7638888888888888		
episode: 72	score: 4.0 memory length: 14238	epsilon: 1.0	steps: 300
lr: 0.0001	evaluation reward: 1.7945205479452055		
episode: 73	score: 2.0 memory length: 14436	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.7972972972972974		
episode: 74	score: 2.0 memory length: 14616	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.8		
episode: 75	score: 2.0 memory length: 14814	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.8026315789473684		
episode: 76	score: 0.0 memory length: 14937	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7792207792207793		
episode: 77	score: 2.0 memory length: 15134	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.7820512820512822		
episode: 78	score: 5.0 memory length: 15478	epsilon: 1.0	steps: 344
lr: 0.0001	evaluation reward: 1.8227848101265822		
episode: 79	score: 0.0 memory length: 15600	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.8		
episode: 80	score: 0.0 memory length: 15722	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.7777777777777777		
episode: 81	score: 2.0	memory length: 15920	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.7804878048780488		
episode: 82	score: 3.0	memory length: 16168	epsilon: 1.0 steps: 248
lr: 0.0001	evaluation reward: 1.7951807228915662		
episode: 83	score: 0.0	memory length: 16291	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7738095238095237		
episode: 84	score: 2.0	memory length: 16508	epsilon: 1.0 steps: 217
lr: 0.0001	evaluation reward: 1.776470588235294		
episode: 85	score: 2.0	memory length: 16705	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.7790697674418605		
episode: 86	score: 4.0	memory length: 17003	epsilon: 1.0 steps: 298
lr: 0.0001	evaluation reward: 1.8045977011494252		
episode: 87	score: 1.0	memory length: 17154	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.7954545454545454		
episode: 88	score: 1.0	memory length: 17305	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.7865168539325842		
episode: 89	score: 0.0	memory length: 17428	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7666666666666666		
episode: 90	score: 4.0	memory length: 17708	epsilon: 1.0 steps: 280
lr: 0.0001	evaluation reward: 1.7912087912087913		
episode: 91	score: 0.0	memory length: 17830	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.7717391304347827		
episode: 92	score: 4.0	memory length: 18147	epsilon: 1.0 steps: 317
lr: 0.0001	evaluation reward: 1.7956989247311828		
episode: 93	score: 2.0	memory length: 18344	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.797872340425532		
episode: 94	score: 0.0	memory length: 18467	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.7789473684210526		
episode: 95	score: 2.0	memory length: 18667	epsilon: 1.0 steps: 200
lr: 0.0001	evaluation reward: 1.78125		
episode: 96	score: 3.0	memory length: 18913	epsilon: 1.0 steps: 246
lr: 0.0001	evaluation reward: 1.7938144329896908		
episode: 97	score: 1.0	memory length: 19064	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.7857142857142858		
episode: 98	score: 2.0	memory length: 19283	epsilon: 1.0 steps: 219
lr: 0.0001	evaluation reward: 1.7878787878787878		
episode: 99	score: 4.0	memory length: 19582	epsilon: 1.0 steps: 299
lr: 0.0001	evaluation reward: 1.81		
episode: 100	score: 1.0	memory length: 19734	epsilon: 1.0 steps: 152
lr: 0.0001	evaluation reward: 1.8		
episode: 101	score: 2.0	memory length: 19949	epsilon: 1.0 steps: 215
lr: 0.0001	evaluation reward: 1.81		
episode: 102	score: 2.0	memory length: 20146	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.82		
episode: 103	score: 3.0	memory length: 20412	epsilon: 1.0 steps: 266
lr: 0.0001	evaluation reward: 1.82		
episode: 104	score: 2.0	memory length: 20610	epsilon: 1.0 steps: 198

lr: 0.0001	evaluation reward: 1.84		
episode: 105	score: 1.0 memory length: 20780	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.84		
episode: 106	score: 0.0 memory length: 20902	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.84		
episode: 107	score: 1.0 memory length: 21074	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.85		
episode: 108	score: 1.0 memory length: 21225	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.85		
episode: 109	score: 3.0 memory length: 21453	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.86		
episode: 110	score: 0.0 memory length: 21575	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.84		
episode: 111	score: 0.0 memory length: 21698	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.83		
episode: 112	score: 0.0 memory length: 21821	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.82		
episode: 113	score: 0.0 memory length: 21944	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.79		
episode: 114	score: 2.0 memory length: 22164	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.75		
episode: 115	score: 3.0 memory length: 22426	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.75		
episode: 116	score: 0.0 memory length: 22548	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.71		
episode: 117	score: 1.0 memory length: 22717	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.72		
episode: 118	score: 2.0 memory length: 22915	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.7		
episode: 119	score: 2.0 memory length: 23117	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.67		
episode: 120	score: 2.0 memory length: 23315	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.69		
episode: 121	score: 0.0 memory length: 23437	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.65		
episode: 122	score: 1.0 memory length: 23609	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.65		
episode: 123	score: 2.0 memory length: 23807	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.63		
episode: 124	score: 1.0 memory length: 23979	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.64		
episode: 125	score: 1.0 memory length: 24148	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.64		
episode: 126	score: 1.0 memory length: 24317	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.65		
episode: 127	score: 3.0 memory length: 24548	epsilon: 1.0	steps: 231
lr: 0.0001	evaluation reward: 1.66		
episode: 128	score: 2.0 memory length: 24746	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.64		
episode: 129	score: 1.0 memory length: 24916	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.64		
episode: 130	score: 1.0 memory length: 25084	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.64		
episode: 131	score: 4.0 memory length: 25340	epsilon: 1.0	steps: 256
lr: 0.0001	evaluation reward: 1.68		
episode: 132	score: 1.0 memory length: 25512	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.65		
episode: 133	score: 3.0 memory length: 25759	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.67		
episode: 134	score: 2.0 memory length: 25956	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.68		
episode: 135	score: 1.0 memory length: 26107	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.68		
episode: 136	score: 3.0 memory length: 26351	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.68		
episode: 137	score: 2.0 memory length: 26569	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.68		
episode: 138	score: 4.0 memory length: 26864	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.7		
episode: 139	score: 1.0 memory length: 27015	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.7		
episode: 140	score: 1.0 memory length: 27166	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.7		
episode: 141	score: 1.0 memory length: 27335	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.67		
episode: 142	score: 3.0 memory length: 27582	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.66		
episode: 143	score: 2.0 memory length: 27802	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.67		
episode: 144	score: 0.0 memory length: 27925	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65		
episode: 145	score: 3.0 memory length: 28172	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.67		
episode: 146	score: 1.0 memory length: 28342	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.67		
episode: 147	score: 2.0 memory length: 28539	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.67		
episode: 148	score: 3.0 memory length: 28804	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.69		
episode: 149	score: 0.0 memory length: 28927	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.68		
episode: 150	score: 1.0 memory length: 29078	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.68		
episode: 151	score: 0.0 memory length: 29201	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.67		
episode: 152	score: 4.0 memory length: 29497	epsilon: 1.0	steps: 296

lr: 0.0001	evaluation reward: 1.71			
episode: 153	score: 1.0	memory length: 29648	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.69			
episode: 154	score: 0.0	memory length: 29771	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65			
episode: 155	score: 0.0	memory length: 29894	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.63			
episode: 156	score: 0.0	memory length: 30017	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61			
episode: 157	score: 2.0	memory length: 30232	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.62			
episode: 158	score: 0.0	memory length: 30355	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61			
episode: 159	score: 1.0	memory length: 30506	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.6			
episode: 160	score: 2.0	memory length: 30724	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.59			
episode: 161	score: 5.0	memory length: 31050	epsilon: 1.0	steps: 326
lr: 0.0001	evaluation reward: 1.61			
episode: 162	score: 1.0	memory length: 31219	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.58			
episode: 163	score: 1.0	memory length: 31370	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.56			
episode: 164	score: 1.0	memory length: 31538	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.55			
episode: 165	score: 3.0	memory length: 31785	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.58			
episode: 166	score: 3.0	memory length: 31994	epsilon: 1.0	steps: 209
lr: 0.0001	evaluation reward: 1.59			
episode: 167	score: 2.0	memory length: 32191	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.61			
episode: 168	score: 4.0	memory length: 32468	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.65			
episode: 169	score: 2.0	memory length: 32666	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65			
episode: 170	score: 4.0	memory length: 32961	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.69			
episode: 171	score: 3.0	memory length: 33208	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.72			
episode: 172	score: 2.0	memory length: 33425	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.7			
episode: 173	score: 0.0	memory length: 33547	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.68			
episode: 174	score: 3.0	memory length: 33776	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.69			
episode: 175	score: 3.0	memory length: 34020	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.7			
episode: 176	score: 0.0	memory length: 34143	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.7		
episode: 177	score: 0.0 memory length: 34265	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.68		
episode: 178	score: 5.0 memory length: 34589	epsilon: 1.0	steps: 324
lr: 0.0001	evaluation reward: 1.68		
episode: 179	score: 2.0 memory length: 34786	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.7		
episode: 180	score: 4.0 memory length: 35099	epsilon: 1.0	steps: 313
lr: 0.0001	evaluation reward: 1.74		
episode: 181	score: 1.0 memory length: 35267	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.73		
episode: 182	score: 1.0 memory length: 35436	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.71		
episode: 183	score: 1.0 memory length: 35586	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.72		
episode: 184	score: 1.0 memory length: 35737	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.71		
episode: 185	score: 0.0 memory length: 35859	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.69		
episode: 186	score: 0.0 memory length: 35982	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.65		
episode: 187	score: 2.0 memory length: 36201	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.66		
episode: 188	score: 2.0 memory length: 36400	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.67		
episode: 189	score: 0.0 memory length: 36522	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.67		
episode: 190	score: 3.0 memory length: 36767	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.66		
episode: 191	score: 2.0 memory length: 36965	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.68		
episode: 192	score: 0.0 memory length: 37088	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64		
episode: 193	score: 0.0 memory length: 37210	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.62		
episode: 194	score: 2.0 memory length: 37428	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.64		
episode: 195	score: 0.0 memory length: 37551	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 196	score: 2.0 memory length: 37770	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.61		
episode: 197	score: 0.0 memory length: 37893	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6		
episode: 198	score: 0.0 memory length: 38015	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.58		
episode: 199	score: 3.0 memory length: 38283	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.57		
episode: 200	score: 2.0 memory length: 38504	epsilon: 1.0	steps: 221

lr: 0.0001	evaluation reward: 1.58		
episode: 201	score: 4.0 memory length: 38804	epsilon: 1.0	steps: 300
lr: 0.0001	evaluation reward: 1.6		
episode: 202	score: 1.0 memory length: 38975	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.59		
episode: 203	score: 1.0 memory length: 39145	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.57		
episode: 204	score: 1.0 memory length: 39314	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.56		
episode: 205	score: 2.0 memory length: 39511	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.57		
episode: 206	score: 2.0 memory length: 39711	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.59		
episode: 207	score: 3.0 memory length: 39937	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.61		
episode: 208	score: 0.0 memory length: 40060	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6		
episode: 209	score: 0.0 memory length: 40183	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57		
episode: 210	score: 5.0 memory length: 40527	epsilon: 1.0	steps: 344
lr: 0.0001	evaluation reward: 1.62		
episode: 211	score: 0.0 memory length: 40650	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 212	score: 0.0 memory length: 40773	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 213	score: 0.0 memory length: 40896	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.62		
episode: 214	score: 3.0 memory length: 41158	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.63		
episode: 215	score: 1.0 memory length: 41327	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.61		
episode: 216	score: 1.0 memory length: 41478	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.62		
episode: 217	score: 0.0 memory length: 41601	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61		
episode: 218	score: 1.0 memory length: 41751	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.6		
episode: 219	score: 0.0 memory length: 41874	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58		
episode: 220	score: 0.0 memory length: 41997	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56		
episode: 221	score: 6.0 memory length: 42357	epsilon: 1.0	steps: 360
lr: 0.0001	evaluation reward: 1.62		
episode: 222	score: 1.0 memory length: 42526	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.62		
episode: 223	score: 2.0 memory length: 42724	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.62		
episode: 224	score: 3.0 memory length: 42950	epsilon: 1.0	steps: 226

lr: 0.0001	evaluation reward: 1.64		
episode: 225	score: 1.0 memory length: 43119	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.64		
episode: 226	score: 1.0 memory length: 43290	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.64		
episode: 227	score: 1.0 memory length: 43459	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.62		
episode: 228	score: 1.0 memory length: 43629	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.61		
episode: 229	score: 0.0 memory length: 43752	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6		
episode: 230	score: 3.0 memory length: 43999	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.62		
episode: 231	score: 0.0 memory length: 44122	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58		
episode: 232	score: 1.0 memory length: 44294	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.58		
episode: 233	score: 2.0 memory length: 44494	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.57		
episode: 234	score: 0.0 memory length: 44617	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55		
episode: 235	score: 2.0 memory length: 44814	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.56		
episode: 236	score: 0.0 memory length: 44936	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.53		
episode: 237	score: 1.0 memory length: 45106	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.52		
episode: 238	score: 2.0 memory length: 45304	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5		
episode: 239	score: 2.0 memory length: 45522	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.51		
episode: 240	score: 3.0 memory length: 45765	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.53		
episode: 241	score: 2.0 memory length: 45983	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.54		
episode: 242	score: 3.0 memory length: 46209	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.54		
episode: 243	score: 2.0 memory length: 46406	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.54		
episode: 244	score: 1.0 memory length: 46574	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.55		
episode: 245	score: 1.0 memory length: 46745	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.53		
episode: 246	score: 1.0 memory length: 46896	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.53		
episode: 247	score: 2.0 memory length: 47114	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.53		
episode: 248	score: 0.0 memory length: 47237	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.5			
episode: 249	score: 2.0	memory length: 47454	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.52			
episode: 250	score: 4.0	memory length: 47751	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.55			
episode: 251	score: 0.0	memory length: 47874	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55			
episode: 252	score: 2.0	memory length: 48095	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.53			
episode: 253	score: 1.0	memory length: 48247	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.53			
episode: 254	score: 11.0	memory length: 48687	epsilon: 1.0	steps: 440
lr: 0.0001	evaluation reward: 1.64			
episode: 255	score: 2.0	memory length: 48885	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.66			
episode: 256	score: 2.0	memory length: 49085	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.68			
episode: 257	score: 1.0	memory length: 49257	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.67			
episode: 258	score: 1.0	memory length: 49426	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.68			
episode: 259	score: 3.0	memory length: 49673	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.7			
episode: 260	score: 0.0	memory length: 49796	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.68			
episode: 261	score: 3.0	memory length: 50023	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.66			
episode: 262	score: 1.0	memory length: 50191	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.66			
episode: 263	score: 2.0	memory length: 50389	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.67			
episode: 264	score: 2.0	memory length: 50609	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.68			
episode: 265	score: 2.0	memory length: 50825	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.67			
episode: 266	score: 0.0	memory length: 50948	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.64			
episode: 267	score: 2.0	memory length: 51164	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.64			
episode: 268	score: 3.0	memory length: 51412	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.63			
episode: 269	score: 2.0	memory length: 51610	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.63			
episode: 270	score: 0.0	memory length: 51732	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.59			
episode: 271	score: 2.0	memory length: 51929	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.58			
episode: 272	score: 2.0	memory length: 52150	epsilon: 1.0	steps: 221

lr: 0.0001	evaluation reward: 1.58		
episode: 273	score: 2.0 memory length: 52348	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6		
episode: 274	score: 2.0 memory length: 52545	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.59		
episode: 275	score: 0.0 memory length: 52668	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56		
episode: 276	score: 1.0 memory length: 52836	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.57		
episode: 277	score: 0.0 memory length: 52959	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57		
episode: 278	score: 1.0 memory length: 53127	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.53		
episode: 279	score: 2.0 memory length: 53325	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.53		
episode: 280	score: 2.0 memory length: 53522	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.51		
episode: 281	score: 0.0 memory length: 53645	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 282	score: 0.0 memory length: 53768	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 283	score: 3.0 memory length: 53994	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.51		
episode: 284	score: 1.0 memory length: 54145	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.51		
episode: 285	score: 2.0 memory length: 54347	epsilon: 1.0	steps: 202
lr: 0.0001	evaluation reward: 1.53		
episode: 286	score: 0.0 memory length: 54470	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53		
episode: 287	score: 1.0 memory length: 54640	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.52		
episode: 288	score: 1.0 memory length: 54809	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.51		
episode: 289	score: 2.0 memory length: 55009	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.53		
episode: 290	score: 0.0 memory length: 55132	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 291	score: 2.0 memory length: 55330	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5		
episode: 292	score: 1.0 memory length: 55480	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51		
episode: 293	score: 0.0 memory length: 55603	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 294	score: 1.0 memory length: 55772	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5		
episode: 295	score: 0.0 memory length: 55894	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.5		
episode: 296	score: 2.0 memory length: 56092	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.5		
episode: 297	score: 1.0 memory length: 56242	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51		
episode: 298	score: 1.0 memory length: 56393	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.52		
episode: 299	score: 1.0 memory length: 56563	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5		
episode: 300	score: 3.0 memory length: 56792	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.51		
episode: 301	score: 2.0 memory length: 56990	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49		
episode: 302	score: 2.0 memory length: 57211	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.5		
episode: 303	score: 0.0 memory length: 57333	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.49		
episode: 304	score: 1.0 memory length: 57505	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.49		
episode: 305	score: 4.0 memory length: 57800	epsilon: 1.0	steps: 295
lr: 0.0001	evaluation reward: 1.51		
episode: 306	score: 1.0 memory length: 57970	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5		
episode: 307	score: 3.0 memory length: 58235	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.5		
episode: 308	score: 2.0 memory length: 58433	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.52		
episode: 309	score: 0.0 memory length: 58556	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 310	score: 8.0 memory length: 58878	epsilon: 1.0	steps: 322
lr: 0.0001	evaluation reward: 1.55		
episode: 311	score: 1.0 memory length: 59047	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.56		
episode: 312	score: 2.0 memory length: 59264	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.58		
episode: 313	score: 0.0 memory length: 59387	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.58		
episode: 314	score: 1.0 memory length: 59538	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.56		
episode: 315	score: 2.0 memory length: 59736	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.57		
episode: 316	score: 0.0 memory length: 59859	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56		
episode: 317	score: 1.0 memory length: 60028	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.57		
episode: 318	score: 0.0 memory length: 60150	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.56		
episode: 319	score: 2.0 memory length: 60348	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.58		
episode: 320	score: 1.0 memory length: 60516	epsilon: 1.0	steps: 168

lr: 0.0001	evaluation reward: 1.59		
episode: 321	score: 0.0 memory length: 60639	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53		
episode: 322	score: 2.0 memory length: 60837	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.54		
episode: 323	score: 0.0 memory length: 60960	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 324	score: 1.0 memory length: 61112	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.5		
episode: 325	score: 2.0 memory length: 61330	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.51		
episode: 326	score: 2.0 memory length: 61528	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.52		
episode: 327	score: 1.0 memory length: 61696	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.52		
episode: 328	score: 2.0 memory length: 61893	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.53		
episode: 329	score: 0.0 memory length: 62016	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53		
episode: 330	score: 1.0 memory length: 62185	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.51		
episode: 331	score: 2.0 memory length: 62384	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.53		
episode: 332	score: 2.0 memory length: 62582	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.54		
episode: 333	score: 2.0 memory length: 62783	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.54		
episode: 334	score: 4.0 memory length: 63100	epsilon: 1.0	steps: 317
lr: 0.0001	evaluation reward: 1.58		
episode: 335	score: 1.0 memory length: 63250	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.57		
episode: 336	score: 1.0 memory length: 63420	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.58		
episode: 337	score: 2.0 memory length: 63618	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.59		
episode: 338	score: 0.0 memory length: 63740	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.57		
episode: 339	score: 0.0 memory length: 63863	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55		
episode: 340	score: 0.0 memory length: 63985	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.52		
episode: 341	score: 0.0 memory length: 64108	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 342	score: 0.0 memory length: 64230	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.47		
episode: 343	score: 3.0 memory length: 64475	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.48		
episode: 344	score: 5.0 memory length: 64821	epsilon: 1.0	steps: 346

lr: 0.0001	evaluation reward: 1.52		
episode: 345	score: 0.0 memory length: 64944	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 346	score: 2.0 memory length: 65144	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.52		
episode: 347	score: 1.0 memory length: 65295	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.51		
episode: 348	score: 1.0 memory length: 65464	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.52		
episode: 349	score: 3.0 memory length: 65709	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.53		
episode: 350	score: 3.0 memory length: 65959	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.52		
episode: 351	score: 1.0 memory length: 66129	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.53		
episode: 352	score: 0.0 memory length: 66252	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 353	score: 1.0 memory length: 66402	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51		
episode: 354	score: 2.0 memory length: 66599	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.42		
episode: 355	score: 0.0 memory length: 66722	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 356	score: 2.0 memory length: 66920	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4		
episode: 357	score: 5.0 memory length: 67258	epsilon: 1.0	steps: 338
lr: 0.0001	evaluation reward: 1.44		
episode: 358	score: 4.0 memory length: 67512	epsilon: 1.0	steps: 254
lr: 0.0001	evaluation reward: 1.47		
episode: 359	score: 0.0 memory length: 67635	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 360	score: 2.0 memory length: 67833	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.46		
episode: 361	score: 3.0 memory length: 68100	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.46		
episode: 362	score: 2.0 memory length: 68297	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.47		
episode: 363	score: 2.0 memory length: 68497	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.47		
episode: 364	score: 1.0 memory length: 68647	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.46		
episode: 365	score: 0.0 memory length: 68770	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 366	score: 1.0 memory length: 68938	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.45		
episode: 367	score: 3.0 memory length: 69185	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.46		
episode: 368	score: 0.0 memory length: 69308	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.43		
episode: 369	score: 1.0 memory length: 69476	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.42		
episode: 370	score: 2.0 memory length: 69691	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.44		
episode: 371	score: 4.0 memory length: 69987	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.46		
episode: 372	score: 3.0 memory length: 70256	epsilon: 1.0	steps: 269
lr: 0.0001	evaluation reward: 1.47		
episode: 373	score: 3.0 memory length: 70501	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.48		
episode: 374	score: 0.0 memory length: 70623	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.46		
episode: 375	score: 1.0 memory length: 70791	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.47		
episode: 376	score: 0.0 memory length: 70914	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 377	score: 1.0 memory length: 71065	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.47		
episode: 378	score: 1.0 memory length: 71215	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.47		
episode: 379	score: 2.0 memory length: 71413	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 380	score: 0.0 memory length: 71535	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45		
episode: 381	score: 3.0 memory length: 71783	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.48		
episode: 382	score: 0.0 memory length: 71906	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 383	score: 2.0 memory length: 72106	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.47		
episode: 384	score: 2.0 memory length: 72304	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.48		
episode: 385	score: 1.0 memory length: 72476	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.47		
episode: 386	score: 0.0 memory length: 72599	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 387	score: 1.0 memory length: 72770	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.47		
episode: 388	score: 2.0 memory length: 72968	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.48		
episode: 389	score: 5.0 memory length: 73303	epsilon: 1.0	steps: 335
lr: 0.0001	evaluation reward: 1.51		
episode: 390	score: 2.0 memory length: 73483	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.53		
episode: 391	score: 3.0 memory length: 73730	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.54		
episode: 392	score: 0.0 memory length: 73853	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.53		
episode: 393	score: 0.0 memory length: 73975	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.53		
episode: 394	score: 0.0 memory length: 74098	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 395	score: 2.0 memory length: 74319	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.54		
episode: 396	score: 0.0 memory length: 74442	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 397	score: 0.0 memory length: 74564	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.51		
episode: 398	score: 1.0 memory length: 74732	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.51		
episode: 399	score: 0.0 memory length: 74855	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 400	score: 0.0 memory length: 74978	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 401	score: 2.0 memory length: 75194	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.47		
episode: 402	score: 0.0 memory length: 75316	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45		
episode: 403	score: 8.0 memory length: 75612	epsilon: 1.0	steps: 296
lr: 0.0001	evaluation reward: 1.53		
episode: 404	score: 3.0 memory length: 75861	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.55		
episode: 405	score: 2.0 memory length: 76077	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.53		
episode: 406	score: 2.0 memory length: 76296	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.54		
episode: 407	score: 0.0 memory length: 76419	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 408	score: 0.0 memory length: 76541	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.49		
episode: 409	score: 0.0 memory length: 76663	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.49		
episode: 410	score: 2.0 memory length: 76845	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.43		
episode: 411	score: 2.0 memory length: 77043	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.44		
episode: 412	score: 3.0 memory length: 77287	epsilon: 1.0	steps: 244
lr: 0.0001	evaluation reward: 1.45		
episode: 413	score: 0.0 memory length: 77410	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45		
episode: 414	score: 0.0 memory length: 77533	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 415	score: 2.0 memory length: 77750	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.44		
episode: 416	score: 1.0 memory length: 77901	epsilon: 1.0	steps: 151

lr: 0.0001	evaluation reward: 1.45			
episode: 417	score: 1.0	memory length: 78052	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.45			
episode: 418	score: 2.0	memory length: 78249	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.47			
episode: 419	score: 2.0	memory length: 78464	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.47			
episode: 420	score: 4.0	memory length: 78740	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.5			
episode: 421	score: 1.0	memory length: 78908	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.51			
episode: 422	score: 2.0	memory length: 79105	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.51			
episode: 423	score: 2.0	memory length: 79287	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.53			
episode: 424	score: 0.0	memory length: 79410	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52			
episode: 425	score: 0.0	memory length: 79532	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.5			
episode: 426	score: 0.0	memory length: 79655	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48			
episode: 427	score: 2.0	memory length: 79853	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49			
episode: 428	score: 1.0	memory length: 80004	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48			
episode: 429	score: 3.0	memory length: 80250	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.51			
episode: 430	score: 1.0	memory length: 80422	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.51			
episode: 431	score: 0.0	memory length: 80545	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 432	score: 1.0	memory length: 80717	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.48			
episode: 433	score: 3.0	memory length: 80944	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.49			
episode: 434	score: 0.0	memory length: 81067	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45			
episode: 435	score: 2.0	memory length: 81265	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.46			
episode: 436	score: 0.0	memory length: 81387	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45			
episode: 437	score: 2.0	memory length: 81603	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.45			
episode: 438	score: 0.0	memory length: 81725	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45			
episode: 439	score: 2.0	memory length: 81943	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.47			
episode: 440	score: 3.0	memory length: 82189	epsilon: 1.0	steps: 246

lr: 0.0001	evaluation reward: 1.5		
episode: 441	score: 1.0 memory length: 82339	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51		
episode: 442	score: 1.0 memory length: 82490	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.52		
episode: 443	score: 0.0 memory length: 82613	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 444	score: 3.0 memory length: 82838	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.47		
episode: 445	score: 1.0 memory length: 82989	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48		
episode: 446	score: 4.0 memory length: 83286	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.5		
episode: 447	score: 0.0 memory length: 83408	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.49		
episode: 448	score: 3.0 memory length: 83634	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.51		
episode: 449	score: 0.0 memory length: 83757	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 450	score: 1.0 memory length: 83929	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.46		
episode: 451	score: 2.0 memory length: 84113	epsilon: 1.0	steps: 184
lr: 0.0001	evaluation reward: 1.47		
episode: 452	score: 2.0 memory length: 84313	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.49		
episode: 453	score: 2.0 memory length: 84532	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.5		
episode: 454	score: 2.0 memory length: 84748	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.5		
episode: 455	score: 8.0 memory length: 85069	epsilon: 1.0	steps: 321
lr: 0.0001	evaluation reward: 1.58		
episode: 456	score: 0.0 memory length: 85191	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.56		
episode: 457	score: 0.0 memory length: 85314	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 458	score: 1.0 memory length: 85482	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.48		
episode: 459	score: 0.0 memory length: 85605	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 460	score: 0.0 memory length: 85728	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 461	score: 4.0 memory length: 86002	epsilon: 1.0	steps: 274
lr: 0.0001	evaluation reward: 1.47		
episode: 462	score: 0.0 memory length: 86124	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45		
episode: 463	score: 4.0 memory length: 86400	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.47		
episode: 464	score: 2.0 memory length: 86599	epsilon: 1.0	steps: 199

lr: 0.0001	evaluation reward: 1.48		
episode: 465	score: 3.0 memory length: 86846	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.51		
episode: 466	score: 1.0 memory length: 86996	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.51		
episode: 467	score: 4.0 memory length: 87257	epsilon: 1.0	steps: 261
lr: 0.0001	evaluation reward: 1.52		
episode: 468	score: 0.0 memory length: 87380	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 469	score: 2.0 memory length: 87599	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.53		
episode: 470	score: 2.0 memory length: 87818	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.53		
episode: 471	score: 1.0 memory length: 87987	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5		
episode: 472	score: 0.0 memory length: 88110	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 473	score: 0.0 memory length: 88232	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.44		
episode: 474	score: 0.0 memory length: 88355	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 475	score: 1.0 memory length: 88526	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.44		
episode: 476	score: 4.0 memory length: 88789	epsilon: 1.0	steps: 263
lr: 0.0001	evaluation reward: 1.48		
episode: 477	score: 1.0 memory length: 88959	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.48		
episode: 478	score: 0.0 memory length: 89081	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.47		
episode: 479	score: 1.0 memory length: 89250	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.46		
episode: 480	score: 3.0 memory length: 89459	epsilon: 1.0	steps: 209
lr: 0.0001	evaluation reward: 1.49		
episode: 481	score: 2.0 memory length: 89679	epsilon: 1.0	steps: 220
lr: 0.0001	evaluation reward: 1.48		
episode: 482	score: 0.0 memory length: 89801	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.48		
episode: 483	score: 2.0 memory length: 89999	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.48		
episode: 484	score: 1.0 memory length: 90167	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.47		
episode: 485	score: 4.0 memory length: 90482	epsilon: 1.0	steps: 315
lr: 0.0001	evaluation reward: 1.5		
episode: 486	score: 1.0 memory length: 90652	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.51		
episode: 487	score: 1.0 memory length: 90804	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.51		
episode: 488	score: 2.0 memory length: 91022	epsilon: 1.0	steps: 218

lr: 0.0001	evaluation reward: 1.51		
episode: 489	score: 0.0 memory length: 91145	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 490	score: 0.0 memory length: 91268	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.44		
episode: 491	score: 3.0 memory length: 91515	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.44		
episode: 492	score: 1.0 memory length: 91666	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.45		
episode: 493	score: 0.0 memory length: 91789	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45		
episode: 494	score: 0.0 memory length: 91912	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.45		
episode: 495	score: 2.0 memory length: 92110	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45		
episode: 496	score: 4.0 memory length: 92376	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.49		
episode: 497	score: 2.0 memory length: 92574	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.51		
episode: 498	score: 1.0 memory length: 92742	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.51		
episode: 499	score: 0.0 memory length: 92864	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.51		
episode: 500	score: 1.0 memory length: 93014	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.52		
episode: 501	score: 5.0 memory length: 93340	epsilon: 1.0	steps: 326
lr: 0.0001	evaluation reward: 1.55		
episode: 502	score: 0.0 memory length: 93463	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.55		
episode: 503	score: 5.0 memory length: 93790	epsilon: 1.0	steps: 327
lr: 0.0001	evaluation reward: 1.52		
episode: 504	score: 3.0 memory length: 94016	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.52		
episode: 505	score: 1.0 memory length: 94187	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.51		
episode: 506	score: 1.0 memory length: 94338	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.5		
episode: 507	score: 0.0 memory length: 94461	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 508	score: 3.0 memory length: 94708	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.53		
episode: 509	score: 2.0 memory length: 94906	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.55		
episode: 510	score: 2.0 memory length: 95124	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.55		
episode: 511	score: 1.0 memory length: 95293	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.54		
episode: 512	score: 0.0 memory length: 95415	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.51			
episode: 513	score: 1.0	memory length: 95565	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.52			
episode: 514	score: 0.0	memory length: 95687	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.52			
episode: 515	score: 0.0	memory length: 95810	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5			
episode: 516	score: 1.0	memory length: 95961	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.5			
episode: 517	score: 2.0	memory length: 96161	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.51			
episode: 518	score: 0.0	memory length: 96284	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49			
episode: 519	score: 4.0	memory length: 96599	epsilon: 1.0	steps: 315
lr: 0.0001	evaluation reward: 1.51			
episode: 520	score: 3.0	memory length: 96826	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.5			
episode: 521	score: 3.0	memory length: 97055	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.52			
episode: 522	score: 3.0	memory length: 97281	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.53			
episode: 523	score: 0.0	memory length: 97403	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.51			
episode: 524	score: 1.0	memory length: 97572	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.52			
episode: 525	score: 2.0	memory length: 97791	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.54			
episode: 526	score: 4.0	memory length: 98068	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.58			
episode: 527	score: 0.0	memory length: 98191	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.56			
episode: 528	score: 3.0	memory length: 98437	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.58			
episode: 529	score: 3.0	memory length: 98686	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.58			
episode: 530	score: 0.0	memory length: 98809	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 531	score: 0.0	memory length: 98932	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 532	score: 1.0	memory length: 99083	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.57			
episode: 533	score: 3.0	memory length: 99334	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.57			
episode: 534	score: 0.0	memory length: 99457	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.57			
episode: 535	score: 3.0	memory length: 99683	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.58			
episode: 536	score: 2.0	memory length: 99880	epsilon: 1.0	steps: 197

```

lr: 0.0001      evaluation reward: 1.6
episode: 537    score: 5.0    memory length: 100204    epsilon: 0.9995941000000088
steps: 324      lr: 0.0001      evaluation reward: 1.63
episode: 538    score: 2.0    memory length: 100402    epsilon: 0.9992020600000173
steps: 198      lr: 0.0001      evaluation reward: 1.65
episode: 539    score: 4.0    memory length: 100694    epsilon: 0.9986239000000299
steps: 292      lr: 0.0001      evaluation reward: 1.67
episode: 540    score: 1.0    memory length: 100864    epsilon: 0.9982873000000372
steps: 170      lr: 0.0001      evaluation reward: 1.65
episode: 541    score: 2.0    memory length: 101066    epsilon: 0.9978873400000459
steps: 202      lr: 0.0001      evaluation reward: 1.66
episode: 542    score: 2.0    memory length: 101284    epsilon: 0.9974557000000552
steps: 218      lr: 0.0001      evaluation reward: 1.67
episode: 543    score: 0.0    memory length: 101407    epsilon: 0.9972121600000605
steps: 123      lr: 0.0001      evaluation reward: 1.67
episode: 544    score: 3.0    memory length: 101653    epsilon: 0.9967250800000711
steps: 246      lr: 0.0001      evaluation reward: 1.67
episode: 545    score: 0.0    memory length: 101776    epsilon: 0.9964815400000764
steps: 123      lr: 0.0001      evaluation reward: 1.66
episode: 546    score: 0.0    memory length: 101899    epsilon: 0.9962380000000817
steps: 123      lr: 0.0001      evaluation reward: 1.62
episode: 547    score: 2.0    memory length: 102116    epsilon: 0.9958083400000091
steps: 217      lr: 0.0001      evaluation reward: 1.64
episode: 548    score: 0.0    memory length: 102238    epsilon: 0.99556678000000962
steps: 122      lr: 0.0001      evaluation reward: 1.61
episode: 549    score: 3.0    memory length: 102483    epsilon: 0.99508168000001068
steps: 245      lr: 0.0001      evaluation reward: 1.64
episode: 550    score: 0.0    memory length: 102606    epsilon: 0.99483814000001121
steps: 123      lr: 0.0001      evaluation reward: 1.63
episode: 551    score: 0.0    memory length: 102729    epsilon: 0.99459460000001173
steps: 123      lr: 0.0001      evaluation reward: 1.61
episode: 552    score: 2.0    memory length: 102927    epsilon: 0.99420256000001259
steps: 198      lr: 0.0001      evaluation reward: 1.61
episode: 553    score: 0.0    memory length: 103049    epsilon: 0.99396100000001311
steps: 122      lr: 0.0001      evaluation reward: 1.59
episode: 554    score: 0.0    memory length: 103172    epsilon: 0.99371746000001364
steps: 123      lr: 0.0001      evaluation reward: 1.57
episode: 555    score: 1.0    memory length: 103341    epsilon: 0.99338284000001437
steps: 169      lr: 0.0001      evaluation reward: 1.5
episode: 556    score: 0.0    memory length: 103464    epsilon: 0.99313930000001489
steps: 123      lr: 0.0001      evaluation reward: 1.5
episode: 557    score: 2.0    memory length: 103681    epsilon: 0.99270964000001583
steps: 217      lr: 0.0001      evaluation reward: 1.52
episode: 558    score: 1.0    memory length: 103832    epsilon: 0.99241066000001648
steps: 151      lr: 0.0001      evaluation reward: 1.52
episode: 559    score: 3.0    memory length: 104057    epsilon: 0.99196516000001744
steps: 225      lr: 0.0001      evaluation reward: 1.55
episode: 560    score: 1.0    memory length: 104208    epsilon: 0.99166618000001809

```

steps: 151	lr: 0.0001	evaluation reward: 1.56
episode: 561	score: 2.0	memory length: 104429 epsilon: 0.9912286000001904
steps: 221	lr: 0.0001	evaluation reward: 1.54
episode: 562	score: 1.0	memory length: 104580 epsilon: 0.9909296200001969
steps: 151	lr: 0.0001	evaluation reward: 1.55
episode: 563	score: 0.0	memory length: 104703 epsilon: 0.9906860800002022
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 564	score: 1.0	memory length: 104871 epsilon: 0.9903534400002094
steps: 168	lr: 0.0001	evaluation reward: 1.5
episode: 565	score: 1.0	memory length: 105041 epsilon: 0.9900168400002167
steps: 170	lr: 0.0001	evaluation reward: 1.48
episode: 566	score: 1.0	memory length: 105210 epsilon: 0.989682220000224
steps: 169	lr: 0.0001	evaluation reward: 1.48
episode: 567	score: 0.0	memory length: 105333 epsilon: 0.9894386800002293
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 568	score: 0.0	memory length: 105456 epsilon: 0.9891951400002346
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 569	score: 2.0	memory length: 105672 epsilon: 0.9887674600002438
steps: 216	lr: 0.0001	evaluation reward: 1.44
episode: 570	score: 3.0	memory length: 105940 epsilon: 0.9882368200002554
steps: 268	lr: 0.0001	evaluation reward: 1.45
episode: 571	score: 1.0	memory length: 106109 epsilon: 0.9879022000002626
steps: 169	lr: 0.0001	evaluation reward: 1.45
episode: 572	score: 2.0	memory length: 106307 epsilon: 0.9875101600002711
steps: 198	lr: 0.0001	evaluation reward: 1.47
episode: 573	score: 2.0	memory length: 106525 epsilon: 0.9870785200002805
steps: 218	lr: 0.0001	evaluation reward: 1.49
episode: 574	score: 3.0	memory length: 106795 epsilon: 0.9865439200002921
steps: 270	lr: 0.0001	evaluation reward: 1.52
episode: 575	score: 1.0	memory length: 106964 epsilon: 0.9862093000002994
steps: 169	lr: 0.0001	evaluation reward: 1.52
episode: 576	score: 1.0	memory length: 107133 epsilon: 0.9858746800003066
steps: 169	lr: 0.0001	evaluation reward: 1.49
episode: 577	score: 0.0	memory length: 107255 epsilon: 0.9856331200003119
steps: 122	lr: 0.0001	evaluation reward: 1.48
episode: 578	score: 3.0	memory length: 107486 epsilon: 0.9851757400003218
steps: 231	lr: 0.0001	evaluation reward: 1.51
episode: 579	score: 1.0	memory length: 107637 epsilon: 0.9848767600003283
steps: 151	lr: 0.0001	evaluation reward: 1.51
episode: 580	score: 0.0	memory length: 107759 epsilon: 0.9846352000003336
steps: 122	lr: 0.0001	evaluation reward: 1.48
episode: 581	score: 2.0	memory length: 107974 epsilon: 0.9842095000003428
steps: 215	lr: 0.0001	evaluation reward: 1.48
episode: 582	score: 0.0	memory length: 108097 epsilon: 0.9839659600003481
steps: 123	lr: 0.0001	evaluation reward: 1.48
episode: 583	score: 0.0	memory length: 108220 epsilon: 0.9837224200003534
steps: 123	lr: 0.0001	evaluation reward: 1.46
episode: 584	score: 2.0	memory length: 108420 epsilon: 0.983326420000362

steps: 200	lr: 0.0001	evaluation reward: 1.47
episode: 585	score: 2.0	memory length: 108602 epsilon: 0.9829660600003698
steps: 182	lr: 0.0001	evaluation reward: 1.45
episode: 586	score: 1.0	memory length: 108770 epsilon: 0.982633420000377
steps: 168	lr: 0.0001	evaluation reward: 1.45
episode: 587	score: 1.0	memory length: 108938 epsilon: 0.9823007800003842
steps: 168	lr: 0.0001	evaluation reward: 1.45
episode: 588	score: 2.0	memory length: 109136 epsilon: 0.9819087400003927
steps: 198	lr: 0.0001	evaluation reward: 1.45
episode: 589	score: 1.0	memory length: 109305 epsilon: 0.9815741200004
steps: 169	lr: 0.0001	evaluation reward: 1.46
episode: 590	score: 1.0	memory length: 109455 epsilon: 0.9812771200004065
steps: 150	lr: 0.0001	evaluation reward: 1.47
episode: 591	score: 2.0	memory length: 109653 epsilon: 0.980885080000415
steps: 198	lr: 0.0001	evaluation reward: 1.46
episode: 592	score: 1.0	memory length: 109823 epsilon: 0.9805484800004223
steps: 170	lr: 0.0001	evaluation reward: 1.46
episode: 593	score: 0.0	memory length: 109945 epsilon: 0.9803069200004275
steps: 122	lr: 0.0001	evaluation reward: 1.46
episode: 594	score: 2.0	memory length: 110143 epsilon: 0.979914880000436
steps: 198	lr: 0.0001	evaluation reward: 1.48
episode: 595	score: 0.0	memory length: 110265 epsilon: 0.9796733200004413
steps: 122	lr: 0.0001	evaluation reward: 1.46
episode: 596	score: 0.0	memory length: 110388 epsilon: 0.9794297800004466
steps: 123	lr: 0.0001	evaluation reward: 1.42
episode: 597	score: 2.0	memory length: 110585 epsilon: 0.979039720000455
steps: 197	lr: 0.0001	evaluation reward: 1.42
episode: 598	score: 0.0	memory length: 110708 epsilon: 0.9787961800004603
steps: 123	lr: 0.0001	evaluation reward: 1.41
episode: 599	score: 2.0	memory length: 110908 epsilon: 0.9784001800004689
steps: 200	lr: 0.0001	evaluation reward: 1.43
episode: 600	score: 0.0	memory length: 111030 epsilon: 0.9781586200004742
steps: 122	lr: 0.0001	evaluation reward: 1.42
episode: 601	score: 2.0	memory length: 111228 epsilon: 0.9777665800004827
steps: 198	lr: 0.0001	evaluation reward: 1.39
episode: 602	score: 7.0	memory length: 111528 epsilon: 0.9771725800004956
steps: 300	lr: 0.0001	evaluation reward: 1.46
episode: 603	score: 0.0	memory length: 111651 epsilon: 0.9769290400005008
steps: 123	lr: 0.0001	evaluation reward: 1.41
episode: 604	score: 1.0	memory length: 111819 epsilon: 0.9765964000005081
steps: 168	lr: 0.0001	evaluation reward: 1.39
episode: 605	score: 5.0	memory length: 112161 epsilon: 0.9759192400005228
steps: 342	lr: 0.0001	evaluation reward: 1.43
episode: 606	score: 0.0	memory length: 112283 epsilon: 0.975677680000528
steps: 122	lr: 0.0001	evaluation reward: 1.42
episode: 607	score: 2.0	memory length: 112481 epsilon: 0.9752856400005365
steps: 198	lr: 0.0001	evaluation reward: 1.44
episode: 608	score: 3.0	memory length: 112725 epsilon: 0.974802520000547

steps: 244	lr: 0.0001	evaluation reward: 1.44
episode: 609	score: 2.0	memory length: 112927 epsilon: 0.9744025600005557
steps: 202	lr: 0.0001	evaluation reward: 1.44
episode: 610	score: 2.0	memory length: 113130 epsilon: 0.9740006200005644
steps: 203	lr: 0.0001	evaluation reward: 1.44
episode: 611	score: 1.0	memory length: 113298 epsilon: 0.9736679800005716
steps: 168	lr: 0.0001	evaluation reward: 1.44
episode: 612	score: 1.0	memory length: 113466 epsilon: 0.9733353400005789
steps: 168	lr: 0.0001	evaluation reward: 1.45
episode: 613	score: 2.0	memory length: 113664 epsilon: 0.9729433000005874
steps: 198	lr: 0.0001	evaluation reward: 1.46
episode: 614	score: 1.0	memory length: 113815 epsilon: 0.9726443200005939
steps: 151	lr: 0.0001	evaluation reward: 1.47
episode: 615	score: 1.0	memory length: 113986 epsilon: 0.9723057400006012
steps: 171	lr: 0.0001	evaluation reward: 1.48
episode: 616	score: 2.0	memory length: 114184 epsilon: 0.9719137000006097
steps: 198	lr: 0.0001	evaluation reward: 1.49
episode: 617	score: 3.0	memory length: 114411 epsilon: 0.9714642400006195
steps: 227	lr: 0.0001	evaluation reward: 1.5
episode: 618	score: 3.0	memory length: 114680 epsilon: 0.970931620000631
steps: 269	lr: 0.0001	evaluation reward: 1.53
episode: 619	score: 4.0	memory length: 114942 epsilon: 0.9704128600006423
steps: 262	lr: 0.0001	evaluation reward: 1.53
episode: 620	score: 9.0	memory length: 115301 epsilon: 0.9697020400006577
steps: 359	lr: 0.0001	evaluation reward: 1.59
episode: 621	score: 2.0	memory length: 115499 epsilon: 0.9693100000006662
steps: 198	lr: 0.0001	evaluation reward: 1.58
episode: 622	score: 0.0	memory length: 115622 epsilon: 0.9690664600006715
steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 623	score: 0.0	memory length: 115744 epsilon: 0.9688249000006768
steps: 122	lr: 0.0001	evaluation reward: 1.55
episode: 624	score: 1.0	memory length: 115912 epsilon: 0.968492260000684
steps: 168	lr: 0.0001	evaluation reward: 1.55
episode: 625	score: 2.0	memory length: 116131 epsilon: 0.9680586400006934
steps: 219	lr: 0.0001	evaluation reward: 1.55
episode: 626	score: 3.0	memory length: 116378 epsilon: 0.967569580000704
steps: 247	lr: 0.0001	evaluation reward: 1.54
episode: 627	score: 4.0	memory length: 116652 epsilon: 0.9670270600007158
steps: 274	lr: 0.0001	evaluation reward: 1.58
episode: 628	score: 2.0	memory length: 116849 epsilon: 0.9666370000007243
steps: 197	lr: 0.0001	evaluation reward: 1.57
episode: 629	score: 3.0	memory length: 117095 epsilon: 0.9661499200007349
steps: 246	lr: 0.0001	evaluation reward: 1.57
episode: 630	score: 2.0	memory length: 117275 epsilon: 0.9657935200007426
steps: 180	lr: 0.0001	evaluation reward: 1.59
episode: 631	score: 0.0	memory length: 117397 epsilon: 0.9655519600007478
steps: 122	lr: 0.0001	evaluation reward: 1.59
episode: 632	score: 1.0	memory length: 117548 epsilon: 0.9652529800007543

steps: 151	lr: 0.0001	evaluation reward: 1.59
episode: 633	score: 3.0	memory length: 117794 epsilon: 0.9647659000007649
steps: 246	lr: 0.0001	evaluation reward: 1.59
episode: 634	score: 0.0	memory length: 117917 epsilon: 0.9645223600007702
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 635	score: 2.0	memory length: 118115 epsilon: 0.9641303200007787
steps: 198	lr: 0.0001	evaluation reward: 1.58
episode: 636	score: 0.0	memory length: 118238 epsilon: 0.963886780000784
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 637	score: 3.0	memory length: 118464 epsilon: 0.9634393000007937
steps: 226	lr: 0.0001	evaluation reward: 1.54
episode: 638	score: 3.0	memory length: 118711 epsilon: 0.9629502400008043
steps: 247	lr: 0.0001	evaluation reward: 1.55
episode: 639	score: 1.0	memory length: 118863 epsilon: 0.9626492800008108
steps: 152	lr: 0.0001	evaluation reward: 1.52
episode: 640	score: 3.0	memory length: 119125 epsilon: 0.9621305200008221
steps: 262	lr: 0.0001	evaluation reward: 1.54
episode: 641	score: 2.0	memory length: 119323 epsilon: 0.9617384800008306
steps: 198	lr: 0.0001	evaluation reward: 1.54
episode: 642	score: 2.0	memory length: 119521 epsilon: 0.9613464400008391
steps: 198	lr: 0.0001	evaluation reward: 1.54
episode: 643	score: 1.0	memory length: 119672 epsilon: 0.9610474600008456
steps: 151	lr: 0.0001	evaluation reward: 1.55
episode: 644	score: 2.0	memory length: 119870 epsilon: 0.9606554200008541
steps: 198	lr: 0.0001	evaluation reward: 1.54
episode: 645	score: 1.0	memory length: 120022 epsilon: 0.9603544600008607
steps: 152	lr: 0.0001	evaluation reward: 1.55
episode: 646	score: 0.0	memory length: 120145 epsilon: 0.960110920000866
steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 647	score: 3.0	memory length: 120411 epsilon: 0.9595842400008774
steps: 266	lr: 0.0001	evaluation reward: 1.56
episode: 648	score: 1.0	memory length: 120583 epsilon: 0.9592436800008848
steps: 172	lr: 0.0001	evaluation reward: 1.57
episode: 649	score: 2.0	memory length: 120763 epsilon: 0.9588872800008925
steps: 180	lr: 0.0001	evaluation reward: 1.56
episode: 650	score: 3.0	memory length: 120990 epsilon: 0.9584378200009023
steps: 227	lr: 0.0001	evaluation reward: 1.59
episode: 651	score: 0.0	memory length: 121113 epsilon: 0.9581942800009076
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 652	score: 0.0	memory length: 121235 epsilon: 0.9579527200009128
steps: 122	lr: 0.0001	evaluation reward: 1.57
episode: 653	score: 2.0	memory length: 121432 epsilon: 0.9575626600009213
steps: 197	lr: 0.0001	evaluation reward: 1.59
episode: 654	score: 2.0	memory length: 121632 epsilon: 0.9571666600009299
steps: 200	lr: 0.0001	evaluation reward: 1.61
episode: 655	score: 0.0	memory length: 121755 epsilon: 0.9569231200009352
steps: 123	lr: 0.0001	evaluation reward: 1.6
episode: 656	score: 1.0	memory length: 121924 epsilon: 0.9565885000009424

steps: 169	lr: 0.0001	evaluation reward: 1.61	
episode: 657	score: 2.0	memory length: 122144	epsilon: 0.9561529000009519
steps: 220	lr: 0.0001	evaluation reward: 1.61	
episode: 658	score: 0.0	memory length: 122267	epsilon: 0.9559093600009572
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 659	score: 0.0	memory length: 122389	epsilon: 0.9556678000009624
steps: 122	lr: 0.0001	evaluation reward: 1.57	
episode: 660	score: 2.0	memory length: 122587	epsilon: 0.9552757600009709
steps: 198	lr: 0.0001	evaluation reward: 1.58	
episode: 661	score: 2.0	memory length: 122808	epsilon: 0.9548381800009804
steps: 221	lr: 0.0001	evaluation reward: 1.58	
episode: 662	score: 2.0	memory length: 123005	epsilon: 0.9544481200009889
steps: 197	lr: 0.0001	evaluation reward: 1.59	
episode: 663	score: 1.0	memory length: 123156	epsilon: 0.9541491400009954
steps: 151	lr: 0.0001	evaluation reward: 1.6	
episode: 664	score: 2.0	memory length: 123373	epsilon: 0.9537194800010047
steps: 217	lr: 0.0001	evaluation reward: 1.61	
episode: 665	score: 0.0	memory length: 123496	epsilon: 0.95347594000101
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 666	score: 1.0	memory length: 123646	epsilon: 0.9531789400010164
steps: 150	lr: 0.0001	evaluation reward: 1.6	
episode: 667	score: 0.0	memory length: 123768	epsilon: 0.9529373800010217
steps: 122	lr: 0.0001	evaluation reward: 1.6	
episode: 668	score: 0.0	memory length: 123891	epsilon: 0.952693840001027
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 669	score: 5.0	memory length: 124180	epsilon: 0.9521216200010394
steps: 289	lr: 0.0001	evaluation reward: 1.63	
episode: 670	score: 1.0	memory length: 124330	epsilon: 0.9518246200010458
steps: 150	lr: 0.0001	evaluation reward: 1.61	
episode: 671	score: 1.0	memory length: 124500	epsilon: 0.9514880200010531
steps: 170	lr: 0.0001	evaluation reward: 1.61	
episode: 672	score: 4.0	memory length: 124776	epsilon: 0.950941540001065
steps: 276	lr: 0.0001	evaluation reward: 1.63	
episode: 673	score: 0.0	memory length: 124899	epsilon: 0.9506980000010703
steps: 123	lr: 0.0001	evaluation reward: 1.61	
episode: 674	score: 1.0	memory length: 125069	epsilon: 0.9503614000010776
steps: 170	lr: 0.0001	evaluation reward: 1.59	
episode: 675	score: 2.0	memory length: 125290	epsilon: 0.9499238200010871
steps: 221	lr: 0.0001	evaluation reward: 1.6	
episode: 676	score: 2.0	memory length: 125488	epsilon: 0.9495317800010956
steps: 198	lr: 0.0001	evaluation reward: 1.61	
episode: 677	score: 1.0	memory length: 125660	epsilon: 0.949191220001103
steps: 172	lr: 0.0001	evaluation reward: 1.62	
episode: 678	score: 0.0	memory length: 125782	epsilon: 0.9489496600011083
steps: 122	lr: 0.0001	evaluation reward: 1.59	
episode: 679	score: 1.0	memory length: 125933	epsilon: 0.9486506800011147
steps: 151	lr: 0.0001	evaluation reward: 1.59	
episode: 680	score: 4.0	memory length: 126249	epsilon: 0.9480250000011283

steps: 316	lr: 0.0001	evaluation reward: 1.63
episode: 681	score: 3.0	memory length: 126497 epsilon: 0.947533960001139
steps: 248	lr: 0.0001	evaluation reward: 1.64
episode: 682	score: 4.0	memory length: 126770 epsilon: 0.9469934200011507
steps: 273	lr: 0.0001	evaluation reward: 1.68
episode: 683	score: 1.0	memory length: 126939 epsilon: 0.946658800001158
steps: 169	lr: 0.0001	evaluation reward: 1.69
episode: 684	score: 2.0	memory length: 127137 epsilon: 0.9462667600011665
steps: 198	lr: 0.0001	evaluation reward: 1.69
episode: 685	score: 1.0	memory length: 127288 epsilon: 0.945967780001173
steps: 151	lr: 0.0001	evaluation reward: 1.68
episode: 686	score: 4.0	memory length: 127542 epsilon: 0.9454648600011839
steps: 254	lr: 0.0001	evaluation reward: 1.71
episode: 687	score: 3.0	memory length: 127772 epsilon: 0.9450094600011938
steps: 230	lr: 0.0001	evaluation reward: 1.73
episode: 688	score: 2.0	memory length: 127990 epsilon: 0.9445778200012032
steps: 218	lr: 0.0001	evaluation reward: 1.73
episode: 689	score: 0.0	memory length: 128113 epsilon: 0.9443342800012084
steps: 123	lr: 0.0001	evaluation reward: 1.72
episode: 690	score: 1.0	memory length: 128264 epsilon: 0.9440353000012149
steps: 151	lr: 0.0001	evaluation reward: 1.72
episode: 691	score: 2.0	memory length: 128462 epsilon: 0.9436432600012234
steps: 198	lr: 0.0001	evaluation reward: 1.72
episode: 692	score: 0.0	memory length: 128584 epsilon: 0.9434017000012287
steps: 122	lr: 0.0001	evaluation reward: 1.71
episode: 693	score: 3.0	memory length: 128834 epsilon: 0.9429067000012394
steps: 250	lr: 0.0001	evaluation reward: 1.74
episode: 694	score: 3.0	memory length: 129082 epsilon: 0.9424156600012501
steps: 248	lr: 0.0001	evaluation reward: 1.75
episode: 695	score: 1.0	memory length: 129232 epsilon: 0.9421186600012565
steps: 150	lr: 0.0001	evaluation reward: 1.76
episode: 696	score: 2.0	memory length: 129430 epsilon: 0.9417266200012651
steps: 198	lr: 0.0001	evaluation reward: 1.78
episode: 697	score: 0.0	memory length: 129553 epsilon: 0.9414830800012703
steps: 123	lr: 0.0001	evaluation reward: 1.76
episode: 698	score: 3.0	memory length: 129820 epsilon: 0.9409544200012818
steps: 267	lr: 0.0001	evaluation reward: 1.79
episode: 699	score: 3.0	memory length: 130047 epsilon: 0.9405049600012916
steps: 227	lr: 0.0001	evaluation reward: 1.8
episode: 700	score: 2.0	memory length: 130265 epsilon: 0.940073320001301
steps: 218	lr: 0.0001	evaluation reward: 1.82
episode: 701	score: 4.0	memory length: 130520 epsilon: 0.9395684200013119
steps: 255	lr: 0.0001	evaluation reward: 1.84
episode: 702	score: 1.0	memory length: 130671 epsilon: 0.9392694400013184
steps: 151	lr: 0.0001	evaluation reward: 1.78
episode: 703	score: 3.0	memory length: 130899 epsilon: 0.9388180000013282
steps: 228	lr: 0.0001	evaluation reward: 1.81
episode: 704	score: 3.0	memory length: 131164 epsilon: 0.9382933000013396

steps: 265	lr: 0.0001	evaluation reward: 1.83
episode: 705	score: 3.0	memory length: 131374 epsilon: 0.9378775000013486
steps: 210	lr: 0.0001	evaluation reward: 1.81
episode: 706	score: 0.0	memory length: 131496 epsilon: 0.9376359400013539
steps: 122	lr: 0.0001	evaluation reward: 1.81
episode: 707	score: 2.0	memory length: 131694 epsilon: 0.9372439000013624
steps: 198	lr: 0.0001	evaluation reward: 1.81
episode: 708	score: 0.0	memory length: 131816 epsilon: 0.9370023400013676
steps: 122	lr: 0.0001	evaluation reward: 1.78
episode: 709	score: 2.0	memory length: 132014 epsilon: 0.9366103000013761
steps: 198	lr: 0.0001	evaluation reward: 1.78
episode: 710	score: 1.0	memory length: 132183 epsilon: 0.9362756800013834
steps: 169	lr: 0.0001	evaluation reward: 1.77
episode: 711	score: 5.0	memory length: 132527 epsilon: 0.9355945600013982
steps: 344	lr: 0.0001	evaluation reward: 1.81
episode: 712	score: 2.0	memory length: 132743 epsilon: 0.9351668800014075
steps: 216	lr: 0.0001	evaluation reward: 1.82
episode: 713	score: 2.0	memory length: 132940 epsilon: 0.9347768200014159
steps: 197	lr: 0.0001	evaluation reward: 1.82
episode: 714	score: 0.0	memory length: 133062 epsilon: 0.9345352600014212
steps: 122	lr: 0.0001	evaluation reward: 1.81
episode: 715	score: 1.0	memory length: 133233 epsilon: 0.9341966800014285
steps: 171	lr: 0.0001	evaluation reward: 1.81
episode: 716	score: 1.0	memory length: 133384 epsilon: 0.933897700001435
steps: 151	lr: 0.0001	evaluation reward: 1.8
episode: 717	score: 0.0	memory length: 133507 epsilon: 0.9336541600014403
steps: 123	lr: 0.0001	evaluation reward: 1.77
episode: 718	score: 1.0	memory length: 133678 epsilon: 0.9333155800014477
steps: 171	lr: 0.0001	evaluation reward: 1.75
episode: 719	score: 0.0	memory length: 133801 epsilon: 0.9330720400014529
steps: 123	lr: 0.0001	evaluation reward: 1.71
episode: 720	score: 3.0	memory length: 134068 epsilon: 0.9325433800014644
steps: 267	lr: 0.0001	evaluation reward: 1.65
episode: 721	score: 1.0	memory length: 134218 epsilon: 0.9322463800014709
steps: 150	lr: 0.0001	evaluation reward: 1.64
episode: 722	score: 2.0	memory length: 134418 epsilon: 0.9318503800014795
steps: 200	lr: 0.0001	evaluation reward: 1.66
episode: 723	score: 1.0	memory length: 134568 epsilon: 0.9315533800014859
steps: 150	lr: 0.0001	evaluation reward: 1.67
episode: 724	score: 4.0	memory length: 134863 epsilon: 0.9309692800014986
steps: 295	lr: 0.0001	evaluation reward: 1.7
episode: 725	score: 0.0	memory length: 134985 epsilon: 0.9307277200015038
steps: 122	lr: 0.0001	evaluation reward: 1.68
episode: 726	score: 2.0	memory length: 135183 epsilon: 0.9303356800015123
steps: 198	lr: 0.0001	evaluation reward: 1.67
episode: 727	score: 1.0	memory length: 135335 epsilon: 0.9300347200015189
steps: 152	lr: 0.0001	evaluation reward: 1.64
episode: 728	score: 4.0	memory length: 135629 epsilon: 0.9294526000015315

steps: 294	lr: 0.0001	evaluation reward: 1.66
episode: 729	score: 1.0	memory length: 135780 epsilon: 0.929153620001538
steps: 151	lr: 0.0001	evaluation reward: 1.64
episode: 730	score: 0.0	memory length: 135903 epsilon: 0.9289100800015433
steps: 123	lr: 0.0001	evaluation reward: 1.62
episode: 731	score: 2.0	memory length: 136100 epsilon: 0.9285200200015518
steps: 197	lr: 0.0001	evaluation reward: 1.64
episode: 732	score: 5.0	memory length: 136410 epsilon: 0.9279062200015651
steps: 310	lr: 0.0001	evaluation reward: 1.68
episode: 733	score: 2.0	memory length: 136608 epsilon: 0.9275141800015736
steps: 198	lr: 0.0001	evaluation reward: 1.67
episode: 734	score: 0.0	memory length: 136730 epsilon: 0.9272726200015788
steps: 122	lr: 0.0001	evaluation reward: 1.67
episode: 735	score: 3.0	memory length: 136974 epsilon: 0.9267895000015893
steps: 244	lr: 0.0001	evaluation reward: 1.68
episode: 736	score: 2.0	memory length: 137191 epsilon: 0.9263598400015987
steps: 217	lr: 0.0001	evaluation reward: 1.7
episode: 737	score: 2.0	memory length: 137409 epsilon: 0.925928200001608
steps: 218	lr: 0.0001	evaluation reward: 1.69
episode: 738	score: 2.0	memory length: 137611 epsilon: 0.9255282400016167
steps: 202	lr: 0.0001	evaluation reward: 1.68
episode: 739	score: 2.0	memory length: 137809 epsilon: 0.9251362000016252
steps: 198	lr: 0.0001	evaluation reward: 1.69
episode: 740	score: 0.0	memory length: 137932 epsilon: 0.9248926600016305
steps: 123	lr: 0.0001	evaluation reward: 1.66
episode: 741	score: 1.0	memory length: 138101 epsilon: 0.9245580400016378
steps: 169	lr: 0.0001	evaluation reward: 1.65
episode: 742	score: 0.0	memory length: 138223 epsilon: 0.924316480001643
steps: 122	lr: 0.0001	evaluation reward: 1.63
episode: 743	score: 2.0	memory length: 138423 epsilon: 0.9239204800016516
steps: 200	lr: 0.0001	evaluation reward: 1.64
episode: 744	score: 2.0	memory length: 138621 epsilon: 0.9235284400016601
steps: 198	lr: 0.0001	evaluation reward: 1.64
episode: 745	score: 0.0	memory length: 138744 epsilon: 0.9232849000016654
steps: 123	lr: 0.0001	evaluation reward: 1.63
episode: 746	score: 3.0	memory length: 138989 epsilon: 0.9227998000016759
steps: 245	lr: 0.0001	evaluation reward: 1.66
episode: 747	score: 0.0	memory length: 139111 epsilon: 0.9225582400016812
steps: 122	lr: 0.0001	evaluation reward: 1.63
episode: 748	score: 1.0	memory length: 139279 epsilon: 0.9222256000016884
steps: 168	lr: 0.0001	evaluation reward: 1.63
episode: 749	score: 2.0	memory length: 139477 epsilon: 0.9218335600016969
steps: 198	lr: 0.0001	evaluation reward: 1.63
episode: 750	score: 1.0	memory length: 139646 epsilon: 0.9214989400017042
steps: 169	lr: 0.0001	evaluation reward: 1.61
episode: 751	score: 0.0	memory length: 139769 epsilon: 0.9212554000017095
steps: 123	lr: 0.0001	evaluation reward: 1.61
episode: 752	score: 2.0	memory length: 139988 epsilon: 0.9208217800017189

steps: 219	lr: 0.0001	evaluation reward: 1.63
episode: 753	score: 2.0	memory length: 140185 epsilon: 0.9204317200017273
steps: 197	lr: 0.0001	evaluation reward: 1.63
episode: 754	score: 0.0	memory length: 140308 epsilon: 0.9201881800017326
steps: 123	lr: 0.0001	evaluation reward: 1.61
episode: 755	score: 1.0	memory length: 140458 epsilon: 0.9198911800017391
steps: 150	lr: 0.0001	evaluation reward: 1.62
episode: 756	score: 1.0	memory length: 140627 epsilon: 0.9195565600017463
steps: 169	lr: 0.0001	evaluation reward: 1.62
episode: 757	score: 1.0	memory length: 140778 epsilon: 0.9192575800017528
steps: 151	lr: 0.0001	evaluation reward: 1.61
episode: 758	score: 2.0	memory length: 140981 epsilon: 0.9188556400017616
steps: 203	lr: 0.0001	evaluation reward: 1.63
episode: 759	score: 2.0	memory length: 141179 epsilon: 0.9184636000017701
steps: 198	lr: 0.0001	evaluation reward: 1.65
episode: 760	score: 2.0	memory length: 141376 epsilon: 0.9180735400017785
steps: 197	lr: 0.0001	evaluation reward: 1.65
episode: 761	score: 5.0	memory length: 141710 epsilon: 0.9174122200017929
steps: 334	lr: 0.0001	evaluation reward: 1.68
episode: 762	score: 2.0	memory length: 141907 epsilon: 0.9170221600018014
steps: 197	lr: 0.0001	evaluation reward: 1.68
episode: 763	score: 1.0	memory length: 142077 epsilon: 0.9166855600018087
steps: 170	lr: 0.0001	evaluation reward: 1.68
episode: 764	score: 1.0	memory length: 142246 epsilon: 0.9163509400018159
steps: 169	lr: 0.0001	evaluation reward: 1.67
episode: 765	score: 2.0	memory length: 142446 epsilon: 0.9159549400018245
steps: 200	lr: 0.0001	evaluation reward: 1.69
episode: 766	score: 0.0	memory length: 142569 epsilon: 0.9157114000018298
steps: 123	lr: 0.0001	evaluation reward: 1.68
episode: 767	score: 0.0	memory length: 142691 epsilon: 0.9154698400018351
steps: 122	lr: 0.0001	evaluation reward: 1.68
episode: 768	score: 4.0	memory length: 142964 epsilon: 0.9149293000018468
steps: 273	lr: 0.0001	evaluation reward: 1.72
episode: 769	score: 0.0	memory length: 143087 epsilon: 0.9146857600018521
steps: 123	lr: 0.0001	evaluation reward: 1.67
episode: 770	score: 1.0	memory length: 143258 epsilon: 0.9143471800018594
steps: 171	lr: 0.0001	evaluation reward: 1.67
episode: 771	score: 1.0	memory length: 143427 epsilon: 0.9140125600018667
steps: 169	lr: 0.0001	evaluation reward: 1.67
episode: 772	score: 3.0	memory length: 143674 epsilon: 0.9135235000018773
steps: 247	lr: 0.0001	evaluation reward: 1.66
episode: 773	score: 1.0	memory length: 143844 epsilon: 0.9131869000018846
steps: 170	lr: 0.0001	evaluation reward: 1.67
episode: 774	score: 0.0	memory length: 143966 epsilon: 0.9129453400018899
steps: 122	lr: 0.0001	evaluation reward: 1.66
episode: 775	score: 0.0	memory length: 144089 epsilon: 0.9127018000018952
steps: 123	lr: 0.0001	evaluation reward: 1.64
episode: 776	score: 3.0	memory length: 144337 epsilon: 0.9122107600019058

steps: 248	lr: 0.0001	evaluation reward: 1.65
episode: 777	score: 1.0	memory length: 144506 epsilon: 0.9118761400019131
steps: 169	lr: 0.0001	evaluation reward: 1.65
episode: 778	score: 3.0	memory length: 144751 epsilon: 0.9113910400019236
steps: 245	lr: 0.0001	evaluation reward: 1.68
episode: 779	score: 0.0	memory length: 144874 epsilon: 0.9111475000019289
steps: 123	lr: 0.0001	evaluation reward: 1.67
episode: 780	score: 2.0	memory length: 145072 epsilon: 0.9107554600019374
steps: 198	lr: 0.0001	evaluation reward: 1.65
episode: 781	score: 0.0	memory length: 145195 epsilon: 0.9105119200019427
steps: 123	lr: 0.0001	evaluation reward: 1.62
episode: 782	score: 2.0	memory length: 145412 epsilon: 0.910082260001952
steps: 217	lr: 0.0001	evaluation reward: 1.6
episode: 783	score: 1.0	memory length: 145581 epsilon: 0.9097476400019593
steps: 169	lr: 0.0001	evaluation reward: 1.6
episode: 784	score: 2.0	memory length: 145799 epsilon: 0.9093160000019687
steps: 218	lr: 0.0001	evaluation reward: 1.6
episode: 785	score: 4.0	memory length: 146115 epsilon: 0.9086903200019822
steps: 316	lr: 0.0001	evaluation reward: 1.63
episode: 786	score: 0.0	memory length: 146238 epsilon: 0.9084467800019875
steps: 123	lr: 0.0001	evaluation reward: 1.59
episode: 787	score: 0.0	memory length: 146361 epsilon: 0.9082032400019928
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 788	score: 3.0	memory length: 146608 epsilon: 0.9077141800020034
steps: 247	lr: 0.0001	evaluation reward: 1.57
episode: 789	score: 0.0	memory length: 146731 epsilon: 0.9074706400020087
steps: 123	lr: 0.0001	evaluation reward: 1.57
episode: 790	score: 3.0	memory length: 146957 epsilon: 0.9070231600020184
steps: 226	lr: 0.0001	evaluation reward: 1.59
episode: 791	score: 0.0	memory length: 147080 epsilon: 0.9067796200020237
steps: 123	lr: 0.0001	evaluation reward: 1.57
episode: 792	score: 2.0	memory length: 147301 epsilon: 0.9063420400020332
steps: 221	lr: 0.0001	evaluation reward: 1.59
episode: 793	score: 3.0	memory length: 147550 epsilon: 0.9058490200020439
steps: 249	lr: 0.0001	evaluation reward: 1.59
episode: 794	score: 2.0	memory length: 147769 epsilon: 0.9054154000020533
steps: 219	lr: 0.0001	evaluation reward: 1.58
episode: 795	score: 0.0	memory length: 147892 epsilon: 0.9051718600020586
steps: 123	lr: 0.0001	evaluation reward: 1.57
episode: 796	score: 3.0	memory length: 148118 epsilon: 0.9047243800020683
steps: 226	lr: 0.0001	evaluation reward: 1.58
episode: 797	score: 2.0	memory length: 148316 epsilon: 0.9043323400020769
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 798	score: 0.0	memory length: 148438 epsilon: 0.9040907800020821
steps: 122	lr: 0.0001	evaluation reward: 1.57
episode: 799	score: 0.0	memory length: 148561 epsilon: 0.9038472400020874
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 800	score: 1.0	memory length: 148733 epsilon: 0.9035066800020948

steps: 172	lr: 0.0001	evaluation reward: 1.53
episode: 801	score: 0.0	memory length: 148855 epsilon: 0.9032651200021
steps: 122	lr: 0.0001	evaluation reward: 1.49
episode: 802	score: 1.0	memory length: 149006 epsilon: 0.9029661400021065
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 803	score: 2.0	memory length: 149204 epsilon: 0.902574100002115
steps: 198	lr: 0.0001	evaluation reward: 1.48
episode: 804	score: 0.0	memory length: 149327 epsilon: 0.9023305600021203
steps: 123	lr: 0.0001	evaluation reward: 1.45
episode: 805	score: 2.0	memory length: 149545 epsilon: 0.9018989200021297
steps: 218	lr: 0.0001	evaluation reward: 1.44
episode: 806	score: 2.0	memory length: 149742 epsilon: 0.9015088600021381
steps: 197	lr: 0.0001	evaluation reward: 1.46
episode: 807	score: 0.0	memory length: 149865 epsilon: 0.9012653200021434
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 808	score: 2.0	memory length: 150081 epsilon: 0.9008376400021527
steps: 216	lr: 0.0001	evaluation reward: 1.46
episode: 809	score: 1.0	memory length: 150250 epsilon: 0.90050302000216
steps: 169	lr: 0.0001	evaluation reward: 1.45
episode: 810	score: 1.0	memory length: 150419 epsilon: 0.9001684000021672
steps: 169	lr: 0.0001	evaluation reward: 1.45
episode: 811	score: 2.0	memory length: 150617 epsilon: 0.8997763600021758
steps: 198	lr: 0.0001	evaluation reward: 1.42
episode: 812	score: 1.0	memory length: 150768 epsilon: 0.8994773800021822
steps: 151	lr: 0.0001	evaluation reward: 1.41
episode: 813	score: 2.0	memory length: 150966 epsilon: 0.8990853400021908
steps: 198	lr: 0.0001	evaluation reward: 1.41
episode: 814	score: 1.0	memory length: 151135 epsilon: 0.898750720002198
steps: 169	lr: 0.0001	evaluation reward: 1.42
episode: 815	score: 0.0	memory length: 151258 epsilon: 0.8985071800022033
steps: 123	lr: 0.0001	evaluation reward: 1.41
episode: 816	score: 0.0	memory length: 151381 epsilon: 0.8982636400022086
steps: 123	lr: 0.0001	evaluation reward: 1.4
episode: 817	score: 3.0	memory length: 151646 epsilon: 0.89773894000222
steps: 265	lr: 0.0001	evaluation reward: 1.43
episode: 818	score: 0.0	memory length: 151769 epsilon: 0.8974954000022253
steps: 123	lr: 0.0001	evaluation reward: 1.42
episode: 819	score: 1.0	memory length: 151919 epsilon: 0.8971984000022317
steps: 150	lr: 0.0001	evaluation reward: 1.43
episode: 820	score: 0.0	memory length: 152042 epsilon: 0.896954860002237
steps: 123	lr: 0.0001	evaluation reward: 1.4
episode: 821	score: 2.0	memory length: 152240 epsilon: 0.8965628200022455
steps: 198	lr: 0.0001	evaluation reward: 1.41
episode: 822	score: 2.0	memory length: 152441 epsilon: 0.8961648400022542
steps: 201	lr: 0.0001	evaluation reward: 1.41
episode: 823	score: 4.0	memory length: 152715 epsilon: 0.8956223200022659
steps: 274	lr: 0.0001	evaluation reward: 1.44
episode: 824	score: 2.0	memory length: 152913 epsilon: 0.8952302800022744

steps: 198	lr: 0.0001	evaluation reward: 1.42
episode: 825	score: 0.0	memory length: 153036 epsilon: 0.8949867400022797
steps: 123	lr: 0.0001	evaluation reward: 1.42
episode: 826	score: 3.0	memory length: 153283 epsilon: 0.8944976800022904
steps: 247	lr: 0.0001	evaluation reward: 1.43
episode: 827	score: 0.0	memory length: 153406 epsilon: 0.8942541400022956
steps: 123	lr: 0.0001	evaluation reward: 1.42
episode: 828	score: 0.0	memory length: 153528 epsilon: 0.8940125800023009
steps: 122	lr: 0.0001	evaluation reward: 1.38
episode: 829	score: 3.0	memory length: 153771 epsilon: 0.8935314400023113
steps: 243	lr: 0.0001	evaluation reward: 1.4
episode: 830	score: 1.0	memory length: 153943 epsilon: 0.8931908800023187
steps: 172	lr: 0.0001	evaluation reward: 1.41
episode: 831	score: 2.0	memory length: 154140 epsilon: 0.8928008200023272
steps: 197	lr: 0.0001	evaluation reward: 1.41
episode: 832	score: 0.0	memory length: 154263 epsilon: 0.8925572800023325
steps: 123	lr: 0.0001	evaluation reward: 1.36
episode: 833	score: 1.0	memory length: 154414 epsilon: 0.8922583000023339
steps: 151	lr: 0.0001	evaluation reward: 1.35
episode: 834	score: 1.0	memory length: 154584 epsilon: 0.8919217000023463
steps: 170	lr: 0.0001	evaluation reward: 1.36
episode: 835	score: 3.0	memory length: 154812 epsilon: 0.8914702600023561
steps: 228	lr: 0.0001	evaluation reward: 1.36
episode: 836	score: 0.0	memory length: 154934 epsilon: 0.8912287000023613
steps: 122	lr: 0.0001	evaluation reward: 1.34
episode: 837	score: 0.0	memory length: 155056 epsilon: 0.8909871400023666
steps: 122	lr: 0.0001	evaluation reward: 1.32
episode: 838	score: 2.0	memory length: 155274 epsilon: 0.8905555000023759
steps: 218	lr: 0.0001	evaluation reward: 1.32
episode: 839	score: 3.0	memory length: 155501 epsilon: 0.8901060400023857
steps: 227	lr: 0.0001	evaluation reward: 1.33
episode: 840	score: 1.0	memory length: 155670 epsilon: 0.889771420002393
steps: 169	lr: 0.0001	evaluation reward: 1.34
episode: 841	score: 0.0	memory length: 155792 epsilon: 0.8895298600023982
steps: 122	lr: 0.0001	evaluation reward: 1.33
episode: 842	score: 9.0	memory length: 156252 epsilon: 0.888619060002418
steps: 460	lr: 0.0001	evaluation reward: 1.42
episode: 843	score: 1.0	memory length: 156423 epsilon: 0.8882804800024253
steps: 171	lr: 0.0001	evaluation reward: 1.41
episode: 844	score: 1.0	memory length: 156595 epsilon: 0.8879399200024327
steps: 172	lr: 0.0001	evaluation reward: 1.4
episode: 845	score: 5.0	memory length: 156914 epsilon: 0.8873083000024464
steps: 319	lr: 0.0001	evaluation reward: 1.45
episode: 846	score: 2.0	memory length: 157112 epsilon: 0.8869162600024549
steps: 198	lr: 0.0001	evaluation reward: 1.44
episode: 847	score: 1.0	memory length: 157263 epsilon: 0.8866172800024614
steps: 151	lr: 0.0001	evaluation reward: 1.45
episode: 848	score: 2.0	memory length: 157460 epsilon: 0.8862272200024699

steps: 197	lr: 0.0001	evaluation reward: 1.46
episode: 849	score: 2.0	memory length: 157657 epsilon: 0.8858371600024784
steps: 197	lr: 0.0001	evaluation reward: 1.46
episode: 850	score: 2.0	memory length: 157854 epsilon: 0.8854471000024868
steps: 197	lr: 0.0001	evaluation reward: 1.47
episode: 851	score: 2.0	memory length: 158072 epsilon: 0.8850154600024962
steps: 218	lr: 0.0001	evaluation reward: 1.49
episode: 852	score: 2.0	memory length: 158272 epsilon: 0.8846194600025048
steps: 200	lr: 0.0001	evaluation reward: 1.49
episode: 853	score: 2.0	memory length: 158490 epsilon: 0.8841878200025142
steps: 218	lr: 0.0001	evaluation reward: 1.49
episode: 854	score: 0.0	memory length: 158612 epsilon: 0.8839462600025194
steps: 122	lr: 0.0001	evaluation reward: 1.49
episode: 855	score: 0.0	memory length: 158734 epsilon: 0.8837047000025247
steps: 122	lr: 0.0001	evaluation reward: 1.48
episode: 856	score: 0.0	memory length: 158856 epsilon: 0.8834631400025299
steps: 122	lr: 0.0001	evaluation reward: 1.47
episode: 857	score: 2.0	memory length: 159038 epsilon: 0.8831027800025377
steps: 182	lr: 0.0001	evaluation reward: 1.48
episode: 858	score: 2.0	memory length: 159238 epsilon: 0.8827067800025463
steps: 200	lr: 0.0001	evaluation reward: 1.48
episode: 859	score: 1.0	memory length: 159389 epsilon: 0.8824078000025528
steps: 151	lr: 0.0001	evaluation reward: 1.47
episode: 860	score: 2.0	memory length: 159607 epsilon: 0.8819761600025622
steps: 218	lr: 0.0001	evaluation reward: 1.47
episode: 861	score: 4.0	memory length: 159900 epsilon: 0.8813960200025748
steps: 293	lr: 0.0001	evaluation reward: 1.46
episode: 862	score: 4.0	memory length: 160155 epsilon: 0.8808911200025857
steps: 255	lr: 0.0001	evaluation reward: 1.48
episode: 863	score: 2.0	memory length: 160373 epsilon: 0.8804594800025951
steps: 218	lr: 0.0001	evaluation reward: 1.49
episode: 864	score: 6.0	memory length: 160746 epsilon: 0.8797209400026111
steps: 373	lr: 0.0001	evaluation reward: 1.54
episode: 865	score: 1.0	memory length: 160898 epsilon: 0.8794199800026177
steps: 152	lr: 0.0001	evaluation reward: 1.53
episode: 866	score: 2.0	memory length: 161113 epsilon: 0.8789942800026269
steps: 215	lr: 0.0001	evaluation reward: 1.55
episode: 867	score: 2.0	memory length: 161331 epsilon: 0.8785626400026363
steps: 218	lr: 0.0001	evaluation reward: 1.57
episode: 868	score: 0.0	memory length: 161454 epsilon: 0.8783191000026416
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 869	score: 0.0	memory length: 161576 epsilon: 0.8780775400026468
steps: 122	lr: 0.0001	evaluation reward: 1.53
episode: 870	score: 2.0	memory length: 161775 epsilon: 0.8776835200026554
steps: 199	lr: 0.0001	evaluation reward: 1.54
episode: 871	score: 2.0	memory length: 161973 epsilon: 0.8772914800026639
steps: 198	lr: 0.0001	evaluation reward: 1.55
episode: 872	score: 3.0	memory length: 162220 epsilon: 0.8768024200026745

steps: 247	lr: 0.0001	evaluation reward: 1.55
episode: 873	score: 3.0	memory length: 162486 epsilon: 0.8762757400026859
steps: 266	lr: 0.0001	evaluation reward: 1.57
episode: 874	score: 0.0	memory length: 162609 epsilon: 0.8760322000026912
steps: 123	lr: 0.0001	evaluation reward: 1.57
episode: 875	score: 1.0	memory length: 162778 epsilon: 0.8756975800026985
steps: 169	lr: 0.0001	evaluation reward: 1.58
episode: 876	score: 0.0	memory length: 162900 epsilon: 0.8754560200027037
steps: 122	lr: 0.0001	evaluation reward: 1.55
episode: 877	score: 3.0	memory length: 163147 epsilon: 0.8749669600027143
steps: 247	lr: 0.0001	evaluation reward: 1.57
episode: 878	score: 0.0	memory length: 163270 epsilon: 0.8747234200027196
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 879	score: 2.0	memory length: 163468 epsilon: 0.8743313800027281
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 880	score: 2.0	memory length: 163666 epsilon: 0.8739393400027367
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 881	score: 1.0	memory length: 163817 epsilon: 0.8736403600027431
steps: 151	lr: 0.0001	evaluation reward: 1.57
episode: 882	score: 1.0	memory length: 163988 epsilon: 0.8733017800027505
steps: 171	lr: 0.0001	evaluation reward: 1.56
episode: 883	score: 0.0	memory length: 164111 epsilon: 0.8730582400027558
steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 884	score: 2.0	memory length: 164312 epsilon: 0.8726602600027644
steps: 201	lr: 0.0001	evaluation reward: 1.55
episode: 885	score: 0.0	memory length: 164435 epsilon: 0.8724167200027697
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 886	score: 1.0	memory length: 164605 epsilon: 0.872080120002777
steps: 170	lr: 0.0001	evaluation reward: 1.52
episode: 887	score: 1.0	memory length: 164777 epsilon: 0.8717395600027844
steps: 172	lr: 0.0001	evaluation reward: 1.53
episode: 888	score: 0.0	memory length: 164900 epsilon: 0.8714960200027897
steps: 123	lr: 0.0001	evaluation reward: 1.5
episode: 889	score: 1.0	memory length: 165069 epsilon: 0.871161400002797
steps: 169	lr: 0.0001	evaluation reward: 1.51
episode: 890	score: 0.0	memory length: 165192 epsilon: 0.8709178600028022
steps: 123	lr: 0.0001	evaluation reward: 1.48
episode: 891	score: 2.0	memory length: 165390 epsilon: 0.8705258200028108
steps: 198	lr: 0.0001	evaluation reward: 1.5
episode: 892	score: 0.0	memory length: 165512 epsilon: 0.870284260002816
steps: 122	lr: 0.0001	evaluation reward: 1.48
episode: 893	score: 2.0	memory length: 165728 epsilon: 0.8698565800028253
steps: 216	lr: 0.0001	evaluation reward: 1.47
episode: 894	score: 3.0	memory length: 165974 epsilon: 0.8693695000028359
steps: 246	lr: 0.0001	evaluation reward: 1.48
episode: 895	score: 2.0	memory length: 166192 epsilon: 0.8689378600028452
steps: 218	lr: 0.0001	evaluation reward: 1.5
episode: 896	score: 1.0	memory length: 166361 epsilon: 0.8686032400028525

steps: 169	lr: 0.0001	evaluation reward: 1.48
episode: 897	score: 2.0	memory length: 166559 epsilon: 0.868211200002861
steps: 198	lr: 0.0001	evaluation reward: 1.48
episode: 898	score: 2.0	memory length: 166777 epsilon: 0.8677795600028704
steps: 218	lr: 0.0001	evaluation reward: 1.5
episode: 899	score: 2.0	memory length: 166996 epsilon: 0.8673459400028798
steps: 219	lr: 0.0001	evaluation reward: 1.52
episode: 900	score: 2.0	memory length: 167212 epsilon: 0.8669182600028891
steps: 216	lr: 0.0001	evaluation reward: 1.53
episode: 901	score: 2.0	memory length: 167430 epsilon: 0.8664866200028984
steps: 218	lr: 0.0001	evaluation reward: 1.55
episode: 902	score: 0.0	memory length: 167553 epsilon: 0.8662430800029037
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 903	score: 2.0	memory length: 167733 epsilon: 0.8658866800029115
steps: 180	lr: 0.0001	evaluation reward: 1.54
episode: 904	score: 2.0	memory length: 167931 epsilon: 0.86549464000292
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 905	score: 1.0	memory length: 168100 epsilon: 0.8651600200029272
steps: 169	lr: 0.0001	evaluation reward: 1.55
episode: 906	score: 2.0	memory length: 168319 epsilon: 0.8647264000029367
steps: 219	lr: 0.0001	evaluation reward: 1.55
episode: 907	score: 3.0	memory length: 168588 epsilon: 0.8641937800029482
steps: 269	lr: 0.0001	evaluation reward: 1.58
episode: 908	score: 0.0	memory length: 168711 epsilon: 0.8639502400029535
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 909	score: 0.0	memory length: 168834 epsilon: 0.8637067000029588
steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 910	score: 3.0	memory length: 169080 epsilon: 0.8632196200029694
steps: 246	lr: 0.0001	evaluation reward: 1.57
episode: 911	score: 4.0	memory length: 169360 epsilon: 0.8626652200029814
steps: 280	lr: 0.0001	evaluation reward: 1.59
episode: 912	score: 2.0	memory length: 169561 epsilon: 0.86226724000299
steps: 201	lr: 0.0001	evaluation reward: 1.6
episode: 913	score: 3.0	memory length: 169786 epsilon: 0.8618217400029997
steps: 225	lr: 0.0001	evaluation reward: 1.61
episode: 914	score: 2.0	memory length: 169984 epsilon: 0.8614297000030082
steps: 198	lr: 0.0001	evaluation reward: 1.62
episode: 915	score: 0.0	memory length: 170107 epsilon: 0.8611861600030135
steps: 123	lr: 0.0001	evaluation reward: 1.62
episode: 916	score: 0.0	memory length: 170229 epsilon: 0.8609446000030188
steps: 122	lr: 0.0001	evaluation reward: 1.62
episode: 917	score: 1.0	memory length: 170397 epsilon: 0.860611960003026
steps: 168	lr: 0.0001	evaluation reward: 1.6
episode: 918	score: 1.0	memory length: 170550 epsilon: 0.8603090200030326
steps: 153	lr: 0.0001	evaluation reward: 1.61
episode: 919	score: 0.0	memory length: 170673 epsilon: 0.8600654800030378
steps: 123	lr: 0.0001	evaluation reward: 1.6
episode: 920	score: 1.0	memory length: 170844 epsilon: 0.8597269000030452

steps: 171	lr: 0.0001	evaluation reward: 1.61
episode: 921	score: 5.0	memory length: 171155 epsilon: 0.8591111200030586
steps: 311	lr: 0.0001	evaluation reward: 1.64
episode: 922	score: 0.0	memory length: 171277 epsilon: 0.8588695600030638
steps: 122	lr: 0.0001	evaluation reward: 1.62
episode: 923	score: 1.0	memory length: 171446 epsilon: 0.8585349400030711
steps: 169	lr: 0.0001	evaluation reward: 1.59
episode: 924	score: 1.0	memory length: 171597 epsilon: 0.8582359600030776
steps: 151	lr: 0.0001	evaluation reward: 1.58
episode: 925	score: 0.0	memory length: 171720 epsilon: 0.8579924200030828
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 926	score: 0.0	memory length: 171842 epsilon: 0.8577508600030881
steps: 122	lr: 0.0001	evaluation reward: 1.55
episode: 927	score: 2.0	memory length: 172040 epsilon: 0.8573588200030966
steps: 198	lr: 0.0001	evaluation reward: 1.57
episode: 928	score: 1.0	memory length: 172210 epsilon: 0.8570222200031039
steps: 170	lr: 0.0001	evaluation reward: 1.58
episode: 929	score: 2.0	memory length: 172428 epsilon: 0.8565905800031133
steps: 218	lr: 0.0001	evaluation reward: 1.57
episode: 930	score: 1.0	memory length: 172597 epsilon: 0.8562559600031205
steps: 169	lr: 0.0001	evaluation reward: 1.57
episode: 931	score: 1.0	memory length: 172748 epsilon: 0.855956980003127
steps: 151	lr: 0.0001	evaluation reward: 1.56
episode: 932	score: 2.0	memory length: 172964 epsilon: 0.8555293000031363
steps: 216	lr: 0.0001	evaluation reward: 1.58
episode: 933	score: 2.0	memory length: 173162 epsilon: 0.8551372600031448
steps: 198	lr: 0.0001	evaluation reward: 1.59
episode: 934	score: 2.0	memory length: 173362 epsilon: 0.8547412600031534
steps: 200	lr: 0.0001	evaluation reward: 1.6
episode: 935	score: 2.0	memory length: 173560 epsilon: 0.8543492200031619
steps: 198	lr: 0.0001	evaluation reward: 1.59
episode: 936	score: 4.0	memory length: 173828 epsilon: 0.8538185800031735
steps: 268	lr: 0.0001	evaluation reward: 1.63
episode: 937	score: 1.0	memory length: 173999 epsilon: 0.8534800000031808
steps: 171	lr: 0.0001	evaluation reward: 1.64
episode: 938	score: 0.0	memory length: 174122 epsilon: 0.8532364600031861
steps: 123	lr: 0.0001	evaluation reward: 1.62
episode: 939	score: 0.0	memory length: 174244 epsilon: 0.8529949000031913
steps: 122	lr: 0.0001	evaluation reward: 1.59
episode: 940	score: 0.0	memory length: 174367 epsilon: 0.8527513600031966
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 941	score: 2.0	memory length: 174565 epsilon: 0.8523593200032051
steps: 198	lr: 0.0001	evaluation reward: 1.6
episode: 942	score: 3.0	memory length: 174809 epsilon: 0.8518762000032156
steps: 244	lr: 0.0001	evaluation reward: 1.54
episode: 943	score: 0.0	memory length: 174932 epsilon: 0.8516326600032209
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 944	score: 2.0	memory length: 175150 epsilon: 0.8512010200032303

steps: 218	lr: 0.0001	evaluation reward: 1.54	
episode: 945	score: 1.0	memory length: 175319	epsilon: 0.8508664000032375
steps: 169	lr: 0.0001	evaluation reward: 1.5	
episode: 946	score: 0.0	memory length: 175442	epsilon: 0.8506228600032428
steps: 123	lr: 0.0001	evaluation reward: 1.48	
episode: 947	score: 2.0	memory length: 175640	epsilon: 0.8502308200032513
steps: 198	lr: 0.0001	evaluation reward: 1.49	
episode: 948	score: 1.0	memory length: 175791	epsilon: 0.8499318400032578
steps: 151	lr: 0.0001	evaluation reward: 1.48	
episode: 949	score: 2.0	memory length: 175991	epsilon: 0.8495358400032664
steps: 200	lr: 0.0001	evaluation reward: 1.48	
episode: 950	score: 3.0	memory length: 176255	epsilon: 0.8490131200032778
steps: 264	lr: 0.0001	evaluation reward: 1.49	
episode: 951	score: 2.0	memory length: 176456	epsilon: 0.8486151400032864
steps: 201	lr: 0.0001	evaluation reward: 1.49	
episode: 952	score: 3.0	memory length: 176683	epsilon: 0.8481656800032962
steps: 227	lr: 0.0001	evaluation reward: 1.5	
episode: 953	score: 3.0	memory length: 176910	epsilon: 0.8477162200033059
steps: 227	lr: 0.0001	evaluation reward: 1.51	
episode: 954	score: 2.0	memory length: 177108	epsilon: 0.8473241800033144
steps: 198	lr: 0.0001	evaluation reward: 1.53	
episode: 955	score: 2.0	memory length: 177329	epsilon: 0.8468866000033239
steps: 221	lr: 0.0001	evaluation reward: 1.55	
episode: 956	score: 2.0	memory length: 177526	epsilon: 0.8464965400033324
steps: 197	lr: 0.0001	evaluation reward: 1.57	
episode: 957	score: 2.0	memory length: 177748	epsilon: 0.846056980003342
steps: 222	lr: 0.0001	evaluation reward: 1.57	
episode: 958	score: 3.0	memory length: 177974	epsilon: 0.8456095000033517
steps: 226	lr: 0.0001	evaluation reward: 1.58	
episode: 959	score: 5.0	memory length: 178301	epsilon: 0.8449620400033657
steps: 327	lr: 0.0001	evaluation reward: 1.62	
episode: 960	score: 5.0	memory length: 178610	epsilon: 0.844350220003379
steps: 309	lr: 0.0001	evaluation reward: 1.65	
episode: 961	score: 2.0	memory length: 178808	epsilon: 0.8439581800033875
steps: 198	lr: 0.0001	evaluation reward: 1.63	
episode: 962	score: 1.0	memory length: 178977	epsilon: 0.8436235600033948
steps: 169	lr: 0.0001	evaluation reward: 1.6	
episode: 963	score: 2.0	memory length: 179194	epsilon: 0.8431939000034041
steps: 217	lr: 0.0001	evaluation reward: 1.6	
episode: 964	score: 7.0	memory length: 179515	epsilon: 0.8425583200034179
steps: 321	lr: 0.0001	evaluation reward: 1.61	
episode: 965	score: 0.0	memory length: 179638	epsilon: 0.8423147800034232
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 966	score: 1.0	memory length: 179809	epsilon: 0.8419762000034305
steps: 171	lr: 0.0001	evaluation reward: 1.59	
episode: 967	score: 2.0	memory length: 180007	epsilon: 0.841584160003439
steps: 198	lr: 0.0001	evaluation reward: 1.59	
episode: 968	score: 0.0	memory length: 180130	epsilon: 0.8413406200034443

steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 969	score: 1.0	memory length: 180300	epsilon: 0.8410040200034516
steps: 170	lr: 0.0001	evaluation reward: 1.6	
episode: 970	score: 3.0	memory length: 180548	epsilon: 0.8405129800034623
steps: 248	lr: 0.0001	evaluation reward: 1.61	
episode: 971	score: 2.0	memory length: 180766	epsilon: 0.8400813400034717
steps: 218	lr: 0.0001	evaluation reward: 1.61	
episode: 972	score: 3.0	memory length: 180996	epsilon: 0.8396259400034816
steps: 230	lr: 0.0001	evaluation reward: 1.61	
episode: 973	score: 1.0	memory length: 181166	epsilon: 0.8392893400034889
steps: 170	lr: 0.0001	evaluation reward: 1.59	
episode: 974	score: 1.0	memory length: 181335	epsilon: 0.8389547200034961
steps: 169	lr: 0.0001	evaluation reward: 1.6	
episode: 975	score: 2.0	memory length: 181556	epsilon: 0.8385171400035056
steps: 221	lr: 0.0001	evaluation reward: 1.61	
episode: 976	score: 1.0	memory length: 181726	epsilon: 0.8381805400035129
steps: 170	lr: 0.0001	evaluation reward: 1.62	
episode: 977	score: 0.0	memory length: 181849	epsilon: 0.8379370000035182
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 978	score: 0.0	memory length: 181972	epsilon: 0.8376934600035235
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 979	score: 2.0	memory length: 182170	epsilon: 0.837301420003532
steps: 198	lr: 0.0001	evaluation reward: 1.59	
episode: 980	score: 0.0	memory length: 182293	epsilon: 0.8370578800035373
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 981	score: 12.0	memory length: 182688	epsilon: 0.8362757800035543
steps: 395	lr: 0.0001	evaluation reward: 1.68	
episode: 982	score: 0.0	memory length: 182811	epsilon: 0.8360322400035596
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 983	score: 2.0	memory length: 183010	epsilon: 0.8356382200035681
steps: 199	lr: 0.0001	evaluation reward: 1.69	
episode: 984	score: 0.0	memory length: 183133	epsilon: 0.8353946800035734
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 985	score: 2.0	memory length: 183333	epsilon: 0.834998680003582
steps: 200	lr: 0.0001	evaluation reward: 1.69	
episode: 986	score: 3.0	memory length: 183579	epsilon: 0.8345116000035926
steps: 246	lr: 0.0001	evaluation reward: 1.71	
episode: 987	score: 1.0	memory length: 183749	epsilon: 0.8341750000035999
steps: 170	lr: 0.0001	evaluation reward: 1.71	
episode: 988	score: 0.0	memory length: 183872	epsilon: 0.8339314600036052
steps: 123	lr: 0.0001	evaluation reward: 1.71	
episode: 989	score: 3.0	memory length: 184100	epsilon: 0.833480020003615
steps: 228	lr: 0.0001	evaluation reward: 1.73	
episode: 990	score: 2.0	memory length: 184316	epsilon: 0.8330523400036243
steps: 216	lr: 0.0001	evaluation reward: 1.75	
episode: 991	score: 2.0	memory length: 184534	epsilon: 0.8326207000036336
steps: 218	lr: 0.0001	evaluation reward: 1.75	
episode: 992	score: 2.0	memory length: 184731	epsilon: 0.8322306400036421

steps: 197	lr: 0.0001	evaluation reward: 1.77
episode: 993	score: 1.0	memory length: 184882 epsilon: 0.8319316600036486
steps: 151	lr: 0.0001	evaluation reward: 1.76
episode: 994	score: 1.0	memory length: 185052 epsilon: 0.8315950600036559
steps: 170	lr: 0.0001	evaluation reward: 1.74
episode: 995	score: 1.0	memory length: 185220 epsilon: 0.8312624200036631
steps: 168	lr: 0.0001	evaluation reward: 1.73
episode: 996	score: 1.0	memory length: 185389 epsilon: 0.8309278000036704
steps: 169	lr: 0.0001	evaluation reward: 1.73
episode: 997	score: 0.0	memory length: 185512 epsilon: 0.8306842600036757
steps: 123	lr: 0.0001	evaluation reward: 1.71
episode: 998	score: 2.0	memory length: 185710 epsilon: 0.8302922200036842
steps: 198	lr: 0.0001	evaluation reward: 1.71
episode: 999	score: 0.0	memory length: 185833 epsilon: 0.8300486800036895
steps: 123	lr: 0.0001	evaluation reward: 1.69
episode: 1000	score: 3.0	memory length: 186100 epsilon: 0.829520020003701
steps: 267	lr: 0.0001	evaluation reward: 1.7
episode: 1001	score: 1.0	memory length: 186270 epsilon: 0.8291834200037083
steps: 170	lr: 0.0001	evaluation reward: 1.69
episode: 1002	score: 0.0	memory length: 186393 epsilon: 0.8289398800037135
steps: 123	lr: 0.0001	evaluation reward: 1.69
episode: 1003	score: 3.0	memory length: 186618 epsilon: 0.8284943800037232
steps: 225	lr: 0.0001	evaluation reward: 1.7
episode: 1004	score: 1.0	memory length: 186787 epsilon: 0.8281597600037305
steps: 169	lr: 0.0001	evaluation reward: 1.69
episode: 1005	score: 3.0	memory length: 187015 epsilon: 0.8277083200037403
steps: 228	lr: 0.0001	evaluation reward: 1.71
episode: 1006	score: 0.0	memory length: 187138 epsilon: 0.8274647800037456
steps: 123	lr: 0.0001	evaluation reward: 1.69
episode: 1007	score: 1.0	memory length: 187288 epsilon: 0.827167780003752
steps: 150	lr: 0.0001	evaluation reward: 1.67
episode: 1008	score: 1.0	memory length: 187439 epsilon: 0.8268688000037585
steps: 151	lr: 0.0001	evaluation reward: 1.68
episode: 1009	score: 1.0	memory length: 187611 epsilon: 0.8265282400037659
steps: 172	lr: 0.0001	evaluation reward: 1.69
episode: 1010	score: 3.0	memory length: 187839 epsilon: 0.8260768000037757
steps: 228	lr: 0.0001	evaluation reward: 1.69
episode: 1011	score: 1.0	memory length: 188008 epsilon: 0.825742180003783
steps: 169	lr: 0.0001	evaluation reward: 1.66
episode: 1012	score: 0.0	memory length: 188131 epsilon: 0.8254986400037883
steps: 123	lr: 0.0001	evaluation reward: 1.64
episode: 1013	score: 3.0	memory length: 188396 epsilon: 0.8249739400037996
steps: 265	lr: 0.0001	evaluation reward: 1.64
episode: 1014	score: 2.0	memory length: 188614 epsilon: 0.824542300003809
steps: 218	lr: 0.0001	evaluation reward: 1.64
episode: 1015	score: 4.0	memory length: 188891 epsilon: 0.8239938400038209
steps: 277	lr: 0.0001	evaluation reward: 1.68
episode: 1016	score: 1.0	memory length: 189043 epsilon: 0.8236928800038275

steps: 152	lr: 0.0001	evaluation reward: 1.69	
episode: 1017	score: 1.0	memory length: 189213	epsilon: 0.8233562800038348
steps: 170	lr: 0.0001	evaluation reward: 1.69	
episode: 1018	score: 0.0	memory length: 189336	epsilon: 0.82311274000384
steps: 123	lr: 0.0001	evaluation reward: 1.68	
episode: 1019	score: 0.0	memory length: 189459	epsilon: 0.8228692000038453
steps: 123	lr: 0.0001	evaluation reward: 1.68	
episode: 1020	score: 3.0	memory length: 189705	epsilon: 0.8223821200038559
steps: 246	lr: 0.0001	evaluation reward: 1.7	
episode: 1021	score: 1.0	memory length: 189856	epsilon: 0.8220831400038624
steps: 151	lr: 0.0001	evaluation reward: 1.66	
episode: 1022	score: 3.0	memory length: 190084	epsilon: 0.8216317000038722
steps: 228	lr: 0.0001	evaluation reward: 1.69	
episode: 1023	score: 1.0	memory length: 190253	epsilon: 0.8212970800038795
steps: 169	lr: 0.0001	evaluation reward: 1.69	
episode: 1024	score: 1.0	memory length: 190422	epsilon: 0.8209624600038867
steps: 169	lr: 0.0001	evaluation reward: 1.69	
episode: 1025	score: 1.0	memory length: 190574	epsilon: 0.8206615000038933
steps: 152	lr: 0.0001	evaluation reward: 1.7	
episode: 1026	score: 2.0	memory length: 190772	epsilon: 0.8202694600039018
steps: 198	lr: 0.0001	evaluation reward: 1.72	
episode: 1027	score: 4.0	memory length: 191089	epsilon: 0.8196418000039154
steps: 317	lr: 0.0001	evaluation reward: 1.74	
episode: 1028	score: 0.0	memory length: 191212	epsilon: 0.8193982600039207
steps: 123	lr: 0.0001	evaluation reward: 1.73	
episode: 1029	score: 0.0	memory length: 191335	epsilon: 0.819154720003926
steps: 123	lr: 0.0001	evaluation reward: 1.71	
episode: 1030	score: 1.0	memory length: 191486	epsilon: 0.8188557400039325
steps: 151	lr: 0.0001	evaluation reward: 1.71	
episode: 1031	score: 2.0	memory length: 191668	epsilon: 0.8184953800039403
steps: 182	lr: 0.0001	evaluation reward: 1.72	
episode: 1032	score: 2.0	memory length: 191866	epsilon: 0.8181033400039488
steps: 198	lr: 0.0001	evaluation reward: 1.72	
episode: 1033	score: 4.0	memory length: 192131	epsilon: 0.8175786400039602
steps: 265	lr: 0.0001	evaluation reward: 1.74	
episode: 1034	score: 0.0	memory length: 192254	epsilon: 0.8173351000039655
steps: 123	lr: 0.0001	evaluation reward: 1.72	
episode: 1035	score: 2.0	memory length: 192454	epsilon: 0.8169391000039741
steps: 200	lr: 0.0001	evaluation reward: 1.72	
episode: 1036	score: 3.0	memory length: 192701	epsilon: 0.8164500400039847
steps: 247	lr: 0.0001	evaluation reward: 1.71	
episode: 1037	score: 6.0	memory length: 193067	epsilon: 0.8157253600040004
steps: 366	lr: 0.0001	evaluation reward: 1.76	
episode: 1038	score: 0.0	memory length: 193190	epsilon: 0.8154818200040057
steps: 123	lr: 0.0001	evaluation reward: 1.76	
episode: 1039	score: 2.0	memory length: 193388	epsilon: 0.8150897800040142
steps: 198	lr: 0.0001	evaluation reward: 1.78	
episode: 1040	score: 1.0	memory length: 193539	epsilon: 0.8147908000040207

steps: 151	lr: 0.0001	evaluation reward: 1.79	
episode: 1041	score: 0.0	memory length: 193661	epsilon: 0.814549240004026
steps: 122	lr: 0.0001	evaluation reward: 1.77	
episode: 1042	score: 1.0	memory length: 193830	epsilon: 0.8142146200040332
steps: 169	lr: 0.0001	evaluation reward: 1.75	
episode: 1043	score: 5.0	memory length: 194137	epsilon: 0.8136067600040464
steps: 307	lr: 0.0001	evaluation reward: 1.8	
episode: 1044	score: 0.0	memory length: 194260	epsilon: 0.8133632200040517
steps: 123	lr: 0.0001	evaluation reward: 1.78	
episode: 1045	score: 0.0	memory length: 194383	epsilon: 0.813119680004057
steps: 123	lr: 0.0001	evaluation reward: 1.77	
episode: 1046	score: 2.0	memory length: 194580	epsilon: 0.8127296200040655
steps: 197	lr: 0.0001	evaluation reward: 1.79	
episode: 1047	score: 2.0	memory length: 194797	epsilon: 0.8122999600040748
steps: 217	lr: 0.0001	evaluation reward: 1.79	
episode: 1048	score: 1.0	memory length: 194947	epsilon: 0.8120029600040812
steps: 150	lr: 0.0001	evaluation reward: 1.79	
episode: 1049	score: 2.0	memory length: 195145	epsilon: 0.8116109200040897
steps: 198	lr: 0.0001	evaluation reward: 1.79	
episode: 1050	score: 0.0	memory length: 195268	epsilon: 0.811367380004095
steps: 123	lr: 0.0001	evaluation reward: 1.76	
episode: 1051	score: 2.0	memory length: 195466	epsilon: 0.8109753400041035
steps: 198	lr: 0.0001	evaluation reward: 1.76	
episode: 1052	score: 1.0	memory length: 195617	epsilon: 0.81067636000411
steps: 151	lr: 0.0001	evaluation reward: 1.74	
episode: 1053	score: 2.0	memory length: 195815	epsilon: 0.8102843200041185
steps: 198	lr: 0.0001	evaluation reward: 1.73	
episode: 1054	score: 2.0	memory length: 196033	epsilon: 0.8098526800041279
steps: 218	lr: 0.0001	evaluation reward: 1.73	
episode: 1055	score: 3.0	memory length: 196280	epsilon: 0.8093636200041385
steps: 247	lr: 0.0001	evaluation reward: 1.74	
episode: 1056	score: 2.0	memory length: 196478	epsilon: 0.808971580004147
steps: 198	lr: 0.0001	evaluation reward: 1.74	
episode: 1057	score: 3.0	memory length: 196725	epsilon: 0.8084825200041577
steps: 247	lr: 0.0001	evaluation reward: 1.75	
episode: 1058	score: 3.0	memory length: 196969	epsilon: 0.8079994000041681
steps: 244	lr: 0.0001	evaluation reward: 1.75	
episode: 1059	score: 0.0	memory length: 197092	epsilon: 0.8077558600041734
steps: 123	lr: 0.0001	evaluation reward: 1.7	
episode: 1060	score: 2.0	memory length: 197290	epsilon: 0.8073638200041819
steps: 198	lr: 0.0001	evaluation reward: 1.67	
episode: 1061	score: 0.0	memory length: 197412	epsilon: 0.8071222600041872
steps: 122	lr: 0.0001	evaluation reward: 1.65	
episode: 1062	score: 0.0	memory length: 197534	epsilon: 0.8068807000041924
steps: 122	lr: 0.0001	evaluation reward: 1.64	
episode: 1063	score: 0.0	memory length: 197657	epsilon: 0.8066371600041977
steps: 123	lr: 0.0001	evaluation reward: 1.62	
episode: 1064	score: 1.0	memory length: 197826	epsilon: 0.806302540004205

steps: 169	lr: 0.0001	evaluation reward: 1.56	
episode: 1065	score: 3.0	memory length: 198054	epsilon: 0.8058511000042148
steps: 228	lr: 0.0001	evaluation reward: 1.59	
episode: 1066	score: 3.0	memory length: 198317	epsilon: 0.8053303600042261
steps: 263	lr: 0.0001	evaluation reward: 1.61	
episode: 1067	score: 3.0	memory length: 198564	epsilon: 0.8048413000042367
steps: 247	lr: 0.0001	evaluation reward: 1.62	
episode: 1068	score: 2.0	memory length: 198781	epsilon: 0.804411640004246
steps: 217	lr: 0.0001	evaluation reward: 1.64	
episode: 1069	score: 1.0	memory length: 198932	epsilon: 0.8041126600042525
steps: 151	lr: 0.0001	evaluation reward: 1.64	
episode: 1070	score: 1.0	memory length: 199101	epsilon: 0.8037780400042598
steps: 169	lr: 0.0001	evaluation reward: 1.62	
episode: 1071	score: 2.0	memory length: 199298	epsilon: 0.8033879800042683
steps: 197	lr: 0.0001	evaluation reward: 1.62	
episode: 1072	score: 1.0	memory length: 199467	epsilon: 0.8030533600042755
steps: 169	lr: 0.0001	evaluation reward: 1.6	
episode: 1073	score: 0.0	memory length: 199590	epsilon: 0.8028098200042808
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 1074	score: 2.0	memory length: 199790	epsilon: 0.8024138200042894
steps: 200	lr: 0.0001	evaluation reward: 1.6	
episode: 1075	score: 0.0	memory length: 199912	epsilon: 0.8021722600042946
steps: 122	lr: 0.0001	evaluation reward: 1.58	
episode: 1076	score: 0.0	memory length: 200035	epsilon: 0.8019287200042999
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 1077	score: 3.0	memory length: 200282	epsilon: 0.8014396600043106
steps: 247	lr: 0.0001	evaluation reward: 1.6	
episode: 1078	score: 4.0	memory length: 200595	epsilon: 0.800819920004324
steps: 313	lr: 0.0001	evaluation reward: 1.64	
episode: 1079	score: 4.0	memory length: 200868	epsilon: 0.8002793800043357
steps: 273	lr: 0.0001	evaluation reward: 1.66	
episode: 1080	score: 2.0	memory length: 201084	epsilon: 0.799851700004345
steps: 216	lr: 0.0001	evaluation reward: 1.68	
episode: 1081	score: 0.0	memory length: 201207	epsilon: 0.7996081600043503
steps: 123	lr: 0.0001	evaluation reward: 1.56	
episode: 1082	score: 1.0	memory length: 201375	epsilon: 0.7992755200043575
steps: 168	lr: 0.0001	evaluation reward: 1.57	
episode: 1083	score: 4.0	memory length: 201652	epsilon: 0.7987270600043694
steps: 277	lr: 0.0001	evaluation reward: 1.59	
episode: 1084	score: 2.0	memory length: 201869	epsilon: 0.7982974000043788
steps: 217	lr: 0.0001	evaluation reward: 1.61	
episode: 1085	score: 3.0	memory length: 202097	epsilon: 0.7978459600043886
steps: 228	lr: 0.0001	evaluation reward: 1.62	
episode: 1086	score: 3.0	memory length: 202323	epsilon: 0.7973984800043983
steps: 226	lr: 0.0001	evaluation reward: 1.62	
episode: 1087	score: 0.0	memory length: 202446	epsilon: 0.7971549400044036
steps: 123	lr: 0.0001	evaluation reward: 1.61	
episode: 1088	score: 3.0	memory length: 202714	epsilon: 0.7966243000044151

steps: 268	lr: 0.0001	evaluation reward: 1.64	
episode: 1089	score: 0.0	memory length: 202837	epsilon: 0.7963807600044204
steps: 123	lr: 0.0001	evaluation reward: 1.61	
episode: 1090	score: 4.0	memory length: 203132	epsilon: 0.795796660004433
steps: 295	lr: 0.0001	evaluation reward: 1.63	
episode: 1091	score: 0.0	memory length: 203255	epsilon: 0.7955531200044383
steps: 123	lr: 0.0001	evaluation reward: 1.61	
episode: 1092	score: 2.0	memory length: 203453	epsilon: 0.7951610800044469
steps: 198	lr: 0.0001	evaluation reward: 1.61	
episode: 1093	score: 2.0	memory length: 203651	epsilon: 0.7947690400044554
steps: 198	lr: 0.0001	evaluation reward: 1.62	
episode: 1094	score: 2.0	memory length: 203849	epsilon: 0.7943770000044639
steps: 198	lr: 0.0001	evaluation reward: 1.63	
episode: 1095	score: 2.0	memory length: 204067	epsilon: 0.7939453600044732
steps: 218	lr: 0.0001	evaluation reward: 1.64	
episode: 1096	score: 7.0	memory length: 204368	epsilon: 0.7933493800044862
steps: 301	lr: 0.0001	evaluation reward: 1.7	
episode: 1097	score: 2.0	memory length: 204584	epsilon: 0.7929217000044955
steps: 216	lr: 0.0001	evaluation reward: 1.72	
episode: 1098	score: 1.0	memory length: 204735	epsilon: 0.792622720004502
steps: 151	lr: 0.0001	evaluation reward: 1.71	
episode: 1099	score: 1.0	memory length: 204886	epsilon: 0.7923237400045084
steps: 151	lr: 0.0001	evaluation reward: 1.72	
episode: 1100	score: 2.0	memory length: 205084	epsilon: 0.791931700004517
steps: 198	lr: 0.0001	evaluation reward: 1.71	
episode: 1101	score: 4.0	memory length: 205372	epsilon: 0.7913614600045293
steps: 288	lr: 0.0001	evaluation reward: 1.74	
episode: 1102	score: 2.0	memory length: 205569	epsilon: 0.7909714000045378
steps: 197	lr: 0.0001	evaluation reward: 1.76	
episode: 1103	score: 9.0	memory length: 205933	epsilon: 0.7902506800045535
steps: 364	lr: 0.0001	evaluation reward: 1.82	
episode: 1104	score: 1.0	memory length: 206084	epsilon: 0.7899517000045599
steps: 151	lr: 0.0001	evaluation reward: 1.82	
episode: 1105	score: 3.0	memory length: 206310	epsilon: 0.7895042200045697
steps: 226	lr: 0.0001	evaluation reward: 1.82	
episode: 1106	score: 3.0	memory length: 206573	epsilon: 0.788983480004581
steps: 263	lr: 0.0001	evaluation reward: 1.85	
episode: 1107	score: 2.0	memory length: 206772	epsilon: 0.7885894600045895
steps: 199	lr: 0.0001	evaluation reward: 1.86	
episode: 1108	score: 4.0	memory length: 207048	epsilon: 0.7880429800046014
steps: 276	lr: 0.0001	evaluation reward: 1.89	
episode: 1109	score: 1.0	memory length: 207219	epsilon: 0.7877044000046087
steps: 171	lr: 0.0001	evaluation reward: 1.89	
episode: 1110	score: 3.0	memory length: 207445	epsilon: 0.7872569200046184
steps: 226	lr: 0.0001	evaluation reward: 1.89	
episode: 1111	score: 2.0	memory length: 207642	epsilon: 0.7868668600046269
steps: 197	lr: 0.0001	evaluation reward: 1.9	
episode: 1112	score: 2.0	memory length: 207857	epsilon: 0.7864411600046362

steps: 215	lr: 0.0001	evaluation reward: 1.92	
episode: 1113	score: 0.0	memory length: 207980	epsilon: 0.7861976200046414
steps: 123	lr: 0.0001	evaluation reward: 1.89	
episode: 1114	score: 2.0	memory length: 208178	epsilon: 0.78580558000465
steps: 198	lr: 0.0001	evaluation reward: 1.89	
episode: 1115	score: 1.0	memory length: 208329	epsilon: 0.7855066000046564
steps: 151	lr: 0.0001	evaluation reward: 1.86	
episode: 1116	score: 0.0	memory length: 208452	epsilon: 0.7852630600046617
steps: 123	lr: 0.0001	evaluation reward: 1.85	
episode: 1117	score: 1.0	memory length: 208624	epsilon: 0.7849225000046691
steps: 172	lr: 0.0001	evaluation reward: 1.85	
episode: 1118	score: 0.0	memory length: 208747	epsilon: 0.7846789600046744
steps: 123	lr: 0.0001	evaluation reward: 1.85	
episode: 1119	score: 1.0	memory length: 208898	epsilon: 0.7843799800046809
steps: 151	lr: 0.0001	evaluation reward: 1.86	
episode: 1120	score: 3.0	memory length: 209142	epsilon: 0.7838968600046914
steps: 244	lr: 0.0001	evaluation reward: 1.86	
episode: 1121	score: 1.0	memory length: 209312	epsilon: 0.7835602600046987
steps: 170	lr: 0.0001	evaluation reward: 1.86	
episode: 1122	score: 2.0	memory length: 209510	epsilon: 0.7831682200047072
steps: 198	lr: 0.0001	evaluation reward: 1.85	
episode: 1123	score: 0.0	memory length: 209633	epsilon: 0.7829246800047125
steps: 123	lr: 0.0001	evaluation reward: 1.84	
episode: 1124	score: 3.0	memory length: 209879	epsilon: 0.7824376000047231
steps: 246	lr: 0.0001	evaluation reward: 1.86	
episode: 1125	score: 2.0	memory length: 210097	epsilon: 0.7820059600047324
steps: 218	lr: 0.0001	evaluation reward: 1.87	
episode: 1126	score: 3.0	memory length: 210323	epsilon: 0.7815584800047422
steps: 226	lr: 0.0001	evaluation reward: 1.88	
episode: 1127	score: 2.0	memory length: 210540	epsilon: 0.7811288200047515
steps: 217	lr: 0.0001	evaluation reward: 1.86	
episode: 1128	score: 1.0	memory length: 210711	epsilon: 0.7807902400047588
steps: 171	lr: 0.0001	evaluation reward: 1.87	
episode: 1129	score: 4.0	memory length: 211003	epsilon: 0.7802120800047714
steps: 292	lr: 0.0001	evaluation reward: 1.91	
episode: 1130	score: 2.0	memory length: 211201	epsilon: 0.7798200400047799
steps: 198	lr: 0.0001	evaluation reward: 1.92	
episode: 1131	score: 2.0	memory length: 211401	epsilon: 0.7794240400047885
steps: 200	lr: 0.0001	evaluation reward: 1.92	
episode: 1132	score: 0.0	memory length: 211524	epsilon: 0.7791805000047938
steps: 123	lr: 0.0001	evaluation reward: 1.9	
episode: 1133	score: 0.0	memory length: 211646	epsilon: 0.778938940004799
steps: 122	lr: 0.0001	evaluation reward: 1.86	
episode: 1134	score: 5.0	memory length: 211954	epsilon: 0.7783291000048123
steps: 308	lr: 0.0001	evaluation reward: 1.91	
episode: 1135	score: 3.0	memory length: 212200	epsilon: 0.7778420200048228
steps: 246	lr: 0.0001	evaluation reward: 1.92	
episode: 1136	score: 2.0	memory length: 212398	epsilon: 0.7774499800048313

steps: 198	lr: 0.0001	evaluation reward: 1.91	
episode: 1137	score: 1.0	memory length: 212549	epsilon: 0.7771510000048378
steps: 151	lr: 0.0001	evaluation reward: 1.86	
episode: 1138	score: 1.0	memory length: 212717	epsilon: 0.776818360004845
steps: 168	lr: 0.0001	evaluation reward: 1.87	
episode: 1139	score: 3.0	memory length: 212943	epsilon: 0.7763708800048548
steps: 226	lr: 0.0001	evaluation reward: 1.88	
episode: 1140	score: 2.0	memory length: 213161	epsilon: 0.7759392400048641
steps: 218	lr: 0.0001	evaluation reward: 1.89	
episode: 1141	score: 2.0	memory length: 213379	epsilon: 0.7755076000048735
steps: 218	lr: 0.0001	evaluation reward: 1.91	
episode: 1142	score: 1.0	memory length: 213549	epsilon: 0.7751710000048808
steps: 170	lr: 0.0001	evaluation reward: 1.91	
episode: 1143	score: 2.0	memory length: 213746	epsilon: 0.7747809400048893
steps: 197	lr: 0.0001	evaluation reward: 1.88	
episode: 1144	score: 2.0	memory length: 213964	epsilon: 0.7743493000048987
steps: 218	lr: 0.0001	evaluation reward: 1.9	
episode: 1145	score: 3.0	memory length: 214208	epsilon: 0.7738661800049091
steps: 244	lr: 0.0001	evaluation reward: 1.93	
episode: 1146	score: 0.0	memory length: 214330	epsilon: 0.7736246200049144
steps: 122	lr: 0.0001	evaluation reward: 1.91	
episode: 1147	score: 1.0	memory length: 214499	epsilon: 0.7732900000049217
steps: 169	lr: 0.0001	evaluation reward: 1.9	
episode: 1148	score: 2.0	memory length: 214697	epsilon: 0.7728979600049302
steps: 198	lr: 0.0001	evaluation reward: 1.91	
episode: 1149	score: 3.0	memory length: 214945	epsilon: 0.7724069200049408
steps: 248	lr: 0.0001	evaluation reward: 1.92	
episode: 1150	score: 0.0	memory length: 215067	epsilon: 0.7721653600049461
steps: 122	lr: 0.0001	evaluation reward: 1.92	
episode: 1151	score: 0.0	memory length: 215189	epsilon: 0.7719238000049513
steps: 122	lr: 0.0001	evaluation reward: 1.9	
episode: 1152	score: 1.0	memory length: 215339	epsilon: 0.7716268000049578
steps: 150	lr: 0.0001	evaluation reward: 1.9	
episode: 1153	score: 2.0	memory length: 215542	epsilon: 0.7712248600049665
steps: 203	lr: 0.0001	evaluation reward: 1.9	
episode: 1154	score: 2.0	memory length: 215740	epsilon: 0.770832820004975
steps: 198	lr: 0.0001	evaluation reward: 1.9	
episode: 1155	score: 2.0	memory length: 215938	epsilon: 0.7704407800049835
steps: 198	lr: 0.0001	evaluation reward: 1.89	
episode: 1156	score: 4.0	memory length: 216254	epsilon: 0.7698151000049971
steps: 316	lr: 0.0001	evaluation reward: 1.91	
episode: 1157	score: 1.0	memory length: 216405	epsilon: 0.7695161200050036
steps: 151	lr: 0.0001	evaluation reward: 1.89	
episode: 1158	score: 1.0	memory length: 216574	epsilon: 0.7691815000050108
steps: 169	lr: 0.0001	evaluation reward: 1.87	
episode: 1159	score: 1.0	memory length: 216725	epsilon: 0.7688825200050173
steps: 151	lr: 0.0001	evaluation reward: 1.88	
episode: 1160	score: 3.0	memory length: 216951	epsilon: 0.768435040005027

steps: 226	lr: 0.0001	evaluation reward: 1.89	
episode: 1161	score: 1.0	memory length: 217120	epsilon: 0.7681004200050343
steps: 169	lr: 0.0001	evaluation reward: 1.9	
episode: 1162	score: 3.0	memory length: 217367	epsilon: 0.7676113600050449
steps: 247	lr: 0.0001	evaluation reward: 1.93	
episode: 1163	score: 1.0	memory length: 217535	epsilon: 0.7672787200050522
steps: 168	lr: 0.0001	evaluation reward: 1.94	
episode: 1164	score: 4.0	memory length: 217846	epsilon: 0.7666629400050655
steps: 311	lr: 0.0001	evaluation reward: 1.97	
episode: 1165	score: 1.0	memory length: 217997	epsilon: 0.766363960005072
steps: 151	lr: 0.0001	evaluation reward: 1.95	
episode: 1166	score: 0.0	memory length: 218120	epsilon: 0.7661204200050773
steps: 123	lr: 0.0001	evaluation reward: 1.92	
episode: 1167	score: 1.0	memory length: 218288	epsilon: 0.7657877800050845
steps: 168	lr: 0.0001	evaluation reward: 1.9	
episode: 1168	score: 4.0	memory length: 218564	epsilon: 0.7652413000050964
steps: 276	lr: 0.0001	evaluation reward: 1.92	
episode: 1169	score: 1.0	memory length: 218716	epsilon: 0.7649403400051029
steps: 152	lr: 0.0001	evaluation reward: 1.92	
episode: 1170	score: 4.0	memory length: 219012	epsilon: 0.7643542600051156
steps: 296	lr: 0.0001	evaluation reward: 1.95	
episode: 1171	score: 4.0	memory length: 219311	epsilon: 0.7637622400051285
steps: 299	lr: 0.0001	evaluation reward: 1.97	
episode: 1172	score: 5.0	memory length: 219635	epsilon: 0.7631207200051424
steps: 324	lr: 0.0001	evaluation reward: 2.01	
episode: 1173	score: 3.0	memory length: 219860	epsilon: 0.7626752200051521
steps: 225	lr: 0.0001	evaluation reward: 2.04	
episode: 1174	score: 0.0	memory length: 219982	epsilon: 0.7624336600051573
steps: 122	lr: 0.0001	evaluation reward: 2.02	
episode: 1175	score: 0.0	memory length: 220105	epsilon: 0.7621901200051626
steps: 123	lr: 0.0001	evaluation reward: 2.02	
episode: 1176	score: 1.0	memory length: 220273	epsilon: 0.7618574800051698
steps: 168	lr: 0.0001	evaluation reward: 2.03	
episode: 1177	score: 2.0	memory length: 220471	epsilon: 0.7614654400051784
steps: 198	lr: 0.0001	evaluation reward: 2.02	
episode: 1178	score: 0.0	memory length: 220593	epsilon: 0.7612238800051836
steps: 122	lr: 0.0001	evaluation reward: 1.98	
episode: 1179	score: 1.0	memory length: 220761	epsilon: 0.7608912400051908
steps: 168	lr: 0.0001	evaluation reward: 1.95	
episode: 1180	score: 7.0	memory length: 221081	epsilon: 0.7602576400052046
steps: 320	lr: 0.0001	evaluation reward: 2.0	
episode: 1181	score: 1.0	memory length: 221250	epsilon: 0.7599230200052118
steps: 169	lr: 0.0001	evaluation reward: 2.01	
episode: 1182	score: 5.0	memory length: 221597	epsilon: 0.7592359600052268
steps: 347	lr: 0.0001	evaluation reward: 2.05	
episode: 1183	score: 2.0	memory length: 221814	epsilon: 0.7588063000052361
steps: 217	lr: 0.0001	evaluation reward: 2.03	
episode: 1184	score: 6.0	memory length: 222156	epsilon: 0.7581291400052508

steps: 342	lr: 0.0001	evaluation reward: 2.07	
episode: 1185	score: 0.0	memory length: 222278	epsilon: 0.757887580005256
steps: 122	lr: 0.0001	evaluation reward: 2.04	
episode: 1186	score: 1.0	memory length: 222449	epsilon: 0.7575490000052634
steps: 171	lr: 0.0001	evaluation reward: 2.02	
episode: 1187	score: 0.0	memory length: 222572	epsilon: 0.7573054600052687
steps: 123	lr: 0.0001	evaluation reward: 2.02	
episode: 1188	score: 1.0	memory length: 222723	epsilon: 0.7570064800052752
steps: 151	lr: 0.0001	evaluation reward: 2.0	
episode: 1189	score: 7.0	memory length: 223151	epsilon: 0.7561590400052935
steps: 428	lr: 0.0001	evaluation reward: 2.07	
episode: 1190	score: 1.0	memory length: 223319	epsilon: 0.7558264000053008
steps: 168	lr: 0.0001	evaluation reward: 2.04	
episode: 1191	score: 2.0	memory length: 223536	epsilon: 0.7553967400053101
steps: 217	lr: 0.0001	evaluation reward: 2.06	
episode: 1192	score: 4.0	memory length: 223830	epsilon: 0.7548146200053227
steps: 294	lr: 0.0001	evaluation reward: 2.08	
episode: 1193	score: 1.0	memory length: 223999	epsilon: 0.75448000000533
steps: 169	lr: 0.0001	evaluation reward: 2.07	
episode: 1194	score: 3.0	memory length: 224245	epsilon: 0.7539929200053406
steps: 246	lr: 0.0001	evaluation reward: 2.08	
episode: 1195	score: 4.0	memory length: 224502	epsilon: 0.7534840600053516
steps: 257	lr: 0.0001	evaluation reward: 2.1	
episode: 1196	score: 1.0	memory length: 224673	epsilon: 0.753145480005359
steps: 171	lr: 0.0001	evaluation reward: 2.04	
episode: 1197	score: 1.0	memory length: 224824	epsilon: 0.7528465000053655
steps: 151	lr: 0.0001	evaluation reward: 2.03	
episode: 1198	score: 0.0	memory length: 224947	epsilon: 0.7526029600053707
steps: 123	lr: 0.0001	evaluation reward: 2.02	
episode: 1199	score: 2.0	memory length: 225127	epsilon: 0.7522465600053785
steps: 180	lr: 0.0001	evaluation reward: 2.03	
episode: 1200	score: 4.0	memory length: 225425	epsilon: 0.7516565200053913
steps: 298	lr: 0.0001	evaluation reward: 2.05	
episode: 1201	score: 2.0	memory length: 225643	epsilon: 0.7512248800054007
steps: 218	lr: 0.0001	evaluation reward: 2.03	
episode: 1202	score: 2.0	memory length: 225865	epsilon: 0.7507853200054102
steps: 222	lr: 0.0001	evaluation reward: 2.03	
episode: 1203	score: 2.0	memory length: 226063	epsilon: 0.7503932800054187
steps: 198	lr: 0.0001	evaluation reward: 1.96	
episode: 1204	score: 0.0	memory length: 226186	epsilon: 0.750149740005424
steps: 123	lr: 0.0001	evaluation reward: 1.95	
episode: 1205	score: 0.0	memory length: 226309	epsilon: 0.7499062000054293
steps: 123	lr: 0.0001	evaluation reward: 1.92	
episode: 1206	score: 0.0	memory length: 226432	epsilon: 0.7496626600054346
steps: 123	lr: 0.0001	evaluation reward: 1.89	
episode: 1207	score: 3.0	memory length: 226643	epsilon: 0.7492448800054436
steps: 211	lr: 0.0001	evaluation reward: 1.9	
episode: 1208	score: 0.0	memory length: 226766	epsilon: 0.7490013400054489

steps: 123	lr: 0.0001	evaluation reward: 1.86	
episode: 1209	score: 2.0	memory length: 226984	epsilon: 0.7485697000054583
steps: 218	lr: 0.0001	evaluation reward: 1.87	
episode: 1210	score: 3.0	memory length: 227195	epsilon: 0.7481519200054674
steps: 211	lr: 0.0001	evaluation reward: 1.87	
episode: 1211	score: 5.0	memory length: 227540	epsilon: 0.7474688200054822
steps: 345	lr: 0.0001	evaluation reward: 1.9	
episode: 1212	score: 1.0	memory length: 227710	epsilon: 0.7471322200054895
steps: 170	lr: 0.0001	evaluation reward: 1.89	
episode: 1213	score: 1.0	memory length: 227860	epsilon: 0.746835220005496
steps: 150	lr: 0.0001	evaluation reward: 1.9	
episode: 1214	score: 2.0	memory length: 228040	epsilon: 0.7464788200055037
steps: 180	lr: 0.0001	evaluation reward: 1.9	
episode: 1215	score: 2.0	memory length: 228238	epsilon: 0.7460867800055122
steps: 198	lr: 0.0001	evaluation reward: 1.91	
episode: 1216	score: 4.0	memory length: 228533	epsilon: 0.7455026800055249
steps: 295	lr: 0.0001	evaluation reward: 1.95	
episode: 1217	score: 2.0	memory length: 228731	epsilon: 0.7451106400055334
steps: 198	lr: 0.0001	evaluation reward: 1.96	
episode: 1218	score: 1.0	memory length: 228900	epsilon: 0.7447760200055407
steps: 169	lr: 0.0001	evaluation reward: 1.97	
episode: 1219	score: 0.0	memory length: 229022	epsilon: 0.7445344600055459
steps: 122	lr: 0.0001	evaluation reward: 1.96	
episode: 1220	score: 3.0	memory length: 229268	epsilon: 0.7440473800055565
steps: 246	lr: 0.0001	evaluation reward: 1.96	
episode: 1221	score: 2.0	memory length: 229489	epsilon: 0.743609800005566
steps: 221	lr: 0.0001	evaluation reward: 1.97	
episode: 1222	score: 5.0	memory length: 229815	epsilon: 0.74296432000558
steps: 326	lr: 0.0001	evaluation reward: 2.0	
episode: 1223	score: 2.0	memory length: 230016	epsilon: 0.7425663400055886
steps: 201	lr: 0.0001	evaluation reward: 2.02	
episode: 1224	score: 3.0	memory length: 230262	epsilon: 0.7420792600055992
steps: 246	lr: 0.0001	evaluation reward: 2.02	
episode: 1225	score: 4.0	memory length: 230576	epsilon: 0.7414575400056127
steps: 314	lr: 0.0001	evaluation reward: 2.04	
episode: 1226	score: 2.0	memory length: 230776	epsilon: 0.7410615400056213
steps: 200	lr: 0.0001	evaluation reward: 2.03	
episode: 1227	score: 1.0	memory length: 230927	epsilon: 0.7407625600056278
steps: 151	lr: 0.0001	evaluation reward: 2.02	
episode: 1228	score: 1.0	memory length: 231078	epsilon: 0.7404635800056343
steps: 151	lr: 0.0001	evaluation reward: 2.02	
episode: 1229	score: 1.0	memory length: 231249	epsilon: 0.7401250000056416
steps: 171	lr: 0.0001	evaluation reward: 1.99	
episode: 1230	score: 2.0	memory length: 231467	epsilon: 0.739693360005651
steps: 218	lr: 0.0001	evaluation reward: 1.99	
episode: 1231	score: 2.0	memory length: 231686	epsilon: 0.7392597400056604
steps: 219	lr: 0.0001	evaluation reward: 1.99	
episode: 1232	score: 4.0	memory length: 231984	epsilon: 0.7386697000056732

steps: 298	lr: 0.0001	evaluation reward: 2.03	
episode: 1233	score: 0.0	memory length: 232107	epsilon: 0.7384261600056785
steps: 123	lr: 0.0001	evaluation reward: 2.03	
episode: 1234	score: 1.0	memory length: 232278	epsilon: 0.7380875800056859
steps: 171	lr: 0.0001	evaluation reward: 1.99	
episode: 1235	score: 2.0	memory length: 232476	epsilon: 0.7376955400056944
steps: 198	lr: 0.0001	evaluation reward: 1.98	
episode: 1236	score: 0.0	memory length: 232599	epsilon: 0.7374520000056997
steps: 123	lr: 0.0001	evaluation reward: 1.96	
episode: 1237	score: 1.0	memory length: 232750	epsilon: 0.7371530200057061
steps: 151	lr: 0.0001	evaluation reward: 1.96	
episode: 1238	score: 2.0	memory length: 232948	epsilon: 0.7367609800057147
steps: 198	lr: 0.0001	evaluation reward: 1.97	
episode: 1239	score: 0.0	memory length: 233071	epsilon: 0.73651744000572
steps: 123	lr: 0.0001	evaluation reward: 1.94	
episode: 1240	score: 0.0	memory length: 233194	epsilon: 0.7362739000057252
steps: 123	lr: 0.0001	evaluation reward: 1.92	
episode: 1241	score: 2.0	memory length: 233392	epsilon: 0.7358818600057337
steps: 198	lr: 0.0001	evaluation reward: 1.92	
episode: 1242	score: 1.0	memory length: 233543	epsilon: 0.7355828800057402
steps: 151	lr: 0.0001	evaluation reward: 1.92	
episode: 1243	score: 3.0	memory length: 233793	epsilon: 0.735087880005751
steps: 250	lr: 0.0001	evaluation reward: 1.93	
episode: 1244	score: 7.0	memory length: 234093	epsilon: 0.7344938800057639
steps: 300	lr: 0.0001	evaluation reward: 1.98	
episode: 1245	score: 4.0	memory length: 234384	epsilon: 0.7339177000057764
steps: 291	lr: 0.0001	evaluation reward: 1.99	
episode: 1246	score: 5.0	memory length: 234698	epsilon: 0.7332959800057899
steps: 314	lr: 0.0001	evaluation reward: 2.04	
episode: 1247	score: 0.0	memory length: 234820	epsilon: 0.7330544200057951
steps: 122	lr: 0.0001	evaluation reward: 2.03	
episode: 1248	score: 1.0	memory length: 234989	epsilon: 0.7327198000058024
steps: 169	lr: 0.0001	evaluation reward: 2.02	
episode: 1249	score: 2.0	memory length: 235207	epsilon: 0.7322881600058118
steps: 218	lr: 0.0001	evaluation reward: 2.01	
episode: 1250	score: 3.0	memory length: 235455	epsilon: 0.7317971200058224
steps: 248	lr: 0.0001	evaluation reward: 2.04	
episode: 1251	score: 3.0	memory length: 235668	epsilon: 0.7313753800058316
steps: 213	lr: 0.0001	evaluation reward: 2.07	
episode: 1252	score: 4.0	memory length: 235946	epsilon: 0.7308249400058435
steps: 278	lr: 0.0001	evaluation reward: 2.1	
episode: 1253	score: 2.0	memory length: 236163	epsilon: 0.7303952800058529
steps: 217	lr: 0.0001	evaluation reward: 2.1	
episode: 1254	score: 3.0	memory length: 236373	epsilon: 0.7299794800058619
steps: 210	lr: 0.0001	evaluation reward: 2.11	
episode: 1255	score: 2.0	memory length: 236591	epsilon: 0.7295478400058713
steps: 218	lr: 0.0001	evaluation reward: 2.11	
episode: 1256	score: 5.0	memory length: 236872	epsilon: 0.7289914600058833

steps: 281	lr: 0.0001	evaluation reward: 2.12	
episode: 1257	score: 3.0	memory length: 237099	epsilon: 0.7285420000058931
steps: 227	lr: 0.0001	evaluation reward: 2.14	
episode: 1258	score: 1.0	memory length: 237249	epsilon: 0.7282450000058995
steps: 150	lr: 0.0001	evaluation reward: 2.14	
episode: 1259	score: 5.0	memory length: 237591	epsilon: 0.7275678400059142
steps: 342	lr: 0.0001	evaluation reward: 2.18	
episode: 1260	score: 2.0	memory length: 237790	epsilon: 0.7271738200059228
steps: 199	lr: 0.0001	evaluation reward: 2.17	
episode: 1261	score: 2.0	memory length: 237988	epsilon: 0.7267817800059313
steps: 198	lr: 0.0001	evaluation reward: 2.18	
episode: 1262	score: 1.0	memory length: 238160	epsilon: 0.7264412200059387
steps: 172	lr: 0.0001	evaluation reward: 2.16	
episode: 1263	score: 3.0	memory length: 238386	epsilon: 0.7259937400059484
steps: 226	lr: 0.0001	evaluation reward: 2.18	
episode: 1264	score: 2.0	memory length: 238604	epsilon: 0.7255621000059578
steps: 218	lr: 0.0001	evaluation reward: 2.16	
episode: 1265	score: 3.0	memory length: 238817	epsilon: 0.7251403600059669
steps: 213	lr: 0.0001	evaluation reward: 2.18	
episode: 1266	score: 3.0	memory length: 239029	epsilon: 0.724720600005976
steps: 212	lr: 0.0001	evaluation reward: 2.21	
episode: 1267	score: 1.0	memory length: 239180	epsilon: 0.7244216200059825
steps: 151	lr: 0.0001	evaluation reward: 2.21	
episode: 1268	score: 0.0	memory length: 239302	epsilon: 0.7241800600059878
steps: 122	lr: 0.0001	evaluation reward: 2.17	
episode: 1269	score: 0.0	memory length: 239425	epsilon: 0.7239365200059931
steps: 123	lr: 0.0001	evaluation reward: 2.16	
episode: 1270	score: 2.0	memory length: 239623	epsilon: 0.7235444800060016
steps: 198	lr: 0.0001	evaluation reward: 2.14	
episode: 1271	score: 0.0	memory length: 239746	epsilon: 0.7233009400060069
steps: 123	lr: 0.0001	evaluation reward: 2.1	
episode: 1272	score: 3.0	memory length: 239972	epsilon: 0.7228534600060166
steps: 226	lr: 0.0001	evaluation reward: 2.08	
episode: 1273	score: 2.0	memory length: 240169	epsilon: 0.722463400006025
steps: 197	lr: 0.0001	evaluation reward: 2.07	
episode: 1274	score: 5.0	memory length: 240462	epsilon: 0.7218832600060376
steps: 293	lr: 0.0001	evaluation reward: 2.12	
episode: 1275	score: 4.0	memory length: 240779	epsilon: 0.7212556000060513
steps: 317	lr: 0.0001	evaluation reward: 2.16	
episode: 1276	score: 3.0	memory length: 241007	epsilon: 0.7208041600060611
steps: 228	lr: 0.0001	evaluation reward: 2.18	
episode: 1277	score: 3.0	memory length: 241233	epsilon: 0.7203566800060708
steps: 226	lr: 0.0001	evaluation reward: 2.19	
episode: 1278	score: 2.0	memory length: 241415	epsilon: 0.7199963200060786
steps: 182	lr: 0.0001	evaluation reward: 2.21	
episode: 1279	score: 0.0	memory length: 241538	epsilon: 0.7197527800060839
steps: 123	lr: 0.0001	evaluation reward: 2.2	
episode: 1280	score: 3.0	memory length: 241768	epsilon: 0.7192973800060938

steps: 230	lr: 0.0001	evaluation reward: 2.16	
episode: 1281	score: 3.0	memory length: 241996	epsilon: 0.7188459400061036
steps: 228	lr: 0.0001	evaluation reward: 2.18	
episode: 1282	score: 1.0	memory length: 242147	epsilon: 0.7185469600061101
steps: 151	lr: 0.0001	evaluation reward: 2.14	
episode: 1283	score: 3.0	memory length: 242373	epsilon: 0.7180994800061198
steps: 226	lr: 0.0001	evaluation reward: 2.15	
episode: 1284	score: 3.0	memory length: 242584	epsilon: 0.7176817000061289
steps: 211	lr: 0.0001	evaluation reward: 2.12	
episode: 1285	score: 4.0	memory length: 242885	epsilon: 0.7170857200061418
steps: 301	lr: 0.0001	evaluation reward: 2.16	
episode: 1286	score: 1.0	memory length: 243054	epsilon: 0.7167511000061491
steps: 169	lr: 0.0001	evaluation reward: 2.16	
episode: 1287	score: 1.0	memory length: 243205	epsilon: 0.7164521200061555
steps: 151	lr: 0.0001	evaluation reward: 2.17	
episode: 1288	score: 2.0	memory length: 243403	epsilon: 0.7160600800061641
steps: 198	lr: 0.0001	evaluation reward: 2.18	
episode: 1289	score: 4.0	memory length: 243699	epsilon: 0.7154740000061768
steps: 296	lr: 0.0001	evaluation reward: 2.15	
episode: 1290	score: 0.0	memory length: 243822	epsilon: 0.7152304600061821
steps: 123	lr: 0.0001	evaluation reward: 2.14	
episode: 1291	score: 3.0	memory length: 244052	epsilon: 0.714775060006192
steps: 230	lr: 0.0001	evaluation reward: 2.15	
episode: 1292	score: 2.0	memory length: 244250	epsilon: 0.7143830200062005
steps: 198	lr: 0.0001	evaluation reward: 2.13	
episode: 1293	score: 3.0	memory length: 244475	epsilon: 0.7139375200062101
steps: 225	lr: 0.0001	evaluation reward: 2.15	
episode: 1294	score: 3.0	memory length: 244745	epsilon: 0.7134029200062217
steps: 270	lr: 0.0001	evaluation reward: 2.15	
episode: 1295	score: 3.0	memory length: 244989	epsilon: 0.7129198000062322
steps: 244	lr: 0.0001	evaluation reward: 2.14	
episode: 1296	score: 1.0	memory length: 245140	epsilon: 0.7126208200062387
steps: 151	lr: 0.0001	evaluation reward: 2.14	
episode: 1297	score: 3.0	memory length: 245366	epsilon: 0.7121733400062484
steps: 226	lr: 0.0001	evaluation reward: 2.16	
episode: 1298	score: 2.0	memory length: 245585	epsilon: 0.7117397200062578
steps: 219	lr: 0.0001	evaluation reward: 2.18	
episode: 1299	score: 1.0	memory length: 245756	epsilon: 0.7114011400062652
steps: 171	lr: 0.0001	evaluation reward: 2.17	
episode: 1300	score: 4.0	memory length: 246021	epsilon: 0.7108764400062766
steps: 265	lr: 0.0001	evaluation reward: 2.17	
episode: 1301	score: 2.0	memory length: 246237	epsilon: 0.7104487600062859
steps: 216	lr: 0.0001	evaluation reward: 2.17	
episode: 1302	score: 1.0	memory length: 246409	epsilon: 0.7101082000062933
steps: 172	lr: 0.0001	evaluation reward: 2.16	
episode: 1303	score: 4.0	memory length: 246688	epsilon: 0.7095557800063053
steps: 279	lr: 0.0001	evaluation reward: 2.18	
episode: 1304	score: 3.0	memory length: 246916	epsilon: 0.7091043400063151

steps: 228	lr: 0.0001	evaluation reward: 2.21	
episode: 1305	score: 1.0	memory length: 247086	epsilon: 0.7087677400063224
steps: 170	lr: 0.0001	evaluation reward: 2.22	
episode: 1306	score: 0.0	memory length: 247209	epsilon: 0.7085242000063277
steps: 123	lr: 0.0001	evaluation reward: 2.22	
episode: 1307	score: 1.0	memory length: 247377	epsilon: 0.7081915600063349
steps: 168	lr: 0.0001	evaluation reward: 2.2	
episode: 1308	score: 3.0	memory length: 247621	epsilon: 0.7077084400063454
steps: 244	lr: 0.0001	evaluation reward: 2.23	
episode: 1309	score: 4.0	memory length: 247880	epsilon: 0.7071956200063565
steps: 259	lr: 0.0001	evaluation reward: 2.25	
episode: 1310	score: 3.0	memory length: 248127	epsilon: 0.7067065600063671
steps: 247	lr: 0.0001	evaluation reward: 2.25	
episode: 1311	score: 3.0	memory length: 248374	epsilon: 0.7062175000063777
steps: 247	lr: 0.0001	evaluation reward: 2.23	
episode: 1312	score: 1.0	memory length: 248543	epsilon: 0.705882880006385
steps: 169	lr: 0.0001	evaluation reward: 2.23	
episode: 1313	score: 2.0	memory length: 248741	epsilon: 0.7054908400063935
steps: 198	lr: 0.0001	evaluation reward: 2.24	
episode: 1314	score: 5.0	memory length: 249071	epsilon: 0.7048374400064077
steps: 330	lr: 0.0001	evaluation reward: 2.27	
episode: 1315	score: 5.0	memory length: 249418	epsilon: 0.7041503800064226
steps: 347	lr: 0.0001	evaluation reward: 2.3	
episode: 1316	score: 2.0	memory length: 249636	epsilon: 0.703718740006432
steps: 218	lr: 0.0001	evaluation reward: 2.28	
episode: 1317	score: 3.0	memory length: 249902	epsilon: 0.7031920600064434
steps: 266	lr: 0.0001	evaluation reward: 2.29	
episode: 1318	score: 1.0	memory length: 250053	epsilon: 0.7028930800064499
steps: 151	lr: 0.0001	evaluation reward: 2.29	
episode: 1319	score: 3.0	memory length: 250300	epsilon: 0.7024040200064605
steps: 247	lr: 0.0001	evaluation reward: 2.32	
episode: 1320	score: 0.0	memory length: 250423	epsilon: 0.7021604800064658
steps: 123	lr: 0.0001	evaluation reward: 2.29	
episode: 1321	score: 0.0	memory length: 250545	epsilon: 0.701918920006471
steps: 122	lr: 0.0001	evaluation reward: 2.27	
episode: 1322	score: 2.0	memory length: 250763	epsilon: 0.7014872800064804
steps: 218	lr: 0.0001	evaluation reward: 2.24	
episode: 1323	score: 1.0	memory length: 250914	epsilon: 0.7011883000064869
steps: 151	lr: 0.0001	evaluation reward: 2.23	
episode: 1324	score: 2.0	memory length: 251112	epsilon: 0.7007962600064954
steps: 198	lr: 0.0001	evaluation reward: 2.22	
episode: 1325	score: 2.0	memory length: 251315	epsilon: 0.7003943200065041
steps: 203	lr: 0.0001	evaluation reward: 2.2	
episode: 1326	score: 0.0	memory length: 251437	epsilon: 0.7001527600065094
steps: 122	lr: 0.0001	evaluation reward: 2.18	
episode: 1327	score: 3.0	memory length: 251668	epsilon: 0.6996953800065193
steps: 231	lr: 0.0001	evaluation reward: 2.2	
episode: 1328	score: 2.0	memory length: 251891	epsilon: 0.6992538400065289

steps: 223	lr: 0.0001	evaluation reward: 2.21	
episode: 1329	score: 2.0	memory length: 252089	epsilon: 0.6988618000065374
steps: 198	lr: 0.0001	evaluation reward: 2.22	
episode: 1330	score: 2.0	memory length: 252307	epsilon: 0.6984301600065468
steps: 218	lr: 0.0001	evaluation reward: 2.22	
episode: 1331	score: 2.0	memory length: 252488	epsilon: 0.6980717800065546
steps: 181	lr: 0.0001	evaluation reward: 2.22	
episode: 1332	score: 3.0	memory length: 252714	epsilon: 0.6976243000065643
steps: 226	lr: 0.0001	evaluation reward: 2.21	
episode: 1333	score: 4.0	memory length: 252989	epsilon: 0.6970798000065761
steps: 275	lr: 0.0001	evaluation reward: 2.25	
episode: 1334	score: 2.0	memory length: 253169	epsilon: 0.6967234000065838
steps: 180	lr: 0.0001	evaluation reward: 2.26	
episode: 1335	score: 1.0	memory length: 253320	epsilon: 0.6964244200065903
steps: 151	lr: 0.0001	evaluation reward: 2.25	
episode: 1336	score: 0.0	memory length: 253443	epsilon: 0.6961808800065956
steps: 123	lr: 0.0001	evaluation reward: 2.25	
episode: 1337	score: 2.0	memory length: 253641	epsilon: 0.6957888400066041
steps: 198	lr: 0.0001	evaluation reward: 2.26	
episode: 1338	score: 2.0	memory length: 253822	epsilon: 0.6954304600066119
steps: 181	lr: 0.0001	evaluation reward: 2.26	
episode: 1339	score: 3.0	memory length: 254089	epsilon: 0.6949018000066234
steps: 267	lr: 0.0001	evaluation reward: 2.29	
episode: 1340	score: 1.0	memory length: 254258	epsilon: 0.6945671800066306
steps: 169	lr: 0.0001	evaluation reward: 2.3	
episode: 1341	score: 1.0	memory length: 254427	epsilon: 0.6942325600066379
steps: 169	lr: 0.0001	evaluation reward: 2.29	
episode: 1342	score: 0.0	memory length: 254550	epsilon: 0.6939890200066432
steps: 123	lr: 0.0001	evaluation reward: 2.28	
episode: 1343	score: 1.0	memory length: 254701	epsilon: 0.6936900400066497
steps: 151	lr: 0.0001	evaluation reward: 2.26	
episode: 1344	score: 5.0	memory length: 255020	epsilon: 0.6930584200066634
steps: 319	lr: 0.0001	evaluation reward: 2.24	
episode: 1345	score: 3.0	memory length: 255246	epsilon: 0.6926109400066731
steps: 226	lr: 0.0001	evaluation reward: 2.23	
episode: 1346	score: 2.0	memory length: 255445	epsilon: 0.6922169200066817
steps: 199	lr: 0.0001	evaluation reward: 2.2	
episode: 1347	score: 3.0	memory length: 255670	epsilon: 0.6917714200066913
steps: 225	lr: 0.0001	evaluation reward: 2.23	
episode: 1348	score: 2.0	memory length: 255854	epsilon: 0.6914071000066992
steps: 184	lr: 0.0001	evaluation reward: 2.24	
episode: 1349	score: 0.0	memory length: 255977	epsilon: 0.6911635600067045
steps: 123	lr: 0.0001	evaluation reward: 2.22	
episode: 1350	score: 4.0	memory length: 256235	epsilon: 0.6906527200067156
steps: 258	lr: 0.0001	evaluation reward: 2.23	
episode: 1351	score: 0.0	memory length: 256358	epsilon: 0.6904091800067209
steps: 123	lr: 0.0001	evaluation reward: 2.2	
episode: 1352	score: 4.0	memory length: 256632	epsilon: 0.6898666600067327

steps: 274	lr: 0.0001	evaluation reward: 2.2	
episode: 1353	score: 4.0	memory length: 256907	epsilon: 0.6893221600067445
steps: 275	lr: 0.0001	evaluation reward: 2.22	
episode: 1354	score: 4.0	memory length: 257172	epsilon: 0.6887974600067559
steps: 265	lr: 0.0001	evaluation reward: 2.23	
episode: 1355	score: 2.0	memory length: 257369	epsilon: 0.6884074000067644
steps: 197	lr: 0.0001	evaluation reward: 2.23	
episode: 1356	score: 2.0	memory length: 257567	epsilon: 0.6880153600067729
steps: 198	lr: 0.0001	evaluation reward: 2.2	
episode: 1357	score: 0.0	memory length: 257689	epsilon: 0.6877738000067781
steps: 122	lr: 0.0001	evaluation reward: 2.17	
episode: 1358	score: 3.0	memory length: 257937	epsilon: 0.6872827600067888
steps: 248	lr: 0.0001	evaluation reward: 2.19	
episode: 1359	score: 2.0	memory length: 258155	epsilon: 0.6868511200067982
steps: 218	lr: 0.0001	evaluation reward: 2.16	
episode: 1360	score: 3.0	memory length: 258381	epsilon: 0.6864036400068079
steps: 226	lr: 0.0001	evaluation reward: 2.17	
episode: 1361	score: 3.0	memory length: 258628	epsilon: 0.6859145800068185
steps: 247	lr: 0.0001	evaluation reward: 2.18	
episode: 1362	score: 1.0	memory length: 258799	epsilon: 0.6855760000068258
steps: 171	lr: 0.0001	evaluation reward: 2.18	
episode: 1363	score: 1.0	memory length: 258950	epsilon: 0.6852770200068323
steps: 151	lr: 0.0001	evaluation reward: 2.16	
episode: 1364	score: 0.0	memory length: 259073	epsilon: 0.6850334800068376
steps: 123	lr: 0.0001	evaluation reward: 2.14	
episode: 1365	score: 1.0	memory length: 259223	epsilon: 0.6847364800068441
steps: 150	lr: 0.0001	evaluation reward: 2.12	
episode: 1366	score: 2.0	memory length: 259405	epsilon: 0.6843761200068519
steps: 182	lr: 0.0001	evaluation reward: 2.11	
episode: 1367	score: 2.0	memory length: 259603	epsilon: 0.6839840800068604
steps: 198	lr: 0.0001	evaluation reward: 2.12	
episode: 1368	score: 3.0	memory length: 259834	epsilon: 0.6835267000068703
steps: 231	lr: 0.0001	evaluation reward: 2.15	
episode: 1369	score: 2.0	memory length: 260051	epsilon: 0.6830970400068797
steps: 217	lr: 0.0001	evaluation reward: 2.17	
episode: 1370	score: 1.0	memory length: 260220	epsilon: 0.6827624200068869
steps: 169	lr: 0.0001	evaluation reward: 2.16	
episode: 1371	score: 2.0	memory length: 260418	epsilon: 0.6823703800068954
steps: 198	lr: 0.0001	evaluation reward: 2.18	
episode: 1372	score: 0.0	memory length: 260540	epsilon: 0.6821288200069007
steps: 122	lr: 0.0001	evaluation reward: 2.15	
episode: 1373	score: 4.0	memory length: 260852	epsilon: 0.6815110600069141
steps: 312	lr: 0.0001	evaluation reward: 2.17	
episode: 1374	score: 1.0	memory length: 261023	epsilon: 0.6811724800069214
steps: 171	lr: 0.0001	evaluation reward: 2.13	
episode: 1375	score: 2.0	memory length: 261203	epsilon: 0.6808160800069292
steps: 180	lr: 0.0001	evaluation reward: 2.11	
episode: 1376	score: 2.0	memory length: 261384	epsilon: 0.680457700006937

steps: 181	lr: 0.0001	evaluation reward: 2.1	
episode: 1377	score: 1.0	memory length: 261535	epsilon: 0.6801587200069434
steps: 151	lr: 0.0001	evaluation reward: 2.08	
episode: 1378	score: 3.0	memory length: 261763	epsilon: 0.6797072800069532
steps: 228	lr: 0.0001	evaluation reward: 2.09	
episode: 1379	score: 2.0	memory length: 261961	epsilon: 0.6793152400069618
steps: 198	lr: 0.0001	evaluation reward: 2.11	
episode: 1380	score: 5.0	memory length: 262286	epsilon: 0.6786717400069757
steps: 325	lr: 0.0001	evaluation reward: 2.13	
episode: 1381	score: 4.0	memory length: 262562	epsilon: 0.6781252600069876
steps: 276	lr: 0.0001	evaluation reward: 2.14	
episode: 1382	score: 4.0	memory length: 262856	epsilon: 0.6775431400070002
steps: 294	lr: 0.0001	evaluation reward: 2.17	
episode: 1383	score: 0.0	memory length: 262979	epsilon: 0.6772996000070055
steps: 123	lr: 0.0001	evaluation reward: 2.14	
episode: 1384	score: 1.0	memory length: 263129	epsilon: 0.677002600007012
steps: 150	lr: 0.0001	evaluation reward: 2.12	
episode: 1385	score: 5.0	memory length: 263473	epsilon: 0.6763214800070267
steps: 344	lr: 0.0001	evaluation reward: 2.13	
episode: 1386	score: 3.0	memory length: 263684	epsilon: 0.6759037000070358
steps: 211	lr: 0.0001	evaluation reward: 2.15	
episode: 1387	score: 5.0	memory length: 264004	epsilon: 0.6752701000070496
steps: 320	lr: 0.0001	evaluation reward: 2.19	
episode: 1388	score: 5.0	memory length: 264329	epsilon: 0.6746266000070635
steps: 325	lr: 0.0001	evaluation reward: 2.22	
episode: 1389	score: 1.0	memory length: 264479	epsilon: 0.67432960000707
steps: 150	lr: 0.0001	evaluation reward: 2.19	
episode: 1390	score: 2.0	memory length: 264677	epsilon: 0.6739375600070785
steps: 198	lr: 0.0001	evaluation reward: 2.21	
episode: 1391	score: 1.0	memory length: 264828	epsilon: 0.673638580007085
steps: 151	lr: 0.0001	evaluation reward: 2.19	
episode: 1392	score: 1.0	memory length: 264996	epsilon: 0.6733059400070922
steps: 168	lr: 0.0001	evaluation reward: 2.18	
episode: 1393	score: 3.0	memory length: 265264	epsilon: 0.6727753000071037
steps: 268	lr: 0.0001	evaluation reward: 2.18	
episode: 1394	score: 1.0	memory length: 265415	epsilon: 0.6724763200071102
steps: 151	lr: 0.0001	evaluation reward: 2.16	
episode: 1395	score: 2.0	memory length: 265613	epsilon: 0.6720842800071187
steps: 198	lr: 0.0001	evaluation reward: 2.15	
episode: 1396	score: 4.0	memory length: 265908	epsilon: 0.6715001800071314
steps: 295	lr: 0.0001	evaluation reward: 2.18	
episode: 1397	score: 1.0	memory length: 266059	epsilon: 0.6712012000071379
steps: 151	lr: 0.0001	evaluation reward: 2.16	
episode: 1398	score: 2.0	memory length: 266257	epsilon: 0.6708091600071464
steps: 198	lr: 0.0001	evaluation reward: 2.16	
episode: 1399	score: 1.0	memory length: 266408	epsilon: 0.6705101800071529
steps: 151	lr: 0.0001	evaluation reward: 2.16	
episode: 1400	score: 3.0	memory length: 266637	epsilon: 0.6700567600071627

steps: 229	lr: 0.0001	evaluation reward: 2.15	
episode: 1401	score: 3.0	memory length: 266884	epsilon: 0.6695677000071734
steps: 247	lr: 0.0001	evaluation reward: 2.16	
episode: 1402	score: 2.0	memory length: 267085	epsilon: 0.669169720007182
steps: 201	lr: 0.0001	evaluation reward: 2.17	
episode: 1403	score: 2.0	memory length: 267283	epsilon: 0.6687776800071905
steps: 198	lr: 0.0001	evaluation reward: 2.15	
episode: 1404	score: 6.0	memory length: 267633	epsilon: 0.6680846800072056
steps: 350	lr: 0.0001	evaluation reward: 2.18	
episode: 1405	score: 6.0	memory length: 268027	epsilon: 0.6673045600072225
steps: 394	lr: 0.0001	evaluation reward: 2.23	
episode: 1406	score: 3.0	memory length: 268291	epsilon: 0.6667818400072338
steps: 264	lr: 0.0001	evaluation reward: 2.26	
episode: 1407	score: 3.0	memory length: 268519	epsilon: 0.6663304000072436
steps: 228	lr: 0.0001	evaluation reward: 2.28	
episode: 1408	score: 2.0	memory length: 268701	epsilon: 0.6659700400072515
steps: 182	lr: 0.0001	evaluation reward: 2.27	
episode: 1409	score: 4.0	memory length: 268974	epsilon: 0.6654295000072632
steps: 273	lr: 0.0001	evaluation reward: 2.27	
episode: 1410	score: 2.0	memory length: 269172	epsilon: 0.6650374600072717
steps: 198	lr: 0.0001	evaluation reward: 2.26	
episode: 1411	score: 5.0	memory length: 269493	epsilon: 0.6644018800072855
steps: 321	lr: 0.0001	evaluation reward: 2.28	
episode: 1412	score: 3.0	memory length: 269739	epsilon: 0.6639148000072961
steps: 246	lr: 0.0001	evaluation reward: 2.3	
episode: 1413	score: 0.0	memory length: 269861	epsilon: 0.6636732400073013
steps: 122	lr: 0.0001	evaluation reward: 2.28	
episode: 1414	score: 3.0	memory length: 270106	epsilon: 0.6631881400073119
steps: 245	lr: 0.0001	evaluation reward: 2.26	
episode: 1415	score: 2.0	memory length: 270304	epsilon: 0.6627961000073204
steps: 198	lr: 0.0001	evaluation reward: 2.23	
episode: 1416	score: 3.0	memory length: 270534	epsilon: 0.6623407000073303
steps: 230	lr: 0.0001	evaluation reward: 2.24	
episode: 1417	score: 2.0	memory length: 270752	epsilon: 0.6619090600073396
steps: 218	lr: 0.0001	evaluation reward: 2.23	
episode: 1418	score: 4.0	memory length: 271027	epsilon: 0.6613645600073514
steps: 275	lr: 0.0001	evaluation reward: 2.26	
episode: 1419	score: 4.0	memory length: 271304	epsilon: 0.6608161000073633
steps: 277	lr: 0.0001	evaluation reward: 2.27	
episode: 1420	score: 4.0	memory length: 271604	epsilon: 0.6602221000073762
steps: 300	lr: 0.0001	evaluation reward: 2.31	
episode: 1421	score: 2.0	memory length: 271786	epsilon: 0.6598617400073841
steps: 182	lr: 0.0001	evaluation reward: 2.33	
episode: 1422	score: 1.0	memory length: 271957	epsilon: 0.6595231600073914
steps: 171	lr: 0.0001	evaluation reward: 2.32	
episode: 1423	score: 3.0	memory length: 272203	epsilon: 0.659036080007402
steps: 246	lr: 0.0001	evaluation reward: 2.34	
episode: 1424	score: 5.0	memory length: 272547	epsilon: 0.6583549600074168

steps: 344	lr: 0.0001	evaluation reward: 2.37	
episode: 1425	score: 3.0	memory length: 272796	epsilon: 0.6578619400074275
steps: 249	lr: 0.0001	evaluation reward: 2.38	
episode: 1426	score: 4.0	memory length: 273042	epsilon: 0.657374860007438
steps: 246	lr: 0.0001	evaluation reward: 2.42	
episode: 1427	score: 6.0	memory length: 273417	epsilon: 0.6566323600074542
steps: 375	lr: 0.0001	evaluation reward: 2.45	
episode: 1428	score: 3.0	memory length: 273648	epsilon: 0.6561749800074641
steps: 231	lr: 0.0001	evaluation reward: 2.46	
episode: 1429	score: 4.0	memory length: 273908	epsilon: 0.6556601800074753
steps: 260	lr: 0.0001	evaluation reward: 2.48	
episode: 1430	score: 1.0	memory length: 274059	epsilon: 0.6553612000074818
steps: 151	lr: 0.0001	evaluation reward: 2.47	
episode: 1431	score: 6.0	memory length: 274401	epsilon: 0.6546840400074965
steps: 342	lr: 0.0001	evaluation reward: 2.51	
episode: 1432	score: 1.0	memory length: 274552	epsilon: 0.654385060007503
steps: 151	lr: 0.0001	evaluation reward: 2.49	
episode: 1433	score: 5.0	memory length: 274912	epsilon: 0.6536722600075184
steps: 360	lr: 0.0001	evaluation reward: 2.5	
episode: 1434	score: 4.0	memory length: 275169	epsilon: 0.6531634000075295
steps: 257	lr: 0.0001	evaluation reward: 2.52	
episode: 1435	score: 3.0	memory length: 275413	epsilon: 0.65268028000754
steps: 244	lr: 0.0001	evaluation reward: 2.54	
episode: 1436	score: 4.0	memory length: 275686	epsilon: 0.6521397400075517
steps: 273	lr: 0.0001	evaluation reward: 2.58	
episode: 1437	score: 4.0	memory length: 275963	epsilon: 0.6515912800075636
steps: 277	lr: 0.0001	evaluation reward: 2.6	
episode: 1438	score: 3.0	memory length: 276189	epsilon: 0.6511438000075733
steps: 226	lr: 0.0001	evaluation reward: 2.61	
episode: 1439	score: 4.0	memory length: 276464	epsilon: 0.6505993000075851
steps: 275	lr: 0.0001	evaluation reward: 2.62	
episode: 1440	score: 1.0	memory length: 276635	epsilon: 0.6502607200075925
steps: 171	lr: 0.0001	evaluation reward: 2.62	
episode: 1441	score: 1.0	memory length: 276806	epsilon: 0.6499221400075998
steps: 171	lr: 0.0001	evaluation reward: 2.62	
episode: 1442	score: 3.0	memory length: 277031	epsilon: 0.6494766400076095
steps: 225	lr: 0.0001	evaluation reward: 2.65	
episode: 1443	score: 3.0	memory length: 277298	epsilon: 0.648947980007621
steps: 267	lr: 0.0001	evaluation reward: 2.67	
episode: 1444	score: 1.0	memory length: 277449	epsilon: 0.6486490000076275
steps: 151	lr: 0.0001	evaluation reward: 2.63	
episode: 1445	score: 2.0	memory length: 277647	epsilon: 0.648256960007636
steps: 198	lr: 0.0001	evaluation reward: 2.62	
episode: 1446	score: 1.0	memory length: 277818	epsilon: 0.6479183800076433
steps: 171	lr: 0.0001	evaluation reward: 2.61	
episode: 1447	score: 3.0	memory length: 278064	epsilon: 0.6474313000076539
steps: 246	lr: 0.0001	evaluation reward: 2.61	
episode: 1448	score: 3.0	memory length: 278295	epsilon: 0.6469739200076638

steps: 231	lr: 0.0001	evaluation reward: 2.62	
episode: 1449	score: 2.0	memory length: 278476	epsilon: 0.6466155400076716
steps: 181	lr: 0.0001	evaluation reward: 2.64	
episode: 1450	score: 1.0	memory length: 278626	epsilon: 0.6463185400076781
steps: 150	lr: 0.0001	evaluation reward: 2.61	
episode: 1451	score: 4.0	memory length: 278922	epsilon: 0.6457324600076908
steps: 296	lr: 0.0001	evaluation reward: 2.65	
episode: 1452	score: 3.0	memory length: 279148	epsilon: 0.6452849800077005
steps: 226	lr: 0.0001	evaluation reward: 2.64	
episode: 1453	score: 2.0	memory length: 279346	epsilon: 0.644892940007709
steps: 198	lr: 0.0001	evaluation reward: 2.62	
episode: 1454	score: 4.0	memory length: 279627	epsilon: 0.6443365600077211
steps: 281	lr: 0.0001	evaluation reward: 2.62	
episode: 1455	score: 4.0	memory length: 279886	epsilon: 0.6438237400077322
steps: 259	lr: 0.0001	evaluation reward: 2.64	
episode: 1456	score: 3.0	memory length: 280111	epsilon: 0.6433782400077419
steps: 225	lr: 0.0001	evaluation reward: 2.65	
episode: 1457	score: 0.0	memory length: 280234	epsilon: 0.6431347000077472
steps: 123	lr: 0.0001	evaluation reward: 2.65	
episode: 1458	score: 4.0	memory length: 280512	epsilon: 0.6425842600077591
steps: 278	lr: 0.0001	evaluation reward: 2.66	
episode: 1459	score: 3.0	memory length: 280759	epsilon: 0.6420952000077698
steps: 247	lr: 0.0001	evaluation reward: 2.67	
episode: 1460	score: 5.0	memory length: 281063	epsilon: 0.6414932800077828
steps: 304	lr: 0.0001	evaluation reward: 2.69	
episode: 1461	score: 2.0	memory length: 281281	epsilon: 0.6410616400077922
steps: 218	lr: 0.0001	evaluation reward: 2.68	
episode: 1462	score: 3.0	memory length: 281494	epsilon: 0.6406399000078014
steps: 213	lr: 0.0001	evaluation reward: 2.7	
episode: 1463	score: 3.0	memory length: 281723	epsilon: 0.6401864800078112
steps: 229	lr: 0.0001	evaluation reward: 2.72	
episode: 1464	score: 0.0	memory length: 281845	epsilon: 0.6399449200078164
steps: 122	lr: 0.0001	evaluation reward: 2.72	
episode: 1465	score: 8.0	memory length: 282311	epsilon: 0.6390222400078365
steps: 466	lr: 0.0001	evaluation reward: 2.79	
episode: 1466	score: 3.0	memory length: 282557	epsilon: 0.638535160007847
steps: 246	lr: 0.0001	evaluation reward: 2.8	
episode: 1467	score: 2.0	memory length: 282774	epsilon: 0.6381055000078564
steps: 217	lr: 0.0001	evaluation reward: 2.8	
episode: 1468	score: 3.0	memory length: 282987	epsilon: 0.6376837600078655
steps: 213	lr: 0.0001	evaluation reward: 2.8	
episode: 1469	score: 4.0	memory length: 283280	epsilon: 0.6371036200078781
steps: 293	lr: 0.0001	evaluation reward: 2.82	
episode: 1470	score: 2.0	memory length: 283497	epsilon: 0.6366739600078875
steps: 217	lr: 0.0001	evaluation reward: 2.83	
episode: 1471	score: 4.0	memory length: 283792	epsilon: 0.6360898600079001
steps: 295	lr: 0.0001	evaluation reward: 2.85	
episode: 1472	score: 2.0	memory length: 284008	epsilon: 0.6356621800079094

steps: 216	lr: 0.0001	evaluation reward: 2.87	
episode: 1473	score: 1.0	memory length: 284159	epsilon: 0.6353632000079159
steps: 151	lr: 0.0001	evaluation reward: 2.84	
episode: 1474	score: 7.0	memory length: 284576	epsilon: 0.6345375400079338
steps: 417	lr: 0.0001	evaluation reward: 2.9	
episode: 1475	score: 2.0	memory length: 284773	epsilon: 0.6341474800079423
steps: 197	lr: 0.0001	evaluation reward: 2.9	
episode: 1476	score: 0.0	memory length: 284896	epsilon: 0.6339039400079476
steps: 123	lr: 0.0001	evaluation reward: 2.88	
episode: 1477	score: 3.0	memory length: 285143	epsilon: 0.6334148800079582
steps: 247	lr: 0.0001	evaluation reward: 2.9	
episode: 1478	score: 7.0	memory length: 285511	epsilon: 0.632686240007974
steps: 368	lr: 0.0001	evaluation reward: 2.94	
episode: 1479	score: 0.0	memory length: 285634	epsilon: 0.6324427000079793
steps: 123	lr: 0.0001	evaluation reward: 2.92	
episode: 1480	score: 3.0	memory length: 285861	epsilon: 0.6319932400079891
steps: 227	lr: 0.0001	evaluation reward: 2.9	
episode: 1481	score: 4.0	memory length: 286136	epsilon: 0.6314487400080009
steps: 275	lr: 0.0001	evaluation reward: 2.9	
episode: 1482	score: 3.0	memory length: 286345	epsilon: 0.6310349200080099
steps: 209	lr: 0.0001	evaluation reward: 2.89	
episode: 1483	score: 1.0	memory length: 286496	epsilon: 0.6307359400080164
steps: 151	lr: 0.0001	evaluation reward: 2.9	
episode: 1484	score: 2.0	memory length: 286693	epsilon: 0.6303458800080248
steps: 197	lr: 0.0001	evaluation reward: 2.91	
episode: 1485	score: 2.0	memory length: 286891	epsilon: 0.6299538400080333
steps: 198	lr: 0.0001	evaluation reward: 2.88	
episode: 1486	score: 2.0	memory length: 287089	epsilon: 0.6295618000080418
steps: 198	lr: 0.0001	evaluation reward: 2.87	
episode: 1487	score: 2.0	memory length: 287308	epsilon: 0.6291281800080513
steps: 219	lr: 0.0001	evaluation reward: 2.84	
episode: 1488	score: 0.0	memory length: 287431	epsilon: 0.6288846400080566
steps: 123	lr: 0.0001	evaluation reward: 2.79	
episode: 1489	score: 5.0	memory length: 287735	epsilon: 0.6282827200080696
steps: 304	lr: 0.0001	evaluation reward: 2.83	
episode: 1490	score: 3.0	memory length: 287966	epsilon: 0.6278253400080795
steps: 231	lr: 0.0001	evaluation reward: 2.84	
episode: 1491	score: 2.0	memory length: 288181	epsilon: 0.6273996400080888
steps: 215	lr: 0.0001	evaluation reward: 2.85	
episode: 1492	score: 1.0	memory length: 288332	epsilon: 0.6271006600080953
steps: 151	lr: 0.0001	evaluation reward: 2.85	
episode: 1493	score: 4.0	memory length: 288592	epsilon: 0.6265858600081065
steps: 260	lr: 0.0001	evaluation reward: 2.86	
episode: 1494	score: 3.0	memory length: 288838	epsilon: 0.626098780008117
steps: 246	lr: 0.0001	evaluation reward: 2.88	
episode: 1495	score: 2.0	memory length: 289037	epsilon: 0.6257047600081256
steps: 199	lr: 0.0001	evaluation reward: 2.88	
episode: 1496	score: 3.0	memory length: 289282	epsilon: 0.6252196600081361

steps: 245	lr: 0.0001	evaluation reward: 2.87	
episode: 1497	score: 2.0	memory length: 289462	epsilon: 0.6248632600081439
steps: 180	lr: 0.0001	evaluation reward: 2.88	
episode: 1498	score: 1.0	memory length: 289613	epsilon: 0.6245642800081503
steps: 151	lr: 0.0001	evaluation reward: 2.87	
episode: 1499	score: 1.0	memory length: 289783	epsilon: 0.6242276800081576
steps: 170	lr: 0.0001	evaluation reward: 2.87	
episode: 1500	score: 2.0	memory length: 289981	epsilon: 0.6238356400081662
steps: 198	lr: 0.0001	evaluation reward: 2.86	
episode: 1501	score: 2.0	memory length: 290179	epsilon: 0.6234436000081747
steps: 198	lr: 0.0001	evaluation reward: 2.85	
episode: 1502	score: 6.0	memory length: 290544	epsilon: 0.6227209000081904
steps: 365	lr: 0.0001	evaluation reward: 2.89	
episode: 1503	score: 2.0	memory length: 290726	epsilon: 0.6223605400081982
steps: 182	lr: 0.0001	evaluation reward: 2.89	
episode: 1504	score: 5.0	memory length: 291030	epsilon: 0.6217586200082112
steps: 304	lr: 0.0001	evaluation reward: 2.88	
episode: 1505	score: 3.0	memory length: 291240	epsilon: 0.6213428200082203
steps: 210	lr: 0.0001	evaluation reward: 2.85	
episode: 1506	score: 6.0	memory length: 291597	epsilon: 0.6206359600082356
steps: 357	lr: 0.0001	evaluation reward: 2.88	
episode: 1507	score: 3.0	memory length: 291822	epsilon: 0.6201904600082453
steps: 225	lr: 0.0001	evaluation reward: 2.88	
episode: 1508	score: 2.0	memory length: 292021	epsilon: 0.6197964400082538
steps: 199	lr: 0.0001	evaluation reward: 2.88	
episode: 1509	score: 0.0	memory length: 292144	epsilon: 0.6195529000082591
steps: 123	lr: 0.0001	evaluation reward: 2.84	
episode: 1510	score: 5.0	memory length: 292469	epsilon: 0.6189094000082731
steps: 325	lr: 0.0001	evaluation reward: 2.87	
episode: 1511	score: 2.0	memory length: 292686	epsilon: 0.6184797400082824
steps: 217	lr: 0.0001	evaluation reward: 2.84	
episode: 1512	score: 3.0	memory length: 292933	epsilon: 0.617990680008293
steps: 247	lr: 0.0001	evaluation reward: 2.84	
episode: 1513	score: 8.0	memory length: 293284	epsilon: 0.6172957000083081
steps: 351	lr: 0.0001	evaluation reward: 2.92	
episode: 1514	score: 6.0	memory length: 293622	epsilon: 0.6166264600083227
steps: 338	lr: 0.0001	evaluation reward: 2.95	
episode: 1515	score: 1.0	memory length: 293792	epsilon: 0.61628986000833
steps: 170	lr: 0.0001	evaluation reward: 2.94	
episode: 1516	score: 2.0	memory length: 294008	epsilon: 0.6158621800083393
steps: 216	lr: 0.0001	evaluation reward: 2.93	
episode: 1517	score: 2.0	memory length: 294206	epsilon: 0.6154701400083478
steps: 198	lr: 0.0001	evaluation reward: 2.93	
episode: 1518	score: 2.0	memory length: 294388	epsilon: 0.6151097800083556
steps: 182	lr: 0.0001	evaluation reward: 2.91	
episode: 1519	score: 4.0	memory length: 294664	epsilon: 0.6145633000083675
steps: 276	lr: 0.0001	evaluation reward: 2.91	
episode: 1520	score: 2.0	memory length: 294844	epsilon: 0.6142069000083752

steps: 180	lr: 0.0001	evaluation reward: 2.89	
episode: 1521	score: 1.0	memory length: 294995	epsilon: 0.6139079200083817
steps: 151	lr: 0.0001	evaluation reward: 2.88	
episode: 1522	score: 9.0	memory length: 295306	epsilon: 0.613292140008395
steps: 311	lr: 0.0001	evaluation reward: 2.96	
episode: 1523	score: 3.0	memory length: 295519	epsilon: 0.6128704000084042
steps: 213	lr: 0.0001	evaluation reward: 2.96	
episode: 1524	score: 2.0	memory length: 295716	epsilon: 0.6124803400084127
steps: 197	lr: 0.0001	evaluation reward: 2.93	
episode: 1525	score: 3.0	memory length: 295961	epsilon: 0.6119952400084232
steps: 245	lr: 0.0001	evaluation reward: 2.93	
episode: 1526	score: 2.0	memory length: 296179	epsilon: 0.6115636000084326
steps: 218	lr: 0.0001	evaluation reward: 2.91	
episode: 1527	score: 1.0	memory length: 296329	epsilon: 0.611266600008439
steps: 150	lr: 0.0001	evaluation reward: 2.86	
episode: 1528	score: 3.0	memory length: 296595	epsilon: 0.6107399200084505
steps: 266	lr: 0.0001	evaluation reward: 2.86	
episode: 1529	score: 4.0	memory length: 296869	epsilon: 0.6101974000084622
steps: 274	lr: 0.0001	evaluation reward: 2.86	
episode: 1530	score: 4.0	memory length: 297161	epsilon: 0.6096192400084748
steps: 292	lr: 0.0001	evaluation reward: 2.89	
episode: 1531	score: 5.0	memory length: 297470	epsilon: 0.6090074200084881
steps: 309	lr: 0.0001	evaluation reward: 2.88	
episode: 1532	score: 5.0	memory length: 297796	epsilon: 0.6083619400085021
steps: 326	lr: 0.0001	evaluation reward: 2.92	
episode: 1533	score: 2.0	memory length: 297976	epsilon: 0.6080055400085098
steps: 180	lr: 0.0001	evaluation reward: 2.89	
episode: 1534	score: 3.0	memory length: 298221	epsilon: 0.6075204400085203
steps: 245	lr: 0.0001	evaluation reward: 2.88	
episode: 1535	score: 6.0	memory length: 298550	epsilon: 0.6068690200085345
steps: 329	lr: 0.0001	evaluation reward: 2.91	
episode: 1536	score: 1.0	memory length: 298701	epsilon: 0.606570040008541
steps: 151	lr: 0.0001	evaluation reward: 2.88	
episode: 1537	score: 3.0	memory length: 298948	epsilon: 0.6060809800085516
steps: 247	lr: 0.0001	evaluation reward: 2.87	
episode: 1538	score: 7.0	memory length: 299327	epsilon: 0.6053305600085679
steps: 379	lr: 0.0001	evaluation reward: 2.91	
episode: 1539	score: 3.0	memory length: 299553	epsilon: 0.6048830800085776
steps: 226	lr: 0.0001	evaluation reward: 2.9	
episode: 1540	score: 5.0	memory length: 299877	epsilon: 0.6042415600085915
steps: 324	lr: 0.0001	evaluation reward: 2.94	
episode: 1541	score: 2.0	memory length: 300059	epsilon: 0.6038812000085994
steps: 182	lr: 0.0001	evaluation reward: 2.95	
episode: 1542	score: 1.0	memory length: 300229	epsilon: 0.6035446000086067
steps: 170	lr: 0.0001	evaluation reward: 2.93	
episode: 1543	score: 6.0	memory length: 300586	epsilon: 0.602837740008622
steps: 357	lr: 0.0001	evaluation reward: 2.96	
episode: 1544	score: 3.0	memory length: 300815	epsilon: 0.6023843200086318

steps: 229	lr: 0.0001	evaluation reward: 2.98	
episode: 1545	score: 4.0	memory length: 301094	epsilon: 0.6018319000086438
steps: 279	lr: 0.0001	evaluation reward: 3.0	
episode: 1546	score: 3.0	memory length: 301305	epsilon: 0.6014141200086529
steps: 211	lr: 0.0001	evaluation reward: 3.02	
episode: 1547	score: 3.0	memory length: 301516	epsilon: 0.600996340008662
steps: 211	lr: 0.0001	evaluation reward: 3.02	
episode: 1548	score: 3.0	memory length: 301780	epsilon: 0.6004736200086733
steps: 264	lr: 0.0001	evaluation reward: 3.02	
episode: 1549	score: 3.0	memory length: 302026	epsilon: 0.5999865400086839
steps: 246	lr: 0.0001	evaluation reward: 3.03	
episode: 1550	score: 3.0	memory length: 302273	epsilon: 0.5994974800086945
steps: 247	lr: 0.0001	evaluation reward: 3.05	
episode: 1551	score: 4.0	memory length: 302550	epsilon: 0.5989490200087064
steps: 277	lr: 0.0001	evaluation reward: 3.05	
episode: 1552	score: 4.0	memory length: 302826	epsilon: 0.5984025400087183
steps: 276	lr: 0.0001	evaluation reward: 3.06	
episode: 1553	score: 5.0	memory length: 303142	epsilon: 0.5977768600087319
steps: 316	lr: 0.0001	evaluation reward: 3.09	
episode: 1554	score: 1.0	memory length: 303293	epsilon: 0.5974778800087384
steps: 151	lr: 0.0001	evaluation reward: 3.06	
episode: 1555	score: 2.0	memory length: 303491	epsilon: 0.5970858400087469
steps: 198	lr: 0.0001	evaluation reward: 3.04	
episode: 1556	score: 4.0	memory length: 303741	epsilon: 0.5965908400087576
steps: 250	lr: 0.0001	evaluation reward: 3.05	
episode: 1557	score: 5.0	memory length: 304056	epsilon: 0.5959671400087712
steps: 315	lr: 0.0001	evaluation reward: 3.1	
episode: 1558	score: 1.0	memory length: 304207	epsilon: 0.5956681600087776
steps: 151	lr: 0.0001	evaluation reward: 3.07	
episode: 1559	score: 4.0	memory length: 304501	epsilon: 0.5950860400087903
steps: 294	lr: 0.0001	evaluation reward: 3.08	
episode: 1560	score: 7.0	memory length: 304856	epsilon: 0.5943831400088055
steps: 355	lr: 0.0001	evaluation reward: 3.1	
episode: 1561	score: 2.0	memory length: 305054	epsilon: 0.593991100008814
steps: 198	lr: 0.0001	evaluation reward: 3.1	
episode: 1562	score: 3.0	memory length: 305267	epsilon: 0.5935693600088232
steps: 213	lr: 0.0001	evaluation reward: 3.1	
episode: 1563	score: 7.0	memory length: 305652	epsilon: 0.5928070600088398
steps: 385	lr: 0.0001	evaluation reward: 3.14	
episode: 1564	score: 1.0	memory length: 305803	epsilon: 0.5925080800088462
steps: 151	lr: 0.0001	evaluation reward: 3.15	
episode: 1565	score: 2.0	memory length: 306006	epsilon: 0.592106140008855
steps: 203	lr: 0.0001	evaluation reward: 3.09	
episode: 1566	score: 0.0	memory length: 306129	epsilon: 0.5918626000088603
steps: 123	lr: 0.0001	evaluation reward: 3.06	
episode: 1567	score: 1.0	memory length: 306298	epsilon: 0.5915279800088675
steps: 169	lr: 0.0001	evaluation reward: 3.05	
episode: 1568	score: 2.0	memory length: 306516	epsilon: 0.5910963400088769

steps: 218	lr: 0.0001	evaluation reward: 3.04	
episode: 1569	score: 3.0	memory length: 306744	epsilon: 0.5906449000088867
steps: 228	lr: 0.0001	evaluation reward: 3.03	
episode: 1570	score: 7.0	memory length: 307067	epsilon: 0.5900053600089006
steps: 323	lr: 0.0001	evaluation reward: 3.08	
episode: 1571	score: 2.0	memory length: 307265	epsilon: 0.5896133200089091
steps: 198	lr: 0.0001	evaluation reward: 3.06	
episode: 1572	score: 3.0	memory length: 307491	epsilon: 0.5891658400089188
steps: 226	lr: 0.0001	evaluation reward: 3.07	
episode: 1573	score: 1.0	memory length: 307660	epsilon: 0.5888312200089261
steps: 169	lr: 0.0001	evaluation reward: 3.07	
episode: 1574	score: 2.0	memory length: 307858	epsilon: 0.5884391800089346
steps: 198	lr: 0.0001	evaluation reward: 3.02	
episode: 1575	score: 6.0	memory length: 308248	epsilon: 0.5876669800089513
steps: 390	lr: 0.0001	evaluation reward: 3.06	
episode: 1576	score: 4.0	memory length: 308527	epsilon: 0.5871145600089633
steps: 279	lr: 0.0001	evaluation reward: 3.1	
episode: 1577	score: 3.0	memory length: 308756	epsilon: 0.5866611400089732
steps: 229	lr: 0.0001	evaluation reward: 3.1	
episode: 1578	score: 0.0	memory length: 308878	epsilon: 0.5864195800089784
steps: 122	lr: 0.0001	evaluation reward: 3.03	
episode: 1579	score: 1.0	memory length: 309028	epsilon: 0.5861225800089849
steps: 150	lr: 0.0001	evaluation reward: 3.04	
episode: 1580	score: 3.0	memory length: 309238	epsilon: 0.5857067800089939
steps: 210	lr: 0.0001	evaluation reward: 3.04	
episode: 1581	score: 7.0	memory length: 309674	epsilon: 0.5848435000090126
steps: 436	lr: 0.0001	evaluation reward: 3.07	
episode: 1582	score: 4.0	memory length: 309951	epsilon: 0.5842950400090245
steps: 277	lr: 0.0001	evaluation reward: 3.08	
episode: 1583	score: 1.0	memory length: 310102	epsilon: 0.583996060009031
steps: 151	lr: 0.0001	evaluation reward: 3.08	
episode: 1584	score: 2.0	memory length: 310300	epsilon: 0.5836040200090395
steps: 198	lr: 0.0001	evaluation reward: 3.08	
episode: 1585	score: 5.0	memory length: 310623	epsilon: 0.5829644800090534
steps: 323	lr: 0.0001	evaluation reward: 3.11	
episode: 1586	score: 3.0	memory length: 310835	epsilon: 0.5825447200090625
steps: 212	lr: 0.0001	evaluation reward: 3.12	
episode: 1587	score: 1.0	memory length: 310986	epsilon: 0.582245740009069
steps: 151	lr: 0.0001	evaluation reward: 3.11	
episode: 1588	score: 5.0	memory length: 311330	epsilon: 0.5815646200090838
steps: 344	lr: 0.0001	evaluation reward: 3.16	
episode: 1589	score: 2.0	memory length: 311527	epsilon: 0.5811745600090923
steps: 197	lr: 0.0001	evaluation reward: 3.13	
episode: 1590	score: 4.0	memory length: 311786	epsilon: 0.5806617400091034
steps: 259	lr: 0.0001	evaluation reward: 3.14	
episode: 1591	score: 3.0	memory length: 312012	epsilon: 0.5802142600091131
steps: 226	lr: 0.0001	evaluation reward: 3.15	
episode: 1592	score: 4.0	memory length: 312269	epsilon: 0.5797054000091242

steps: 257	lr: 0.0001	evaluation reward: 3.18	
episode: 1593	score: 4.0	memory length: 312510	epsilon: 0.5792282200091345
steps: 241	lr: 0.0001	evaluation reward: 3.18	
episode: 1594	score: 1.0	memory length: 312681	epsilon: 0.5788896400091419
steps: 171	lr: 0.0001	evaluation reward: 3.16	
episode: 1595	score: 1.0	memory length: 312832	epsilon: 0.5785906600091484
steps: 151	lr: 0.0001	evaluation reward: 3.15	
episode: 1596	score: 7.0	memory length: 313255	epsilon: 0.5777531200091666
steps: 423	lr: 0.0001	evaluation reward: 3.19	
episode: 1597	score: 3.0	memory length: 313521	epsilon: 0.577226440009178
steps: 266	lr: 0.0001	evaluation reward: 3.2	
episode: 1598	score: 3.0	memory length: 313752	epsilon: 0.5767690600091879
steps: 231	lr: 0.0001	evaluation reward: 3.22	
episode: 1599	score: 6.0	memory length: 314095	epsilon: 0.5760899200092027
steps: 343	lr: 0.0001	evaluation reward: 3.27	
episode: 1600	score: 4.0	memory length: 314390	epsilon: 0.5755058200092154
steps: 295	lr: 0.0001	evaluation reward: 3.29	
episode: 1601	score: 3.0	memory length: 314601	epsilon: 0.5750880400092244
steps: 211	lr: 0.0001	evaluation reward: 3.3	
episode: 1602	score: 5.0	memory length: 314944	epsilon: 0.5744089000092392
steps: 343	lr: 0.0001	evaluation reward: 3.29	
episode: 1603	score: 4.0	memory length: 315218	epsilon: 0.5738663800092509
steps: 274	lr: 0.0001	evaluation reward: 3.31	
episode: 1604	score: 3.0	memory length: 315444	epsilon: 0.5734189000092607
steps: 226	lr: 0.0001	evaluation reward: 3.29	
episode: 1605	score: 3.0	memory length: 315673	epsilon: 0.5729654800092705
steps: 229	lr: 0.0001	evaluation reward: 3.29	
episode: 1606	score: 5.0	memory length: 315981	epsilon: 0.5723556400092837
steps: 308	lr: 0.0001	evaluation reward: 3.28	
episode: 1607	score: 2.0	memory length: 316181	epsilon: 0.5719596400092923
steps: 200	lr: 0.0001	evaluation reward: 3.27	
episode: 1608	score: 0.0	memory length: 316304	epsilon: 0.5717161000092976
steps: 123	lr: 0.0001	evaluation reward: 3.25	
episode: 1609	score: 3.0	memory length: 316532	epsilon: 0.5712646600093074
steps: 228	lr: 0.0001	evaluation reward: 3.28	
episode: 1610	score: 3.0	memory length: 316744	epsilon: 0.5708449000093165
steps: 212	lr: 0.0001	evaluation reward: 3.26	
episode: 1611	score: 8.0	memory length: 317181	epsilon: 0.5699796400093353
steps: 437	lr: 0.0001	evaluation reward: 3.32	
episode: 1612	score: 5.0	memory length: 317491	epsilon: 0.5693658400093486
steps: 310	lr: 0.0001	evaluation reward: 3.34	
episode: 1613	score: 5.0	memory length: 317797	epsilon: 0.5687599600093618
steps: 306	lr: 0.0001	evaluation reward: 3.31	
episode: 1614	score: 4.0	memory length: 318060	epsilon: 0.5682392200093731
steps: 263	lr: 0.0001	evaluation reward: 3.29	
episode: 1615	score: 4.0	memory length: 318336	epsilon: 0.567692740009385
steps: 276	lr: 0.0001	evaluation reward: 3.32	
episode: 1616	score: 6.0	memory length: 318670	epsilon: 0.5670314200093993

steps: 334	lr: 0.0001	evaluation reward: 3.36	
episode: 1617	score: 5.0	memory length: 318996	epsilon: 0.5663859400094133
steps: 326	lr: 0.0001	evaluation reward: 3.39	
episode: 1618	score: 3.0	memory length: 319245	epsilon: 0.565892920009424
steps: 249	lr: 0.0001	evaluation reward: 3.4	
episode: 1619	score: 4.0	memory length: 319518	epsilon: 0.5653523800094358
steps: 273	lr: 0.0001	evaluation reward: 3.4	
episode: 1620	score: 6.0	memory length: 319840	epsilon: 0.5647148200094496
steps: 322	lr: 0.0001	evaluation reward: 3.44	
episode: 1621	score: 5.0	memory length: 320162	epsilon: 0.5640772600094635
steps: 322	lr: 0.0001	evaluation reward: 3.48	
episode: 1622	score: 5.0	memory length: 320458	epsilon: 0.5634911800094762
steps: 296	lr: 0.0001	evaluation reward: 3.44	
episode: 1623	score: 4.0	memory length: 320736	epsilon: 0.5629407400094881
steps: 278	lr: 0.0001	evaluation reward: 3.45	
episode: 1624	score: 2.0	memory length: 320935	epsilon: 0.5625467200094967
steps: 199	lr: 0.0001	evaluation reward: 3.45	
episode: 1625	score: 5.0	memory length: 321226	epsilon: 0.5619705400095092
steps: 291	lr: 0.0001	evaluation reward: 3.47	
episode: 1626	score: 3.0	memory length: 321452	epsilon: 0.5615230600095189
steps: 226	lr: 0.0001	evaluation reward: 3.48	
episode: 1627	score: 5.0	memory length: 321760	epsilon: 0.5609132200095321
steps: 308	lr: 0.0001	evaluation reward: 3.52	
episode: 1628	score: 4.0	memory length: 322034	epsilon: 0.5603707000095439
steps: 274	lr: 0.0001	evaluation reward: 3.53	
episode: 1629	score: 5.0	memory length: 322342	epsilon: 0.5597608600095572
steps: 308	lr: 0.0001	evaluation reward: 3.54	
episode: 1630	score: 2.0	memory length: 322542	epsilon: 0.5593648600095658
steps: 200	lr: 0.0001	evaluation reward: 3.52	
episode: 1631	score: 2.0	memory length: 322740	epsilon: 0.5589728200095743
steps: 198	lr: 0.0001	evaluation reward: 3.49	
episode: 1632	score: 4.0	memory length: 323018	epsilon: 0.5584223800095862
steps: 278	lr: 0.0001	evaluation reward: 3.48	
episode: 1633	score: 3.0	memory length: 323265	epsilon: 0.5579333200095968
steps: 247	lr: 0.0001	evaluation reward: 3.49	
episode: 1634	score: 3.0	memory length: 323491	epsilon: 0.5574858400096065
steps: 226	lr: 0.0001	evaluation reward: 3.49	
episode: 1635	score: 4.0	memory length: 323733	epsilon: 0.557006680009617
steps: 242	lr: 0.0001	evaluation reward: 3.47	
episode: 1636	score: 2.0	memory length: 323931	epsilon: 0.5566146400096255
steps: 198	lr: 0.0001	evaluation reward: 3.48	
episode: 1637	score: 3.0	memory length: 324160	epsilon: 0.5561612200096353
steps: 229	lr: 0.0001	evaluation reward: 3.48	
episode: 1638	score: 5.0	memory length: 324464	epsilon: 0.5555593000096484
steps: 304	lr: 0.0001	evaluation reward: 3.46	
episode: 1639	score: 4.0	memory length: 324721	epsilon: 0.5550504400096594
steps: 257	lr: 0.0001	evaluation reward: 3.47	
episode: 1640	score: 3.0	memory length: 324948	epsilon: 0.5546009800096692

steps: 227	lr: 0.0001	evaluation reward: 3.45	
episode: 1641	score: 3.0	memory length: 325157	epsilon: 0.5541871600096782
steps: 209	lr: 0.0001	evaluation reward: 3.46	
episode: 1642	score: 3.0	memory length: 325386	epsilon: 0.553733740009688
steps: 229	lr: 0.0001	evaluation reward: 3.48	
episode: 1643	score: 6.0	memory length: 325738	epsilon: 0.5530367800097031
steps: 352	lr: 0.0001	evaluation reward: 3.48	
episode: 1644	score: 3.0	memory length: 325985	epsilon: 0.5525477200097137
steps: 247	lr: 0.0001	evaluation reward: 3.48	
episode: 1645	score: 7.0	memory length: 326374	epsilon: 0.5517775000097305
steps: 389	lr: 0.0001	evaluation reward: 3.51	
episode: 1646	score: 0.0	memory length: 326497	epsilon: 0.5515339600097358
steps: 123	lr: 0.0001	evaluation reward: 3.48	
episode: 1647	score: 3.0	memory length: 326724	epsilon: 0.5510845000097455
steps: 227	lr: 0.0001	evaluation reward: 3.48	
episode: 1648	score: 2.0	memory length: 326924	epsilon: 0.5506885000097541
steps: 200	lr: 0.0001	evaluation reward: 3.47	
episode: 1649	score: 2.0	memory length: 327106	epsilon: 0.5503281400097619
steps: 182	lr: 0.0001	evaluation reward: 3.46	
episode: 1650	score: 4.0	memory length: 327397	epsilon: 0.5497519600097744
steps: 291	lr: 0.0001	evaluation reward: 3.47	
episode: 1651	score: 6.0	memory length: 327731	epsilon: 0.5490906400097888
steps: 334	lr: 0.0001	evaluation reward: 3.49	
episode: 1652	score: 4.0	memory length: 327989	epsilon: 0.5485798000097999
steps: 258	lr: 0.0001	evaluation reward: 3.49	
episode: 1653	score: 5.0	memory length: 328292	epsilon: 0.5479798600098129
steps: 303	lr: 0.0001	evaluation reward: 3.49	
episode: 1654	score: 2.0	memory length: 328492	epsilon: 0.5475838600098215
steps: 200	lr: 0.0001	evaluation reward: 3.5	
episode: 1655	score: 5.0	memory length: 328800	epsilon: 0.5469740200098347
steps: 308	lr: 0.0001	evaluation reward: 3.53	
episode: 1656	score: 1.0	memory length: 328952	epsilon: 0.5466730600098413
steps: 152	lr: 0.0001	evaluation reward: 3.5	
episode: 1657	score: 3.0	memory length: 329180	epsilon: 0.5462216200098511
steps: 228	lr: 0.0001	evaluation reward: 3.48	
episode: 1658	score: 2.0	memory length: 329378	epsilon: 0.5458295800098596
steps: 198	lr: 0.0001	evaluation reward: 3.49	
episode: 1659	score: 2.0	memory length: 329596	epsilon: 0.545397940009869
steps: 218	lr: 0.0001	evaluation reward: 3.47	
episode: 1660	score: 6.0	memory length: 329979	epsilon: 0.5446396000098854
steps: 383	lr: 0.0001	evaluation reward: 3.46	
episode: 1661	score: 4.0	memory length: 330225	epsilon: 0.544152520009896
steps: 246	lr: 0.0001	evaluation reward: 3.48	
episode: 1662	score: 2.0	memory length: 330406	epsilon: 0.5437941400099038
steps: 181	lr: 0.0001	evaluation reward: 3.47	
episode: 1663	score: 2.0	memory length: 330605	epsilon: 0.5434001200099123
steps: 199	lr: 0.0001	evaluation reward: 3.42	
episode: 1664	score: 3.0	memory length: 330831	epsilon: 0.542952640009922

steps: 226	lr: 0.0001	evaluation reward: 3.44	
episode: 1665	score: 0.0	memory length: 330954	epsilon: 0.5427091000099273
steps: 123	lr: 0.0001	evaluation reward: 3.42	
episode: 1666	score: 6.0	memory length: 331298	epsilon: 0.5420279800099421
steps: 344	lr: 0.0001	evaluation reward: 3.48	
episode: 1667	score: 1.0	memory length: 331449	epsilon: 0.5417290000099486
steps: 151	lr: 0.0001	evaluation reward: 3.48	
episode: 1668	score: 4.0	memory length: 331727	epsilon: 0.5411785600099606
steps: 278	lr: 0.0001	evaluation reward: 3.5	
episode: 1669	score: 5.0	memory length: 332034	epsilon: 0.5405707000099738
steps: 307	lr: 0.0001	evaluation reward: 3.52	
episode: 1670	score: 4.0	memory length: 332296	epsilon: 0.540051940009985
steps: 262	lr: 0.0001	evaluation reward: 3.49	
episode: 1671	score: 4.0	memory length: 332536	epsilon: 0.5395767400099953
steps: 240	lr: 0.0001	evaluation reward: 3.51	
episode: 1672	score: 5.0	memory length: 332845	epsilon: 0.5389649200100086
steps: 309	lr: 0.0001	evaluation reward: 3.53	
episode: 1673	score: 5.0	memory length: 333153	epsilon: 0.5383550800100219
steps: 308	lr: 0.0001	evaluation reward: 3.57	
episode: 1674	score: 5.0	memory length: 333459	epsilon: 0.537749200010035
steps: 306	lr: 0.0001	evaluation reward: 3.6	
episode: 1675	score: 3.0	memory length: 333684	epsilon: 0.5373037000100447
steps: 225	lr: 0.0001	evaluation reward: 3.57	
episode: 1676	score: 4.0	memory length: 333964	epsilon: 0.5367493000100567
steps: 280	lr: 0.0001	evaluation reward: 3.57	
episode: 1677	score: 2.0	memory length: 334163	epsilon: 0.5363552800100653
steps: 199	lr: 0.0001	evaluation reward: 3.56	
episode: 1678	score: 2.0	memory length: 334363	epsilon: 0.5359592800100739
steps: 200	lr: 0.0001	evaluation reward: 3.58	
episode: 1679	score: 4.0	memory length: 334623	epsilon: 0.535444480010085
steps: 260	lr: 0.0001	evaluation reward: 3.61	
episode: 1680	score: 4.0	memory length: 334883	epsilon: 0.5349296800100962
steps: 260	lr: 0.0001	evaluation reward: 3.62	
episode: 1681	score: 4.0	memory length: 335149	epsilon: 0.5344030000101077
steps: 266	lr: 0.0001	evaluation reward: 3.59	
episode: 1682	score: 5.0	memory length: 335458	epsilon: 0.5337911800101209
steps: 309	lr: 0.0001	evaluation reward: 3.6	
episode: 1683	score: 3.0	memory length: 335703	epsilon: 0.5333060800101315
steps: 245	lr: 0.0001	evaluation reward: 3.62	
episode: 1684	score: 3.0	memory length: 335916	epsilon: 0.5328843400101406
steps: 213	lr: 0.0001	evaluation reward: 3.63	
episode: 1685	score: 5.0	memory length: 336228	epsilon: 0.532266580010154
steps: 312	lr: 0.0001	evaluation reward: 3.63	
episode: 1686	score: 5.0	memory length: 336515	epsilon: 0.5316983200101664
steps: 287	lr: 0.0001	evaluation reward: 3.65	
episode: 1687	score: 1.0	memory length: 336666	epsilon: 0.5313993400101729
steps: 151	lr: 0.0001	evaluation reward: 3.65	
episode: 1688	score: 9.0	memory length: 337008	epsilon: 0.5307221800101876

steps: 342	lr: 0.0001	evaluation reward: 3.69	
episode: 1689	score: 2.0	memory length: 337206	epsilon: 0.5303301400101961
steps: 198	lr: 0.0001	evaluation reward: 3.69	
episode: 1690	score: 4.0	memory length: 337468	epsilon: 0.5298113800102073
steps: 262	lr: 0.0001	evaluation reward: 3.69	
episode: 1691	score: 4.0	memory length: 337761	epsilon: 0.5292312400102199
steps: 293	lr: 0.0001	evaluation reward: 3.7	
episode: 1692	score: 1.0	memory length: 337932	epsilon: 0.5288926600102273
steps: 171	lr: 0.0001	evaluation reward: 3.67	
episode: 1693	score: 3.0	memory length: 338142	epsilon: 0.5284768600102363
steps: 210	lr: 0.0001	evaluation reward: 3.66	
episode: 1694	score: 4.0	memory length: 338402	epsilon: 0.5279620600102475
steps: 260	lr: 0.0001	evaluation reward: 3.69	
episode: 1695	score: 1.0	memory length: 338553	epsilon: 0.527663080010254
steps: 151	lr: 0.0001	evaluation reward: 3.69	
episode: 1696	score: 1.0	memory length: 338721	epsilon: 0.5273304400102612
steps: 168	lr: 0.0001	evaluation reward: 3.63	
episode: 1697	score: 1.0	memory length: 338890	epsilon: 0.5269958200102685
steps: 169	lr: 0.0001	evaluation reward: 3.61	
episode: 1698	score: 4.0	memory length: 339165	epsilon: 0.5264513200102803
steps: 275	lr: 0.0001	evaluation reward: 3.62	
episode: 1699	score: 5.0	memory length: 339490	epsilon: 0.5258078200102942
steps: 325	lr: 0.0001	evaluation reward: 3.61	
episode: 1700	score: 4.0	memory length: 339749	epsilon: 0.5252950000103054
steps: 259	lr: 0.0001	evaluation reward: 3.61	
episode: 1701	score: 4.0	memory length: 340025	epsilon: 0.5247485200103172
steps: 276	lr: 0.0001	evaluation reward: 3.62	
episode: 1702	score: 2.0	memory length: 340223	epsilon: 0.5243564800103258
steps: 198	lr: 0.0001	evaluation reward: 3.59	
episode: 1703	score: 2.0	memory length: 340404	epsilon: 0.5239981000103335
steps: 181	lr: 0.0001	evaluation reward: 3.57	
episode: 1704	score: 4.0	memory length: 340663	epsilon: 0.5234852800103447
steps: 259	lr: 0.0001	evaluation reward: 3.58	
episode: 1705	score: 3.0	memory length: 340894	epsilon: 0.5230279000103546
steps: 231	lr: 0.0001	evaluation reward: 3.58	
episode: 1706	score: 6.0	memory length: 341269	epsilon: 0.5222854000103707
steps: 375	lr: 0.0001	evaluation reward: 3.59	
episode: 1707	score: 2.0	memory length: 341485	epsilon: 0.52185772001038
steps: 216	lr: 0.0001	evaluation reward: 3.59	
episode: 1708	score: 3.0	memory length: 341713	epsilon: 0.5214062800103898
steps: 228	lr: 0.0001	evaluation reward: 3.62	
episode: 1709	score: 6.0	memory length: 342071	epsilon: 0.5206974400104052
steps: 358	lr: 0.0001	evaluation reward: 3.65	
episode: 1710	score: 0.0	memory length: 342193	epsilon: 0.5204558800104104
steps: 122	lr: 0.0001	evaluation reward: 3.62	
episode: 1711	score: 3.0	memory length: 342406	epsilon: 0.5200341400104196
steps: 213	lr: 0.0001	evaluation reward: 3.57	
episode: 1712	score: 5.0	memory length: 342730	epsilon: 0.5193926200104335

steps: 324	lr: 0.0001	evaluation reward: 3.57	
episode: 1713	score: 4.0	memory length: 342987	epsilon: 0.5188837600104446
steps: 257	lr: 0.0001	evaluation reward: 3.56	
episode: 1714	score: 2.0	memory length: 343184	epsilon: 0.518493700010453
steps: 197	lr: 0.0001	evaluation reward: 3.54	
episode: 1715	score: 4.0	memory length: 343480	epsilon: 0.5179076200104658
steps: 296	lr: 0.0001	evaluation reward: 3.54	
episode: 1716	score: 4.0	memory length: 343756	epsilon: 0.5173611400104776
steps: 276	lr: 0.0001	evaluation reward: 3.52	
episode: 1717	score: 4.0	memory length: 344018	epsilon: 0.5168423800104889
steps: 262	lr: 0.0001	evaluation reward: 3.51	
episode: 1718	score: 4.0	memory length: 344277	epsilon: 0.5163295600105
steps: 259	lr: 0.0001	evaluation reward: 3.52	
episode: 1719	score: 7.0	memory length: 344713	epsilon: 0.5154662800105188
steps: 436	lr: 0.0001	evaluation reward: 3.55	
episode: 1720	score: 5.0	memory length: 345051	epsilon: 0.5147970400105333
steps: 338	lr: 0.0001	evaluation reward: 3.54	
episode: 1721	score: 2.0	memory length: 345271	epsilon: 0.5143614400105427
steps: 220	lr: 0.0001	evaluation reward: 3.51	
episode: 1722	score: 6.0	memory length: 345607	epsilon: 0.5136961600105572
steps: 336	lr: 0.0001	evaluation reward: 3.52	
episode: 1723	score: 2.0	memory length: 345788	epsilon: 0.513337780010565
steps: 181	lr: 0.0001	evaluation reward: 3.5	
episode: 1724	score: 0.0	memory length: 345911	epsilon: 0.5130942400105702
steps: 123	lr: 0.0001	evaluation reward: 3.48	
episode: 1725	score: 2.0	memory length: 346091	epsilon: 0.512737840010578
steps: 180	lr: 0.0001	evaluation reward: 3.45	
episode: 1726	score: 4.0	memory length: 346366	epsilon: 0.5121933400105898
steps: 275	lr: 0.0001	evaluation reward: 3.46	
episode: 1727	score: 2.0	memory length: 346566	epsilon: 0.5117973400105984
steps: 200	lr: 0.0001	evaluation reward: 3.43	
episode: 1728	score: 2.0	memory length: 346748	epsilon: 0.5114369800106062
steps: 182	lr: 0.0001	evaluation reward: 3.41	
episode: 1729	score: 2.0	memory length: 346968	epsilon: 0.5110013800106157
steps: 220	lr: 0.0001	evaluation reward: 3.38	
episode: 1730	score: 1.0	memory length: 347119	epsilon: 0.5107024000106222
steps: 151	lr: 0.0001	evaluation reward: 3.37	
episode: 1731	score: 4.0	memory length: 347412	epsilon: 0.5101222600106348
steps: 293	lr: 0.0001	evaluation reward: 3.39	
episode: 1732	score: 4.0	memory length: 347671	epsilon: 0.5096094400106459
steps: 259	lr: 0.0001	evaluation reward: 3.39	
episode: 1733	score: 1.0	memory length: 347822	epsilon: 0.5093104600106524
steps: 151	lr: 0.0001	evaluation reward: 3.37	
episode: 1734	score: 2.0	memory length: 348022	epsilon: 0.508914460010661
steps: 200	lr: 0.0001	evaluation reward: 3.36	
episode: 1735	score: 4.0	memory length: 348281	epsilon: 0.5084016400106721
steps: 259	lr: 0.0001	evaluation reward: 3.36	
episode: 1736	score: 2.0	memory length: 348481	epsilon: 0.5080056400106807

```

steps: 200    lr: 0.0001    evaluation reward: 3.36
episode: 1737  score: 1.0    memory length: 348632    epsilon: 0.5077066600106872
steps: 151    lr: 0.0001    evaluation reward: 3.34
episode: 1738  score: 3.0    memory length: 348880    epsilon: 0.5072156200106979
steps: 248    lr: 0.0001    evaluation reward: 3.32
episode: 1739  score: 3.0    memory length: 349091    epsilon: 0.5067978400107069
steps: 211    lr: 0.0001    evaluation reward: 3.31
episode: 1740  score: 1.0    memory length: 349242    epsilon: 0.5064988600107134
steps: 151    lr: 0.0001    evaluation reward: 3.29
episode: 1741  score: 3.0    memory length: 349470    epsilon: 0.5060474200107232
steps: 228    lr: 0.0001    evaluation reward: 3.29
episode: 1742  score: 4.0    memory length: 349725    epsilon: 0.5055425200107342
steps: 255    lr: 0.0001    evaluation reward: 3.3
episode: 1743  score: 6.0    memory length: 350064    epsilon: 0.5048713000107488
steps: 339    lr: 0.0001    evaluation reward: 3.3
episode: 1744  score: 4.0    memory length: 350323    epsilon: 0.5043584800107599
steps: 259    lr: 0.0001    evaluation reward: 3.31
episode: 1745  score: 4.0    memory length: 350584    epsilon: 0.5038417000107711
steps: 261    lr: 0.0001    evaluation reward: 3.28
episode: 1746  score: 1.0    memory length: 350754    epsilon: 0.5035051000107784
steps: 170    lr: 0.0001    evaluation reward: 3.29
episode: 1747  score: 5.0    memory length: 351063    epsilon: 0.5028932800107917
steps: 309    lr: 0.0001    evaluation reward: 3.31
episode: 1748  score: 3.0    memory length: 351289    epsilon: 0.5024458000108014
steps: 226    lr: 0.0001    evaluation reward: 3.32
episode: 1749  score: 4.0    memory length: 351585    epsilon: 0.5018597200108141
steps: 296    lr: 0.0001    evaluation reward: 3.34
episode: 1750  score: 3.0    memory length: 351796    epsilon: 0.5014419400108232
steps: 211    lr: 0.0001    evaluation reward: 3.33
episode: 1751  score: 3.0    memory length: 352009    epsilon: 0.5010202000108324
steps: 213    lr: 0.0001    evaluation reward: 3.3
episode: 1752  score: 0.0    memory length: 352132    epsilon: 0.5007766600108376
steps: 123    lr: 0.0001    evaluation reward: 3.26
episode: 1753  score: 4.0    memory length: 352409    epsilon: 0.5002282000108496
steps: 277    lr: 0.0001    evaluation reward: 3.25
episode: 1754  score: 3.0    memory length: 352653    epsilon: 0.4997450800108529
steps: 244    lr: 0.0001    evaluation reward: 3.26
episode: 1755  score: 9.0    memory length: 353125    epsilon:
0.49881052001084697    steps: 472    lr: 0.0001    evaluation reward: 3.3
episode: 1756  score: 3.0    memory length: 353338    epsilon: 0.4983887800108443
steps: 213    lr: 0.0001    evaluation reward: 3.32
episode: 1757  score: 2.0    memory length: 353520    epsilon: 0.498028420010842
steps: 182    lr: 0.0001    evaluation reward: 3.31
episode: 1758  score: 3.0    memory length: 353750    epsilon:
0.49757302001083914    steps: 230    lr: 0.0001    evaluation reward: 3.32
episode: 1759  score: 1.0    memory length: 353901    epsilon:
0.49727404001083725    steps: 151    lr: 0.0001    evaluation reward: 3.31
episode: 1760  score: 5.0    memory length: 354228    epsilon:

```

0.49662658001083315 steps: 327 lr: 0.0001 evaluation reward: 3.3
 episode: 1761 score: 4.0 memory length: 354505 epsilon: 0.4960781200108297
 steps: 277 lr: 0.0001 evaluation reward: 3.3
 episode: 1762 score: 1.0 memory length: 354656 epsilon: 0.4957791400108278
 steps: 151 lr: 0.0001 evaluation reward: 3.29
 episode: 1763 score: 2.0 memory length: 354854 epsilon: 0.4953871000108253
 steps: 198 lr: 0.0001 evaluation reward: 3.29
 episode: 1764 score: 2.0 memory length: 355036 epsilon:
 0.49502674001082303 steps: 182 lr: 0.0001 evaluation reward: 3.28
 episode: 1765 score: 5.0 memory length: 355382 epsilon: 0.4943416600108187
 steps: 346 lr: 0.0001 evaluation reward: 3.33
 episode: 1766 score: 3.0 memory length: 355608 epsilon:
 0.49389418001081586 steps: 226 lr: 0.0001 evaluation reward: 3.3
 episode: 1767 score: 1.0 memory length: 355759 epsilon: 0.493595200010814
 steps: 151 lr: 0.0001 evaluation reward: 3.3
 episode: 1768 score: 5.0 memory length: 356087 epsilon:
 0.49294576001080986 steps: 328 lr: 0.0001 evaluation reward: 3.31
 episode: 1769 score: 5.0 memory length: 356396 epsilon: 0.492333940010806
 steps: 309 lr: 0.0001 evaluation reward: 3.31
 episode: 1770 score: 6.0 memory length: 356721 epsilon: 0.4916904400108019
 steps: 325 lr: 0.0001 evaluation reward: 3.33
 episode: 1771 score: 4.0 memory length: 356979 epsilon: 0.4911796000107987
 steps: 258 lr: 0.0001 evaluation reward: 3.33
 episode: 1772 score: 3.0 memory length: 357210 epsilon: 0.4907222200107958
 steps: 231 lr: 0.0001 evaluation reward: 3.31
 episode: 1773 score: 1.0 memory length: 357362 epsilon: 0.4904212600107939
 steps: 152 lr: 0.0001 evaluation reward: 3.27
 episode: 1774 score: 5.0 memory length: 357648 epsilon: 0.4898549800107903
 steps: 286 lr: 0.0001 evaluation reward: 3.27
 episode: 1775 score: 3.0 memory length: 357861 epsilon:
 0.48943324001078764 steps: 213 lr: 0.0001 evaluation reward: 3.27
 episode: 1776 score: 3.0 memory length: 358111 epsilon: 0.4889382400107845
 steps: 250 lr: 0.0001 evaluation reward: 3.26
 episode: 1777 score: 6.0 memory length: 358465 epsilon: 0.4882373200107801
 steps: 354 lr: 0.0001 evaluation reward: 3.3
 episode: 1778 score: 3.0 memory length: 358710 epsilon: 0.487752220010777
 steps: 245 lr: 0.0001 evaluation reward: 3.31
 episode: 1779 score: 3.0 memory length: 358921 epsilon:
 0.48733444001077436 steps: 211 lr: 0.0001 evaluation reward: 3.3
 episode: 1780 score: 4.0 memory length: 359180 epsilon: 0.4868216200107711
 steps: 259 lr: 0.0001 evaluation reward: 3.3
 episode: 1781 score: 3.0 memory length: 359406 epsilon: 0.4863741400107683
 steps: 226 lr: 0.0001 evaluation reward: 3.29
 episode: 1782 score: 5.0 memory length: 359721 epsilon:
 0.48575044001076434 steps: 315 lr: 0.0001 evaluation reward: 3.29
 episode: 1783 score: 3.0 memory length: 359949 epsilon: 0.4852990000107615
 steps: 228 lr: 0.0001 evaluation reward: 3.29
 episode: 1784 score: 2.0 memory length: 360147 epsilon: 0.484906960010759

```

steps: 198    lr: 0.0001    evaluation reward: 3.28
episode: 1785  score: 5.0    memory length: 360490    epsilon: 0.4842278200107547
steps: 343    lr: 0.0001    evaluation reward: 3.28
episode: 1786  score: 3.0    memory length: 360716    epsilon: 0.4837803400107519
steps: 226    lr: 0.0001    evaluation reward: 3.26
episode: 1787  score: 3.0    memory length: 360961    epsilon: 0.4832952400107488
steps: 245    lr: 0.0001    evaluation reward: 3.28
episode: 1788  score: 1.0    memory length: 361112    epsilon: 0.4829962600107469
steps: 151    lr: 0.0001    evaluation reward: 3.2
episode: 1789  score: 3.0    memory length: 361338    epsilon: 0.4825487800107441
steps: 226    lr: 0.0001    evaluation reward: 3.21
episode: 1790  score: 5.0    memory length: 361642    epsilon: 0.4819468600107403
steps: 304    lr: 0.0001    evaluation reward: 3.22
episode: 1791  score: 7.0    memory length: 362030    epsilon: 0.4811786200107354
steps: 388    lr: 0.0001    evaluation reward: 3.25
episode: 1792  score: 5.0    memory length: 362371    epsilon:
0.48050344001073114    steps: 341    lr: 0.0001    evaluation reward: 3.29
episode: 1793  score: 3.0    memory length: 362619    epsilon:
0.48001240001072804    steps: 248    lr: 0.0001    evaluation reward: 3.29
episode: 1794  score: 4.0    memory length: 362897    epsilon:
0.47946196001072455    steps: 278    lr: 0.0001    evaluation reward: 3.29
episode: 1795  score: 3.0    memory length: 363110    epsilon: 0.4790402200107219
steps: 213    lr: 0.0001    evaluation reward: 3.31
episode: 1796  score: 4.0    memory length: 363406    epsilon: 0.4784541400107182
steps: 296    lr: 0.0001    evaluation reward: 3.34
episode: 1797  score: 3.0    memory length: 363634    epsilon: 0.4780027000107153
steps: 228    lr: 0.0001    evaluation reward: 3.36
episode: 1798  score: 4.0    memory length: 363908    epsilon: 0.4774601800107119
steps: 274    lr: 0.0001    evaluation reward: 3.36
episode: 1799  score: 3.0    memory length: 364138    epsilon: 0.477004780010709
steps: 230    lr: 0.0001    evaluation reward: 3.34
episode: 1800  score: 3.0    memory length: 364351    epsilon:
0.47658304001070634    steps: 213    lr: 0.0001    evaluation reward: 3.33
episode: 1801  score: 4.0    memory length: 364627    epsilon: 0.4760365600107029
steps: 276    lr: 0.0001    evaluation reward: 3.33
episode: 1802  score: 3.0    memory length: 364838    epsilon:
0.47561878001070024    steps: 211    lr: 0.0001    evaluation reward: 3.34
episode: 1803  score: 2.0    memory length: 365056    epsilon: 0.4751871400106975
steps: 218    lr: 0.0001    evaluation reward: 3.34
episode: 1804  score: 3.0    memory length: 365269    epsilon:
0.47476540001069484    steps: 213    lr: 0.0001    evaluation reward: 3.33
episode: 1805  score: 3.0    memory length: 365495    epsilon: 0.474317920010692
steps: 226    lr: 0.0001    evaluation reward: 3.33
episode: 1806  score: 4.0    memory length: 365754    epsilon:
0.47380510001068876    steps: 259    lr: 0.0001    evaluation reward: 3.31
episode: 1807  score: 3.0    memory length: 366003    epsilon:
0.47331208001068564    steps: 249    lr: 0.0001    evaluation reward: 3.32
episode: 1808  score: 2.0    memory length: 366184    epsilon: 0.4729537000106834

```



```

steps: 181    lr: 0.0001    evaluation reward: 3.31
episode: 1809  score: 3.0    memory length: 366410    epsilon:
0.47250622001068054    steps: 226    lr: 0.0001    evaluation reward: 3.28
episode: 1810  score: 3.0    memory length: 366641    epsilon:
0.47204884001067765    steps: 231    lr: 0.0001    evaluation reward: 3.31
episode: 1811  score: 3.0    memory length: 366870    epsilon: 0.4715954200106748
steps: 229    lr: 0.0001    evaluation reward: 3.31
episode: 1812  score: 2.0    memory length: 367051    epsilon: 0.4712370400106725
steps: 181    lr: 0.0001    evaluation reward: 3.28
episode: 1813  score: 2.0    memory length: 367232    epsilon:
0.47087866001067025    steps: 181    lr: 0.0001    evaluation reward: 3.26
episode: 1814  score: 3.0    memory length: 367460    epsilon: 0.4704272200106674
steps: 228    lr: 0.0001    evaluation reward: 3.27
episode: 1815  score: 4.0    memory length: 367736    epsilon:
0.46988074001066393    steps: 276    lr: 0.0001    evaluation reward: 3.27
episode: 1816  score: 4.0    memory length: 368030    epsilon:
0.46929862001066025    steps: 294    lr: 0.0001    evaluation reward: 3.27
episode: 1817  score: 4.0    memory length: 368309    epsilon:
0.46874620001065676    steps: 279    lr: 0.0001    evaluation reward: 3.27
episode: 1818  score: 1.0    memory length: 368459    epsilon: 0.4684492000106549
steps: 150    lr: 0.0001    evaluation reward: 3.24
episode: 1819  score: 3.0    memory length: 368687    epsilon: 0.467997760010652
steps: 228    lr: 0.0001    evaluation reward: 3.2
episode: 1820  score: 5.0    memory length: 369017    epsilon: 0.4673443600106479
steps: 330    lr: 0.0001    evaluation reward: 3.2
episode: 1821  score: 4.0    memory length: 369293    epsilon:
0.46679788001064443    steps: 276    lr: 0.0001    evaluation reward: 3.22
episode: 1822  score: 1.0    memory length: 369444    epsilon:
0.46649890001064254    steps: 151    lr: 0.0001    evaluation reward: 3.17
episode: 1823  score: 3.0    memory length: 369657    epsilon:
0.46607716001063987    steps: 213    lr: 0.0001    evaluation reward: 3.18
episode: 1824  score: 5.0    memory length: 369943    epsilon: 0.4655108800106363
steps: 286    lr: 0.0001    evaluation reward: 3.23
episode: 1825  score: 1.0    memory length: 370094    epsilon: 0.4652119000106344
steps: 151    lr: 0.0001    evaluation reward: 3.22
episode: 1826  score: 2.0    memory length: 370275    epsilon: 0.4648535200106321
steps: 181    lr: 0.0001    evaluation reward: 3.2
episode: 1827  score: 5.0    memory length: 370600    epsilon:
0.46421002001062805    steps: 325    lr: 0.0001    evaluation reward: 3.23
episode: 1828  score: 3.0    memory length: 370811    epsilon: 0.4637922400106254
steps: 211    lr: 0.0001    evaluation reward: 3.24
episode: 1829  score: 6.0    memory length: 371176    epsilon:
0.46306954001062084    steps: 365    lr: 0.0001    evaluation reward: 3.28
episode: 1830  score: 3.0    memory length: 371402    epsilon: 0.462622060010618
steps: 226    lr: 0.0001    evaluation reward: 3.3
episode: 1831  score: 4.0    memory length: 371696    epsilon: 0.4620399400106143
steps: 294    lr: 0.0001    evaluation reward: 3.3
episode: 1832  score: 8.0    memory length: 372127    epsilon: 0.4611865600106089

```

```

steps: 431    lr: 0.0001    evaluation reward: 3.34
episode: 1833  score: 5.0    memory length: 372449    epsilon: 0.4605490000106049
steps: 322    lr: 0.0001    evaluation reward: 3.38
episode: 1834  score: 5.0    memory length: 372744    epsilon: 0.4599649000106012
steps: 295    lr: 0.0001    evaluation reward: 3.41
episode: 1835  score: 3.0    memory length: 372990    epsilon: 0.4594778200105981
steps: 246    lr: 0.0001    evaluation reward: 3.4
episode: 1836  score: 3.0    memory length: 373216    epsilon: 0.4590303400105953
steps: 226    lr: 0.0001    evaluation reward: 3.41
episode: 1837  score: 3.0    memory length: 373445    epsilon: 0.4585769200105924
steps: 229    lr: 0.0001    evaluation reward: 3.43
episode: 1838  score: 3.0    memory length: 373690    epsilon:
0.45809182001058935    steps: 245    lr: 0.0001    evaluation reward: 3.43
episode: 1839  score: 3.0    memory length: 373903    epsilon: 0.4576700800105867
steps: 213    lr: 0.0001    evaluation reward: 3.43
episode: 1840  score: 3.0    memory length: 374131    epsilon: 0.4572186400105838
steps: 228    lr: 0.0001    evaluation reward: 3.45
episode: 1841  score: 3.0    memory length: 374360    epsilon:
0.45676522001058095    steps: 229    lr: 0.0001    evaluation reward: 3.45
episode: 1842  score: 3.0    memory length: 374573    epsilon: 0.4563434800105783
steps: 213    lr: 0.0001    evaluation reward: 3.44
episode: 1843  score: 6.0    memory length: 374907    epsilon: 0.4556821600105741
steps: 334    lr: 0.0001    evaluation reward: 3.44
episode: 1844  score: 3.0    memory length: 375134    epsilon:
0.45523270001057126    steps: 227    lr: 0.0001    evaluation reward: 3.43
episode: 1845  score: 3.0    memory length: 375343    epsilon:
0.45481888001056864    steps: 209    lr: 0.0001    evaluation reward: 3.42
episode: 1846  score: 3.0    memory length: 375572    epsilon:
0.45436546001056577    steps: 229    lr: 0.0001    evaluation reward: 3.44
episode: 1847  score: 2.0    memory length: 375770    epsilon: 0.4539734200105633
steps: 198    lr: 0.0001    evaluation reward: 3.41
episode: 1848  score: 3.0    memory length: 376037    epsilon:
0.45344476001055994    steps: 267    lr: 0.0001    evaluation reward: 3.41
episode: 1849  score: 5.0    memory length: 376347    epsilon:
0.45283096001055606    steps: 310    lr: 0.0001    evaluation reward: 3.42
episode: 1850  score: 3.0    memory length: 376560    epsilon: 0.4524092200105534
steps: 213    lr: 0.0001    evaluation reward: 3.42
episode: 1851  score: 5.0    memory length: 376892    epsilon:
0.45175186001054923    steps: 332    lr: 0.0001    evaluation reward: 3.44
episode: 1852  score: 3.0    memory length: 377119    epsilon: 0.4513024000105464
steps: 227    lr: 0.0001    evaluation reward: 3.47
episode: 1853  score: 3.0    memory length: 377346    epsilon:
0.45085294001054355    steps: 227    lr: 0.0001    evaluation reward: 3.46
episode: 1854  score: 4.0    memory length: 377643    epsilon: 0.4502648800105398
steps: 297    lr: 0.0001    evaluation reward: 3.47
episode: 1855  score: 4.0    memory length: 377919    epsilon:
0.44971840001053637    steps: 276    lr: 0.0001    evaluation reward: 3.42
episode: 1856  score: 7.0    memory length: 378283    epsilon: 0.4489976800105318

```

```

steps: 364    lr: 0.0001    evaluation reward: 3.46
episode: 1857  score: 5.0    memory length: 378554    epsilon: 0.4484611000105284
steps: 271    lr: 0.0001    evaluation reward: 3.49
episode: 1858  score: 3.0    memory length: 378782    epsilon:
0.44800966001052556    steps: 228    lr: 0.0001    evaluation reward: 3.49
episode: 1859  score: 3.0    memory length: 379011    epsilon: 0.4475562400105227
steps: 229    lr: 0.0001    evaluation reward: 3.51
episode: 1860  score: 1.0    memory length: 379163    epsilon: 0.4472552800105208
steps: 152    lr: 0.0001    evaluation reward: 3.47
episode: 1861  score: 3.0    memory length: 379393    epsilon: 0.4467998800105179
steps: 230    lr: 0.0001    evaluation reward: 3.46
episode: 1862  score: 3.0    memory length: 379606    epsilon:
0.44637814001051523    steps: 213    lr: 0.0001    evaluation reward: 3.48
episode: 1863  score: 3.0    memory length: 379817    epsilon: 0.4459603600105126
steps: 211    lr: 0.0001    evaluation reward: 3.49
episode: 1864  score: 5.0    memory length: 380138    epsilon:
0.44532478001050857    steps: 321    lr: 0.0001    evaluation reward: 3.52
episode: 1865  score: 3.0    memory length: 380348    epsilon:
0.44490898001050594    steps: 210    lr: 0.0001    evaluation reward: 3.5
episode: 1866  score: 3.0    memory length: 380561    epsilon:
0.44448724001050327    steps: 213    lr: 0.0001    evaluation reward: 3.5
episode: 1867  score: 4.0    memory length: 380837    epsilon: 0.4439407600104998
steps: 276    lr: 0.0001    evaluation reward: 3.53
episode: 1868  score: 3.0    memory length: 381063    epsilon: 0.443493280010497
steps: 226    lr: 0.0001    evaluation reward: 3.51
episode: 1869  score: 5.0    memory length: 381373    epsilon: 0.4428794800104931
steps: 310    lr: 0.0001    evaluation reward: 3.51
episode: 1870  score: 2.0    memory length: 381571    epsilon: 0.4424874400104906
steps: 198    lr: 0.0001    evaluation reward: 3.47
episode: 1871  score: 5.0    memory length: 381914    epsilon: 0.4418083000104863
steps: 343    lr: 0.0001    evaluation reward: 3.48
episode: 1872  score: 4.0    memory length: 382191    epsilon:
0.44125984001048285    steps: 277    lr: 0.0001    evaluation reward: 3.49
episode: 1873  score: 6.0    memory length: 382563    epsilon: 0.4405232800104782
steps: 372    lr: 0.0001    evaluation reward: 3.54
episode: 1874  score: 4.0    memory length: 382829    epsilon:
0.43999660001047486    steps: 266    lr: 0.0001    evaluation reward: 3.53
episode: 1875  score: 3.0    memory length: 383040    epsilon: 0.4395788200104722
steps: 211    lr: 0.0001    evaluation reward: 3.53
episode: 1876  score: 3.0    memory length: 383266    epsilon: 0.4391313400104694
steps: 226    lr: 0.0001    evaluation reward: 3.53
episode: 1877  score: 2.0    memory length: 383447    epsilon: 0.4387729600104671
steps: 181    lr: 0.0001    evaluation reward: 3.49
episode: 1878  score: 4.0    memory length: 383705    epsilon: 0.4382621200104639
steps: 258    lr: 0.0001    evaluation reward: 3.5
episode: 1879  score: 2.0    memory length: 383886    epsilon: 0.4379037400104616
steps: 181    lr: 0.0001    evaluation reward: 3.49
episode: 1880  score: 3.0    memory length: 384133    epsilon: 0.4374146800104585

```

```

steps: 247    lr: 0.0001    evaluation reward: 3.48
episode: 1881  score: 3.0    memory length: 384363    epsilon:
0.43695928001045564    steps: 230    lr: 0.0001    evaluation reward: 3.48
episode: 1882  score: 3.0    memory length: 384591    epsilon: 0.4365078400104528
steps: 228    lr: 0.0001    evaluation reward: 3.46
episode: 1883  score: 3.0    memory length: 384822    epsilon: 0.4360504600104499
steps: 231    lr: 0.0001    evaluation reward: 3.46
episode: 1884  score: 5.0    memory length: 385113    epsilon:
0.43547428001044625    steps: 291    lr: 0.0001    evaluation reward: 3.49
episode: 1885  score: 2.0    memory length: 385293    epsilon: 0.435117880010444
steps: 180    lr: 0.0001    evaluation reward: 3.46
episode: 1886  score: 3.0    memory length: 385540    epsilon: 0.4346288200104409
steps: 247    lr: 0.0001    evaluation reward: 3.46
episode: 1887  score: 3.0    memory length: 385750    epsilon:
0.43421302001043827    steps: 210    lr: 0.0001    evaluation reward: 3.46
episode: 1888  score: 8.0    memory length: 386190    epsilon:
0.43334182001043275    steps: 440    lr: 0.0001    evaluation reward: 3.53
episode: 1889  score: 6.0    memory length: 386545    epsilon: 0.4326389200104283
steps: 355    lr: 0.0001    evaluation reward: 3.56
episode: 1890  score: 3.0    memory length: 386773    epsilon:
0.43218748001042545    steps: 228    lr: 0.0001    evaluation reward: 3.54
episode: 1891  score: 3.0    memory length: 386984    epsilon: 0.4317697000104228
steps: 211    lr: 0.0001    evaluation reward: 3.5
episode: 1892  score: 5.0    memory length: 387272    epsilon: 0.4311994600104192
steps: 288    lr: 0.0001    evaluation reward: 3.5
episode: 1893  score: 3.0    memory length: 387517    epsilon:
0.43071436001041613    steps: 245    lr: 0.0001    evaluation reward: 3.5
episode: 1894  score: 2.0    memory length: 387699    epsilon:
0.43035400001041385    steps: 182    lr: 0.0001    evaluation reward: 3.48
episode: 1895  score: 1.0    memory length: 387868    epsilon:
0.43001938001041173    steps: 169    lr: 0.0001    evaluation reward: 3.46
episode: 1896  score: 2.0    memory length: 388050    epsilon:
0.42965902001040945    steps: 182    lr: 0.0001    evaluation reward: 3.44
episode: 1897  score: 5.0    memory length: 388377    epsilon:
0.42901156001040536    steps: 327    lr: 0.0001    evaluation reward: 3.46
episode: 1898  score: 3.0    memory length: 388587    epsilon: 0.4285957600104027
steps: 210    lr: 0.0001    evaluation reward: 3.45
episode: 1899  score: 5.0    memory length: 388914    epsilon:
0.42794830001039863    steps: 327    lr: 0.0001    evaluation reward: 3.47
episode: 1900  score: 4.0    memory length: 389173    epsilon: 0.4274354800103954
steps: 259    lr: 0.0001    evaluation reward: 3.48
episode: 1901  score: 3.0    memory length: 389403    epsilon: 0.4269800800103925
steps: 230    lr: 0.0001    evaluation reward: 3.47
episode: 1902  score: 3.0    memory length: 389614    epsilon:
0.42656230001038986    steps: 211    lr: 0.0001    evaluation reward: 3.47
episode: 1903  score: 6.0    memory length: 389969    epsilon: 0.4258594000103854
steps: 355    lr: 0.0001    evaluation reward: 3.51
episode: 1904  score: 3.0    memory length: 390197    epsilon:

```

0.42540796001038256 steps: 228 lr: 0.0001 evaluation reward: 3.51
 episode: 1905 score: 4.0 memory length: 390456 epsilon: 0.4248951400103793
 steps: 259 lr: 0.0001 evaluation reward: 3.52
 episode: 1906 score: 2.0 memory length: 390638 epsilon:
 0.42453478001037703 steps: 182 lr: 0.0001 evaluation reward: 3.5
 episode: 1907 score: 3.0 memory length: 390851 epsilon:
 0.42411304001037436 steps: 213 lr: 0.0001 evaluation reward: 3.5
 episode: 1908 score: 3.0 memory length: 391060 epsilon:
 0.42369922001037175 steps: 209 lr: 0.0001 evaluation reward: 3.51
 episode: 1909 score: 4.0 memory length: 391337 epsilon: 0.4231507600103683
 steps: 277 lr: 0.0001 evaluation reward: 3.52
 episode: 1910 score: 4.0 memory length: 391615 epsilon: 0.4226003200103648
 steps: 278 lr: 0.0001 evaluation reward: 3.53
 episode: 1911 score: 3.0 memory length: 391828 epsilon: 0.4221785800103621
 steps: 213 lr: 0.0001 evaluation reward: 3.53
 episode: 1912 score: 4.0 memory length: 392088 epsilon:
 0.42166378001035887 steps: 260 lr: 0.0001 evaluation reward: 3.55
 episode: 1913 score: 5.0 memory length: 392435 epsilon: 0.4209767200103545
 steps: 347 lr: 0.0001 evaluation reward: 3.58
 episode: 1914 score: 3.0 memory length: 392661 epsilon: 0.4205292400103517
 steps: 226 lr: 0.0001 evaluation reward: 3.58
 episode: 1915 score: 3.0 memory length: 392873 epsilon:
 0.42010948001034903 steps: 212 lr: 0.0001 evaluation reward: 3.57
 episode: 1916 score: 3.0 memory length: 393120 epsilon:
 0.41962042001034594 steps: 247 lr: 0.0001 evaluation reward: 3.56
 episode: 1917 score: 3.0 memory length: 393330 epsilon: 0.4192046200103433
 steps: 210 lr: 0.0001 evaluation reward: 3.55
 episode: 1918 score: 4.0 memory length: 393571 epsilon: 0.4187274400103403
 steps: 241 lr: 0.0001 evaluation reward: 3.58
 episode: 1919 score: 5.0 memory length: 393853 epsilon:
 0.41816908001033676 steps: 282 lr: 0.0001 evaluation reward: 3.6
 episode: 1920 score: 4.0 memory length: 394112 epsilon: 0.4176562600103335
 steps: 259 lr: 0.0001 evaluation reward: 3.59
 episode: 1921 score: 3.0 memory length: 394325 epsilon:
 0.41723452001033084 steps: 213 lr: 0.0001 evaluation reward: 3.58
 episode: 1922 score: 4.0 memory length: 394618 epsilon: 0.4166543800103272
 steps: 293 lr: 0.0001 evaluation reward: 3.61
 episode: 1923 score: 5.0 memory length: 394925 epsilon:
 0.41604652001032333 steps: 307 lr: 0.0001 evaluation reward: 3.63
 episode: 1924 score: 4.0 memory length: 395220 epsilon:
 0.41546242001031963 steps: 295 lr: 0.0001 evaluation reward: 3.62
 episode: 1925 score: 5.0 memory length: 395511 epsilon: 0.414886240010316
 steps: 291 lr: 0.0001 evaluation reward: 3.66
 episode: 1926 score: 4.0 memory length: 395787 epsilon:
 0.41433976001031253 steps: 276 lr: 0.0001 evaluation reward: 3.68
 episode: 1927 score: 3.0 memory length: 395997 epsilon: 0.4139239600103099
 steps: 210 lr: 0.0001 evaluation reward: 3.66
 episode: 1928 score: 4.0 memory length: 396255 epsilon:

0.41341312001030667 steps: 258 lr: 0.0001 evaluation reward: 3.67
 episode: 1929 score: 5.0 memory length: 396600 epsilon:
 0.41273002001030235 steps: 345 lr: 0.0001 evaluation reward: 3.66
 episode: 1930 score: 4.0 memory length: 396857 epsilon: 0.4122211600102991
 steps: 257 lr: 0.0001 evaluation reward: 3.67
 episode: 1931 score: 6.0 memory length: 397196 epsilon: 0.4115499400102949
 steps: 339 lr: 0.0001 evaluation reward: 3.69
 episode: 1932 score: 2.0 memory length: 397378 epsilon: 0.4111895800102926
 steps: 182 lr: 0.0001 evaluation reward: 3.63
 episode: 1933 score: 3.0 memory length: 397625 epsilon: 0.4107005200102895
 steps: 247 lr: 0.0001 evaluation reward: 3.61
 episode: 1934 score: 3.0 memory length: 397838 epsilon:
 0.41027878001028684 steps: 213 lr: 0.0001 evaluation reward: 3.59
 episode: 1935 score: 4.0 memory length: 398115 epsilon:
 0.40973032001028337 steps: 277 lr: 0.0001 evaluation reward: 3.6
 episode: 1936 score: 3.0 memory length: 398348 epsilon:
 0.40926898001028045 steps: 233 lr: 0.0001 evaluation reward: 3.6
 episode: 1937 score: 3.0 memory length: 398577 epsilon: 0.4088155600102776
 steps: 229 lr: 0.0001 evaluation reward: 3.6
 episode: 1938 score: 4.0 memory length: 398837 epsilon: 0.4083007600102743
 steps: 260 lr: 0.0001 evaluation reward: 3.61
 episode: 1939 score: 7.0 memory length: 399226 epsilon:
 0.40753054001026945 steps: 389 lr: 0.0001 evaluation reward: 3.65
 episode: 1940 score: 3.0 memory length: 399436 epsilon: 0.4071147400102668
 steps: 210 lr: 0.0001 evaluation reward: 3.65
 episode: 1941 score: 3.0 memory length: 399664 epsilon:
 0.40666330001026396 steps: 228 lr: 0.0001 evaluation reward: 3.65
 episode: 1942 score: 5.0 memory length: 399961 epsilon:
 0.40607524001026024 steps: 297 lr: 0.0001 evaluation reward: 3.67
 episode: 1943 score: 4.0 memory length: 400238 epsilon:
 0.40552678001025677 steps: 277 lr: 0.0001 evaluation reward: 3.65
 episode: 1944 score: 2.0 memory length: 400436 epsilon: 0.4051347400102543
 steps: 198 lr: 0.0001 evaluation reward: 3.64
 episode: 1945 score: 5.0 memory length: 400731 epsilon: 0.4045506400102506
 steps: 295 lr: 0.0001 evaluation reward: 3.66
 episode: 1946 score: 3.0 memory length: 400957 epsilon:
 0.40410316001024776 steps: 226 lr: 0.0001 evaluation reward: 3.66
 episode: 1947 score: 6.0 memory length: 401310 epsilon:
 0.40340422001024334 steps: 353 lr: 0.0001 evaluation reward: 3.7
 episode: 1948 score: 2.0 memory length: 401492 epsilon:
 0.40304386001024106 steps: 182 lr: 0.0001 evaluation reward: 3.69
 episode: 1949 score: 2.0 memory length: 401673 epsilon: 0.4026854800102388
 steps: 181 lr: 0.0001 evaluation reward: 3.66
 episode: 1950 score: 3.0 memory length: 401919 epsilon: 0.4021984000102357
 steps: 246 lr: 0.0001 evaluation reward: 3.66
 episode: 1951 score: 3.0 memory length: 402165 epsilon:
 0.40171132001023263 steps: 246 lr: 0.0001 evaluation reward: 3.64
 episode: 1952 score: 3.0 memory length: 402409 epsilon: 0.4012282000102296

```

steps: 244    lr: 0.0001    evaluation reward: 3.64
episode: 1953  score: 5.0    memory length: 402734    epsilon: 0.4005847000102255
steps: 325    lr: 0.0001    evaluation reward: 3.66
episode: 1954  score: 4.0    memory length: 402976    epsilon:
0.40010554001022247    steps: 242    lr: 0.0001    evaluation reward: 3.66
episode: 1955  score: 3.0    memory length: 403203    epsilon: 0.3996560800102196
steps: 227    lr: 0.0001    evaluation reward: 3.65
episode: 1956  score: 3.0    memory length: 403448    epsilon:
0.39917098001021656    steps: 245    lr: 0.0001    evaluation reward: 3.61
episode: 1957  score: 1.0    memory length: 403599    epsilon:
0.39887200001021467    steps: 151    lr: 0.0001    evaluation reward: 3.57
episode: 1958  score: 4.0    memory length: 403857    epsilon:
0.39836116001021143    steps: 258    lr: 0.0001    evaluation reward: 3.58
episode: 1959  score: 4.0    memory length: 404170    epsilon: 0.3977414200102075
steps: 313    lr: 0.0001    evaluation reward: 3.59
episode: 1960  score: 3.0    memory length: 404420    epsilon: 0.3972464200102044
steps: 250    lr: 0.0001    evaluation reward: 3.61
episode: 1961  score: 2.0    memory length: 404602    epsilon: 0.3968860600102021
steps: 182    lr: 0.0001    evaluation reward: 3.6
episode: 1962  score: 5.0    memory length: 404948    epsilon:
0.39620098001019777    steps: 346    lr: 0.0001    evaluation reward: 3.62
episode: 1963  score: 4.0    memory length: 405225    epsilon: 0.3956525200101943
steps: 277    lr: 0.0001    evaluation reward: 3.63
episode: 1964  score: 2.0    memory length: 405407    epsilon: 0.395292160010192
steps: 182    lr: 0.0001    evaluation reward: 3.6
episode: 1965  score: 2.0    memory length: 405589    epsilon:
0.39493180001018974    steps: 182    lr: 0.0001    evaluation reward: 3.59
episode: 1966  score: 0.0    memory length: 405712    epsilon: 0.3946882600101882
steps: 123    lr: 0.0001    evaluation reward: 3.56
episode: 1967  score: 4.0    memory length: 405991    epsilon: 0.3941358400101847
steps: 279    lr: 0.0001    evaluation reward: 3.56
episode: 1968  score: 3.0    memory length: 406221    epsilon: 0.3936804400101818
steps: 230    lr: 0.0001    evaluation reward: 3.56
episode: 1969  score: 3.0    memory length: 406468    epsilon: 0.3931913800101787
steps: 247    lr: 0.0001    evaluation reward: 3.54
episode: 1970  score: 7.0    memory length: 406841    epsilon:
0.39245284001017405    steps: 373    lr: 0.0001    evaluation reward: 3.59
episode: 1971  score: 4.0    memory length: 407116    epsilon: 0.3919083400101706
steps: 275    lr: 0.0001    evaluation reward: 3.58
episode: 1972  score: 2.0    memory length: 407298    epsilon:
0.39154798001016833    steps: 182    lr: 0.0001    evaluation reward: 3.56
episode: 1973  score: 3.0    memory length: 407511    epsilon:
0.39112624001016566    steps: 213    lr: 0.0001    evaluation reward: 3.53
episode: 1974  score: 3.0    memory length: 407737    epsilon:
0.39067876001016283    steps: 226    lr: 0.0001    evaluation reward: 3.52
episode: 1975  score: 5.0    memory length: 408043    epsilon: 0.390072880010159
steps: 306    lr: 0.0001    evaluation reward: 3.54
episode: 1976  score: 4.0    memory length: 408305    epsilon: 0.3895541200101557

```

```

steps: 262    lr: 0.0001    evaluation reward: 3.55
episode: 1977  score: 3.0    memory length: 408517    epsilon:
0.38913436001015306    steps: 212    lr: 0.0001    evaluation reward: 3.56
episode: 1978  score: 5.0    memory length: 408862    epsilon:
0.38845126001014874    steps: 345    lr: 0.0001    evaluation reward: 3.57
episode: 1979  score: 6.0    memory length: 409238    epsilon: 0.387706780010144
steps: 376    lr: 0.0001    evaluation reward: 3.61
episode: 1980  score: 7.0    memory length: 409643    epsilon:
0.38690488001013895    steps: 405    lr: 0.0001    evaluation reward: 3.65
episode: 1981  score: 5.0    memory length: 409966    epsilon: 0.3862653400101349
steps: 323    lr: 0.0001    evaluation reward: 3.67
episode: 1982  score: 5.0    memory length: 410293    epsilon: 0.3856178800101308
steps: 327    lr: 0.0001    evaluation reward: 3.69
episode: 1983  score: 4.0    memory length: 410570    epsilon:
0.38506942001012734    steps: 277    lr: 0.0001    evaluation reward: 3.7
episode: 1984  score: 3.0    memory length: 410783    epsilon:
0.38464768001012467    steps: 213    lr: 0.0001    evaluation reward: 3.68
episode: 1985  score: 3.0    memory length: 411010    epsilon: 0.3841982200101218
steps: 227    lr: 0.0001    evaluation reward: 3.69
episode: 1986  score: 4.0    memory length: 411266    epsilon: 0.3836913400101186
steps: 256    lr: 0.0001    evaluation reward: 3.7
episode: 1987  score: 7.0    memory length: 411653    epsilon:
0.38292508001011377    steps: 387    lr: 0.0001    evaluation reward: 3.74
episode: 1988  score: 6.0    memory length: 411988    epsilon: 0.3822617800101096
steps: 335    lr: 0.0001    evaluation reward: 3.72
episode: 1989  score: 4.0    memory length: 412268    epsilon:
0.38170738001010607    steps: 280    lr: 0.0001    evaluation reward: 3.7
episode: 1990  score: 6.0    memory length: 412650    epsilon: 0.3809510200101013
steps: 382    lr: 0.0001    evaluation reward: 3.73
episode: 1991  score: 3.0    memory length: 412897    epsilon: 0.3804619600100982
steps: 247    lr: 0.0001    evaluation reward: 3.73
episode: 1992  score: 3.0    memory length: 413145    epsilon: 0.3799709200100951
steps: 248    lr: 0.0001    evaluation reward: 3.71
episode: 1993  score: 3.0    memory length: 413358    epsilon: 0.3795491800100924
steps: 213    lr: 0.0001    evaluation reward: 3.71
episode: 1994  score: 2.0    memory length: 413540    epsilon:
0.37918882001009013    steps: 182    lr: 0.0001    evaluation reward: 3.71
episode: 1995  score: 3.0    memory length: 413769    epsilon:
0.37873540001008726    steps: 229    lr: 0.0001    evaluation reward: 3.73
episode: 1996  score: 5.0    memory length: 414077    epsilon: 0.3781255600100834
steps: 308    lr: 0.0001    evaluation reward: 3.76
episode: 1997  score: 4.0    memory length: 414356    epsilon: 0.3775731400100799
steps: 279    lr: 0.0001    evaluation reward: 3.75
episode: 1998  score: 4.0    memory length: 414631    epsilon:
0.37702864001007647    steps: 275    lr: 0.0001    evaluation reward: 3.76
episode: 1999  score: 4.0    memory length: 414887    epsilon:
0.37652176001007326    steps: 256    lr: 0.0001    evaluation reward: 3.75
episode: 2000  score: 4.0    memory length: 415164    epsilon: 0.3759733000100698

```



```

steps: 277    lr: 0.0001    evaluation reward: 3.75
episode: 2001  score: 2.0    memory length: 415362    epsilon: 0.3755812600100673
steps: 198    lr: 0.0001    evaluation reward: 3.74
episode: 2002  score: 3.0    memory length: 415608    epsilon: 0.3750941800100642
steps: 246    lr: 0.0001    evaluation reward: 3.74
episode: 2003  score: 7.0    memory length: 416003    epsilon: 0.3743120800100593
steps: 395    lr: 0.0001    evaluation reward: 3.75
episode: 2004  score: 3.0    memory length: 416233    epsilon: 0.3738566800100564
steps: 230    lr: 0.0001    evaluation reward: 3.75
episode: 2005  score: 6.0    memory length: 416604    epsilon:
0.37312210001005175    steps: 371    lr: 0.0001    evaluation reward: 3.77
episode: 2006  score: 5.0    memory length: 416912    epsilon: 0.3725122600100479
steps: 308    lr: 0.0001    evaluation reward: 3.8
episode: 2007  score: 2.0    memory length: 417093    epsilon: 0.3721538800100456
steps: 181    lr: 0.0001    evaluation reward: 3.79
episode: 2008  score: 3.0    memory length: 417304    epsilon: 0.371736100010043
steps: 211    lr: 0.0001    evaluation reward: 3.79
episode: 2009  score: 3.0    memory length: 417514    epsilon:
0.37132030001004035    steps: 210    lr: 0.0001    evaluation reward: 3.78
episode: 2010  score: 4.0    memory length: 417773    epsilon: 0.3708074800100371
steps: 259    lr: 0.0001    evaluation reward: 3.78
episode: 2011  score: 4.0    memory length: 418051    epsilon: 0.3702570400100336
steps: 278    lr: 0.0001    evaluation reward: 3.79
episode: 2012  score: 5.0    memory length: 418358    epsilon: 0.3696491800100298
steps: 307    lr: 0.0001    evaluation reward: 3.8
episode: 2013  score: 7.0    memory length: 418731    epsilon: 0.3689106400100251
steps: 373    lr: 0.0001    evaluation reward: 3.82
episode: 2014  score: 4.0    memory length: 418975    epsilon:
0.36842752001002205    steps: 244    lr: 0.0001    evaluation reward: 3.83
episode: 2015  score: 3.0    memory length: 419186    epsilon: 0.3680097400100194
steps: 211    lr: 0.0001    evaluation reward: 3.83
episode: 2016  score: 6.0    memory length: 419502    epsilon:
0.36738406001001545    steps: 316    lr: 0.0001    evaluation reward: 3.86
episode: 2017  score: 3.0    memory length: 419728    epsilon: 0.3669365800100126
steps: 226    lr: 0.0001    evaluation reward: 3.86
episode: 2018  score: 6.0    memory length: 420084    epsilon:
0.36623170001000815    steps: 356    lr: 0.0001    evaluation reward: 3.88
episode: 2019  score: 4.0    memory length: 420344    epsilon: 0.3657169000100049
steps: 260    lr: 0.0001    evaluation reward: 3.87
episode: 2020  score: 4.0    memory length: 420640    epsilon: 0.3651308200100012
steps: 296    lr: 0.0001    evaluation reward: 3.87
episode: 2021  score: 3.0    memory length: 420907    epsilon:
0.36460216000999784    steps: 267    lr: 0.0001    evaluation reward: 3.87
episode: 2022  score: 3.0    memory length: 421120    epsilon: 0.3641804200099952
steps: 213    lr: 0.0001    evaluation reward: 3.86
episode: 2023  score: 5.0    memory length: 421427    epsilon:
0.36357256000999133    steps: 307    lr: 0.0001    evaluation reward: 3.86
episode: 2024  score: 1.0    memory length: 421578    epsilon:

```

0.36327358000998944 steps: 151 lr: 0.0001 evaluation reward: 3.83
 episode: 2025 score: 3.0 memory length: 421788 epsilon: 0.3628577800099868
 steps: 210 lr: 0.0001 evaluation reward: 3.81
 episode: 2026 score: 4.0 memory length: 422048 epsilon:
 0.36234298000998355 steps: 260 lr: 0.0001 evaluation reward: 3.81
 episode: 2027 score: 7.0 memory length: 422463 epsilon:
 0.36152128000997835 steps: 415 lr: 0.0001 evaluation reward: 3.85
 episode: 2028 score: 3.0 memory length: 422674 epsilon: 0.3611035000099757
 steps: 211 lr: 0.0001 evaluation reward: 3.84
 episode: 2029 score: 3.0 memory length: 422941 epsilon:
 0.36057484000997236 steps: 267 lr: 0.0001 evaluation reward: 3.82
 episode: 2030 score: 6.0 memory length: 423332 epsilon:
 0.35980066000996747 steps: 391 lr: 0.0001 evaluation reward: 3.84
 episode: 2031 score: 5.0 memory length: 423626 epsilon: 0.3592185400099638
 steps: 294 lr: 0.0001 evaluation reward: 3.83
 episode: 2032 score: 11.0 memory length: 424186 epsilon:
 0.35810974000995677 steps: 560 lr: 0.0001 evaluation reward: 3.92
 episode: 2033 score: 4.0 memory length: 424500 epsilon:
 0.35748802000995283 steps: 314 lr: 0.0001 evaluation reward: 3.93
 episode: 2034 score: 7.0 memory length: 424872 epsilon: 0.3567514600099482
 steps: 372 lr: 0.0001 evaluation reward: 3.97
 episode: 2035 score: 4.0 memory length: 425169 epsilon:
 0.35616340000994445 steps: 297 lr: 0.0001 evaluation reward: 3.97
 episode: 2036 score: 5.0 memory length: 425479 epsilon:
 0.35554960000994057 steps: 310 lr: 0.0001 evaluation reward: 3.99
 episode: 2037 score: 4.0 memory length: 425738 epsilon: 0.3550367800099373
 steps: 259 lr: 0.0001 evaluation reward: 4.0
 episode: 2038 score: 5.0 memory length: 426046 epsilon:
 0.35442694000993347 steps: 308 lr: 0.0001 evaluation reward: 4.01
 episode: 2039 score: 3.0 memory length: 426258 epsilon: 0.3540071800099308
 steps: 212 lr: 0.0001 evaluation reward: 3.97
 episode: 2040 score: 5.0 memory length: 426562 epsilon: 0.353405260009927
 steps: 304 lr: 0.0001 evaluation reward: 3.99
 episode: 2041 score: 7.0 memory length: 426945 epsilon: 0.3526469200099222
 steps: 383 lr: 0.0001 evaluation reward: 4.03
 episode: 2042 score: 5.0 memory length: 427274 epsilon: 0.3519955000099181
 steps: 329 lr: 0.0001 evaluation reward: 4.03
 episode: 2043 score: 6.0 memory length: 427627 epsilon:
 0.35129656000991366 steps: 353 lr: 0.0001 evaluation reward: 4.05
 episode: 2044 score: 7.0 memory length: 428032 epsilon: 0.3504946600099086
 steps: 405 lr: 0.0001 evaluation reward: 4.1
 episode: 2045 score: 5.0 memory length: 428338 epsilon:
 0.34988878000990475 steps: 306 lr: 0.0001 evaluation reward: 4.1
 episode: 2046 score: 7.0 memory length: 428736 epsilon:
 0.34910074000989977 steps: 398 lr: 0.0001 evaluation reward: 4.14
 episode: 2047 score: 5.0 memory length: 429025 epsilon:
 0.34852852000989615 steps: 289 lr: 0.0001 evaluation reward: 4.13
 episode: 2048 score: 4.0 memory length: 429288 epsilon:

0.34800778000989285 steps: 263 lr: 0.0001 evaluation reward: 4.15
 episode: 2049 score: 2.0 memory length: 429469 epsilon: 0.3476494000098906
 steps: 181 lr: 0.0001 evaluation reward: 4.15
 episode: 2050 score: 5.0 memory length: 429779 epsilon: 0.3470356000098867
 steps: 310 lr: 0.0001 evaluation reward: 4.17
 episode: 2051 score: 3.0 memory length: 430023 epsilon:
 0.34655248000988365 steps: 244 lr: 0.0001 evaluation reward: 4.17
 episode: 2052 score: 3.0 memory length: 430235 epsilon: 0.346132720009881
 steps: 212 lr: 0.0001 evaluation reward: 4.17
 episode: 2053 score: 2.0 memory length: 430417 epsilon: 0.3457723600098787
 steps: 182 lr: 0.0001 evaluation reward: 4.14
 episode: 2054 score: 9.0 memory length: 430775 epsilon: 0.3450635200098742
 steps: 358 lr: 0.0001 evaluation reward: 4.19
 episode: 2055 score: 6.0 memory length: 431153 epsilon: 0.3443150800098695
 steps: 378 lr: 0.0001 evaluation reward: 4.22
 episode: 2056 score: 7.0 memory length: 431583 epsilon: 0.3434636800098641
 steps: 430 lr: 0.0001 evaluation reward: 4.26
 episode: 2057 score: 6.0 memory length: 431938 epsilon:
 0.34276078000985966 steps: 355 lr: 0.0001 evaluation reward: 4.31
 episode: 2058 score: 3.0 memory length: 432168 epsilon: 0.3423053800098568
 steps: 230 lr: 0.0001 evaluation reward: 4.3
 episode: 2059 score: 5.0 memory length: 432512 epsilon:
 0.34162426000985247 steps: 344 lr: 0.0001 evaluation reward: 4.31
 episode: 2060 score: 7.0 memory length: 432903 epsilon:
 0.34085008000984757 steps: 391 lr: 0.0001 evaluation reward: 4.35
 episode: 2061 score: 6.0 memory length: 433208 epsilon:
 0.34024618000984375 steps: 305 lr: 0.0001 evaluation reward: 4.39
 episode: 2062 score: 2.0 memory length: 433389 epsilon: 0.3398878000098415
 steps: 181 lr: 0.0001 evaluation reward: 4.36
 episode: 2063 score: 5.0 memory length: 433733 epsilon:
 0.33920668000983717 steps: 344 lr: 0.0001 evaluation reward: 4.37
 episode: 2064 score: 4.0 memory length: 433995 epsilon: 0.3386879200098339
 steps: 262 lr: 0.0001 evaluation reward: 4.39
 episode: 2065 score: 3.0 memory length: 434206 epsilon:
 0.33827014000983124 steps: 211 lr: 0.0001 evaluation reward: 4.4
 episode: 2066 score: 6.0 memory length: 434565 epsilon:
 0.33755932000982675 steps: 359 lr: 0.0001 evaluation reward: 4.46
 episode: 2067 score: 4.0 memory length: 434840 epsilon: 0.3370148200098233
 steps: 275 lr: 0.0001 evaluation reward: 4.46
 episode: 2068 score: 4.0 memory length: 435100 epsilon:
 0.33650002000982004 steps: 260 lr: 0.0001 evaluation reward: 4.47
 episode: 2069 score: 6.0 memory length: 435438 epsilon: 0.3358307800098158
 steps: 338 lr: 0.0001 evaluation reward: 4.5
 episode: 2070 score: 4.0 memory length: 435716 epsilon: 0.3352803400098123
 steps: 278 lr: 0.0001 evaluation reward: 4.47
 episode: 2071 score: 4.0 memory length: 435973 epsilon: 0.3347714800098091
 steps: 257 lr: 0.0001 evaluation reward: 4.47
 episode: 2072 score: 5.0 memory length: 436303 epsilon: 0.334118080009805

```

steps: 330    lr: 0.0001    evaluation reward: 4.5
episode: 2073  score: 6.0    memory length: 436658    epsilon: 0.3334151800098005
steps: 355    lr: 0.0001    evaluation reward: 4.53
episode: 2074  score: 4.0    memory length: 436953    epsilon:
0.33283108000979683    steps: 295    lr: 0.0001    evaluation reward: 4.54
episode: 2075  score: 3.0    memory length: 437165    epsilon: 0.3324113200097942
steps: 212    lr: 0.0001    evaluation reward: 4.52
episode: 2076  score: 6.0    memory length: 437500    epsilon: 0.33174802000979
steps: 335    lr: 0.0001    evaluation reward: 4.54
episode: 2077  score: 4.0    memory length: 437756    epsilon:
0.33124114000978677    steps: 256    lr: 0.0001    evaluation reward: 4.55
episode: 2078  score: 3.0    memory length: 437969    epsilon: 0.3308194000097841
steps: 213    lr: 0.0001    evaluation reward: 4.53
episode: 2079  score: 4.0    memory length: 438244    epsilon:
0.33027490000978066    steps: 275    lr: 0.0001    evaluation reward: 4.51
episode: 2080  score: 3.0    memory length: 438454    epsilon:
0.32985910000977803    steps: 210    lr: 0.0001    evaluation reward: 4.47
episode: 2081  score: 3.0    memory length: 438667    epsilon:
0.32943736000977536    steps: 213    lr: 0.0001    evaluation reward: 4.45
episode: 2082  score: 3.0    memory length: 438880    epsilon: 0.3290156200097727
steps: 213    lr: 0.0001    evaluation reward: 4.43
episode: 2083  score: 3.0    memory length: 439093    epsilon: 0.32859388000977
steps: 213    lr: 0.0001    evaluation reward: 4.42
episode: 2084  score: 9.0    memory length: 439592    epsilon:
0.32760586000976377    steps: 499    lr: 0.0001    evaluation reward: 4.48
episode: 2085  score: 4.0    memory length: 439855    epsilon: 0.3270851200097605
steps: 263    lr: 0.0001    evaluation reward: 4.49
episode: 2086  score: 5.0    memory length: 440146    epsilon:
0.32650894000975683    steps: 291    lr: 0.0001    evaluation reward: 4.5
episode: 2087  score: 9.0    memory length: 440505    epsilon:
0.32579812000975233    steps: 359    lr: 0.0001    evaluation reward: 4.52
episode: 2088  score: 5.0    memory length: 440801    epsilon: 0.3252120400097486
steps: 296    lr: 0.0001    evaluation reward: 4.51
episode: 2089  score: 3.0    memory length: 441032    epsilon:
0.32475466000974573    steps: 231    lr: 0.0001    evaluation reward: 4.5
episode: 2090  score: 4.0    memory length: 441287    epsilon:
0.32424976000974254    steps: 255    lr: 0.0001    evaluation reward: 4.48
episode: 2091  score: 5.0    memory length: 441615    epsilon:
0.32360032000973843    steps: 328    lr: 0.0001    evaluation reward: 4.5
episode: 2092  score: 5.0    memory length: 441924    epsilon:
0.32298850000973456    steps: 309    lr: 0.0001    evaluation reward: 4.52
episode: 2093  score: 4.0    memory length: 442199    epsilon: 0.3224440000097311
steps: 275    lr: 0.0001    evaluation reward: 4.53
episode: 2094  score: 3.0    memory length: 442409    epsilon: 0.3220282000097285
steps: 210    lr: 0.0001    evaluation reward: 4.54
episode: 2095  score: 4.0    memory length: 442651    epsilon:
0.32154904000972545    steps: 242    lr: 0.0001    evaluation reward: 4.55
episode: 2096  score: 3.0    memory length: 442899    epsilon:

```

0.32105800000972234 steps: 248 lr: 0.0001 evaluation reward: 4.53
 episode: 2097 score: 6.0 memory length: 443255 epsilon: 0.3203531200097179
 steps: 356 lr: 0.0001 evaluation reward: 4.55
 episode: 2098 score: 6.0 memory length: 443630 epsilon: 0.3196106200097132
 steps: 375 lr: 0.0001 evaluation reward: 4.57
 episode: 2099 score: 5.0 memory length: 443936 epsilon:
 0.31900474000970935 steps: 306 lr: 0.0001 evaluation reward: 4.58
 episode: 2100 score: 6.0 memory length: 444292 epsilon: 0.3182998600097049
 steps: 356 lr: 0.0001 evaluation reward: 4.6
 episode: 2101 score: 4.0 memory length: 444567 epsilon:
 0.31775536000970145 steps: 275 lr: 0.0001 evaluation reward: 4.62
 episode: 2102 score: 4.0 memory length: 444809 epsilon: 0.3172762000096984
 steps: 242 lr: 0.0001 evaluation reward: 4.63
 episode: 2103 score: 4.0 memory length: 445067 epsilon: 0.3167653600096952
 steps: 258 lr: 0.0001 evaluation reward: 4.6
 episode: 2104 score: 4.0 memory length: 445346 epsilon: 0.3162129400096917
 steps: 279 lr: 0.0001 evaluation reward: 4.61
 episode: 2105 score: 5.0 memory length: 445632 epsilon: 0.3156466600096881
 steps: 286 lr: 0.0001 evaluation reward: 4.6
 episode: 2106 score: 3.0 memory length: 445862 epsilon: 0.3151912600096852
 steps: 230 lr: 0.0001 evaluation reward: 4.58
 episode: 2107 score: 6.0 memory length: 446181 epsilon:
 0.31455964000968123 steps: 319 lr: 0.0001 evaluation reward: 4.62
 episode: 2108 score: 7.0 memory length: 446585 epsilon:
 0.31375972000967617 steps: 404 lr: 0.0001 evaluation reward: 4.66
 episode: 2109 score: 3.0 memory length: 446815 epsilon: 0.3133043200096733
 steps: 230 lr: 0.0001 evaluation reward: 4.66
 episode: 2110 score: 3.0 memory length: 447028 epsilon: 0.3128825800096706
 steps: 213 lr: 0.0001 evaluation reward: 4.65
 episode: 2111 score: 3.0 memory length: 447241 epsilon:
 0.31246084000966795 steps: 213 lr: 0.0001 evaluation reward: 4.64
 episode: 2112 score: 6.0 memory length: 447596 epsilon: 0.3117579400096635
 steps: 355 lr: 0.0001 evaluation reward: 4.65
 episode: 2113 score: 3.0 memory length: 447825 epsilon:
 0.31130452000966063 steps: 229 lr: 0.0001 evaluation reward: 4.61
 episode: 2114 score: 3.0 memory length: 448055 epsilon:
 0.31084912000965775 steps: 230 lr: 0.0001 evaluation reward: 4.6
 episode: 2115 score: 3.0 memory length: 448282 epsilon: 0.3103996600096549
 steps: 227 lr: 0.0001 evaluation reward: 4.6
 episode: 2116 score: 5.0 memory length: 448606 epsilon:
 0.30975814000965085 steps: 324 lr: 0.0001 evaluation reward: 4.59
 episode: 2117 score: 4.0 memory length: 448885 epsilon:
 0.30920572000964736 steps: 279 lr: 0.0001 evaluation reward: 4.6
 episode: 2118 score: 10.0 memory length: 449381 epsilon:
 0.30822364000964114 steps: 496 lr: 0.0001 evaluation reward: 4.64
 episode: 2119 score: 5.0 memory length: 449687 epsilon: 0.3076177600096373
 steps: 306 lr: 0.0001 evaluation reward: 4.65
 episode: 2120 score: 5.0 memory length: 449993 epsilon: 0.3070118800096335

```

steps: 306    lr: 0.0001    evaluation reward: 4.66
episode: 2121  score: 3.0    memory length: 450206    epsilon: 0.3065901400096308
steps: 213    lr: 0.0001    evaluation reward: 4.66
episode: 2122  score: 5.0    memory length: 450518    epsilon: 0.3059723800096269
steps: 312    lr: 0.0001    evaluation reward: 4.68
episode: 2123  score: 4.0    memory length: 450795    epsilon:
0.30542392000962343    steps: 277    lr: 0.0001    evaluation reward: 4.67
episode: 2124  score: 6.0    memory length: 451166    epsilon: 0.3046893400096188
steps: 371    lr: 0.0001    evaluation reward: 4.72
episode: 2125  score: 7.0    memory length: 451538    epsilon: 0.3039527800096141
steps: 372    lr: 0.0001    evaluation reward: 4.76
episode: 2126  score: 5.0    memory length: 451863    epsilon:
0.30330928000961005    steps: 325    lr: 0.0001    evaluation reward: 4.77
episode: 2127  score: 6.0    memory length: 452163    epsilon: 0.3027152800096063
steps: 300    lr: 0.0001    evaluation reward: 4.76
episode: 2128  score: 7.0    memory length: 452510    epsilon:
0.30202822000960194    steps: 347    lr: 0.0001    evaluation reward: 4.8
episode: 2129  score: 5.0    memory length: 452816    epsilon: 0.3014223400095981
steps: 306    lr: 0.0001    evaluation reward: 4.82
episode: 2130  score: 7.0    memory length: 453202    epsilon: 0.3006580600095933
steps: 386    lr: 0.0001    evaluation reward: 4.83
episode: 2131  score: 5.0    memory length: 453527    epsilon: 0.3000145600095892
steps: 325    lr: 0.0001    evaluation reward: 4.83
episode: 2132  score: 5.0    memory length: 453831    epsilon: 0.2994126400095854
steps: 304    lr: 0.0001    evaluation reward: 4.77
episode: 2133  score: 3.0    memory length: 454079    epsilon: 0.2989216000095823
steps: 248    lr: 0.0001    evaluation reward: 4.76
episode: 2134  score: 6.0    memory length: 454437    epsilon: 0.2982127600095778
steps: 358    lr: 0.0001    evaluation reward: 4.75
episode: 2135  score: 5.0    memory length: 454770    epsilon:
0.29755342000957363    steps: 333    lr: 0.0001    evaluation reward: 4.76
episode: 2136  score: 3.0    memory length: 454983    epsilon:
0.29713168000957096    steps: 213    lr: 0.0001    evaluation reward: 4.74
episode: 2137  score: 10.0    memory length: 455486    epsilon:
0.29613574000956466    steps: 503    lr: 0.0001    evaluation reward: 4.8
episode: 2138  score: 7.0    memory length: 455895    epsilon:
0.29532592000955954    steps: 409    lr: 0.0001    evaluation reward: 4.82
episode: 2139  score: 5.0    memory length: 456219    epsilon: 0.2946844000095555
steps: 324    lr: 0.0001    evaluation reward: 4.84
episode: 2140  score: 6.0    memory length: 456570    epsilon: 0.2939894200095511
steps: 351    lr: 0.0001    evaluation reward: 4.85
episode: 2141  score: 3.0    memory length: 456783    epsilon: 0.2935676800095484
steps: 213    lr: 0.0001    evaluation reward: 4.81
episode: 2142  score: 5.0    memory length: 457092    epsilon:
0.29295586000954454    steps: 309    lr: 0.0001    evaluation reward: 4.81
episode: 2143  score: 7.0    memory length: 457485    epsilon: 0.2921777200095396
steps: 393    lr: 0.0001    evaluation reward: 4.82
episode: 2144  score: 5.0    memory length: 457806    epsilon: 0.2915421400095356

```

```

steps: 321    lr: 0.0001    evaluation reward: 4.8
episode: 2145  score: 5.0    memory length: 458135    epsilon: 0.2908907200095315
steps: 329    lr: 0.0001    evaluation reward: 4.8
episode: 2146  score: 5.0    memory length: 458457    epsilon:
0.29025316000952744    steps: 322    lr: 0.0001    evaluation reward: 4.78
episode: 2147  score: 4.0    memory length: 458737    epsilon:
0.28969876000952394    steps: 280    lr: 0.0001    evaluation reward: 4.77
episode: 2148  score: 5.0    memory length: 459067    epsilon: 0.2890453600095198
steps: 330    lr: 0.0001    evaluation reward: 4.78
episode: 2149  score: 3.0    memory length: 459277    epsilon:
0.28862956000951717    steps: 210    lr: 0.0001    evaluation reward: 4.79
episode: 2150  score: 5.0    memory length: 459550    epsilon:
0.28808902000951375    steps: 273    lr: 0.0001    evaluation reward: 4.79
episode: 2151  score: 9.0    memory length: 459875    epsilon: 0.2874455200095097
steps: 325    lr: 0.0001    evaluation reward: 4.85
episode: 2152  score: 11.0   memory length: 460314    epsilon:
0.2865763000095042    steps: 439    lr: 0.0001    evaluation reward: 4.93
episode: 2153  score: 5.0    memory length: 460613    epsilon:
0.28598428000950044    steps: 299    lr: 0.0001    evaluation reward: 4.96
episode: 2154  score: 6.0    memory length: 460967    epsilon: 0.285283360009496
steps: 354    lr: 0.0001    evaluation reward: 4.93
episode: 2155  score: 6.0    memory length: 461308    epsilon:
0.28460818000949173    steps: 341    lr: 0.0001    evaluation reward: 4.93
episode: 2156  score: 4.0    memory length: 461572    epsilon: 0.2840854600094884
steps: 264    lr: 0.0001    evaluation reward: 4.9
episode: 2157  score: 4.0    memory length: 461832    epsilon:
0.28357066000948516    steps: 260    lr: 0.0001    evaluation reward: 4.88
episode: 2158  score: 3.0    memory length: 462045    epsilon: 0.2831489200094825
steps: 213    lr: 0.0001    evaluation reward: 4.88
episode: 2159  score: 4.0    memory length: 462343    epsilon:
0.28255888000947876    steps: 298    lr: 0.0001    evaluation reward: 4.87
episode: 2160  score: 4.0    memory length: 462587    epsilon: 0.2820757600094757
steps: 244    lr: 0.0001    evaluation reward: 4.84
episode: 2161  score: 4.0    memory length: 462829    epsilon: 0.2815966000094727
steps: 242    lr: 0.0001    evaluation reward: 4.82
episode: 2162  score: 6.0    memory length: 463172    epsilon: 0.2809174600094684
steps: 343    lr: 0.0001    evaluation reward: 4.86
episode: 2163  score: 3.0    memory length: 463402    epsilon: 0.2804620600094655
steps: 230    lr: 0.0001    evaluation reward: 4.84
episode: 2164  score: 4.0    memory length: 463676    epsilon:
0.27991954000946206    steps: 274    lr: 0.0001    evaluation reward: 4.84
episode: 2165  score: 4.0    memory length: 463952    epsilon: 0.2793730600094586
steps: 276    lr: 0.0001    evaluation reward: 4.85
episode: 2166  score: 7.0    memory length: 464299    epsilon:
0.27868600000945426    steps: 347    lr: 0.0001    evaluation reward: 4.86
episode: 2167  score: 4.0    memory length: 464555    epsilon:
0.27817912000945105    steps: 256    lr: 0.0001    evaluation reward: 4.86
episode: 2168  score: 6.0    memory length: 464872    epsilon: 0.2775514600094471

```

```

steps: 317    lr: 0.0001    evaluation reward: 4.88
episode: 2169  score: 4.0    memory length: 465113    epsilon:
0.27707428000944406    steps: 241    lr: 0.0001    evaluation reward: 4.86
episode: 2170  score: 5.0    memory length: 465437    epsilon: 0.27643276000944
steps: 324    lr: 0.0001    evaluation reward: 4.87
episode: 2171  score: 3.0    memory length: 465647    epsilon: 0.2760169600094374
steps: 210    lr: 0.0001    evaluation reward: 4.86
episode: 2172  score: 3.0    memory length: 465859    epsilon: 0.2755972000094347
steps: 212    lr: 0.0001    evaluation reward: 4.84
episode: 2173  score: 4.0    memory length: 466135    epsilon:
0.27505072000943126    steps: 276    lr: 0.0001    evaluation reward: 4.82
episode: 2174  score: 7.0    memory length: 466535    epsilon:
0.27425872000942625    steps: 400    lr: 0.0001    evaluation reward: 4.85
episode: 2175  score: 4.0    memory length: 466776    epsilon:
0.27378154000942323    steps: 241    lr: 0.0001    evaluation reward: 4.86
episode: 2176  score: 3.0    memory length: 467026    epsilon: 0.2732865400094201
steps: 250    lr: 0.0001    evaluation reward: 4.83
episode: 2177  score: 7.0    memory length: 467412    epsilon:
0.27252226000941526    steps: 386    lr: 0.0001    evaluation reward: 4.86
episode: 2178  score: 4.0    memory length: 467670    epsilon:
0.27201142000941203    steps: 258    lr: 0.0001    evaluation reward: 4.87
episode: 2179  score: 5.0    memory length: 467974    epsilon: 0.2714095000094082
steps: 304    lr: 0.0001    evaluation reward: 4.88
episode: 2180  score: 3.0    memory length: 468204    epsilon:
0.27095410000940534    steps: 230    lr: 0.0001    evaluation reward: 4.88
episode: 2181  score: 5.0    memory length: 468527    epsilon: 0.2703145600094013
steps: 323    lr: 0.0001    evaluation reward: 4.9
episode: 2182  score: 6.0    memory length: 468920    epsilon:
0.26953642000939637    steps: 393    lr: 0.0001    evaluation reward: 4.93
episode: 2183  score: 4.0    memory length: 469161    epsilon:
0.26905924000939335    steps: 241    lr: 0.0001    evaluation reward: 4.94
episode: 2184  score: 13.0   memory length: 469678    epsilon:
0.2680355800093869    steps: 517    lr: 0.0001    evaluation reward: 4.98
episode: 2185  score: 4.0    memory length: 469956    epsilon: 0.2674851400093834
steps: 278    lr: 0.0001    evaluation reward: 4.98
episode: 2186  score: 4.0    memory length: 470217    epsilon: 0.2669683600093801
steps: 261    lr: 0.0001    evaluation reward: 4.97
episode: 2187  score: 6.0    memory length: 470574    epsilon:
0.26626150000937565    steps: 357    lr: 0.0001    evaluation reward: 4.94
episode: 2188  score: 4.0    memory length: 470833    epsilon: 0.2657486800093724
steps: 259    lr: 0.0001    evaluation reward: 4.93
episode: 2189  score: 2.0    memory length: 471013    epsilon:
0.26539228000937015    steps: 180    lr: 0.0001    evaluation reward: 4.92
episode: 2190  score: 3.0    memory length: 471261    epsilon:
0.26490124000936704    steps: 248    lr: 0.0001    evaluation reward: 4.91
episode: 2191  score: 6.0    memory length: 471604    epsilon:
0.26422210000936275    steps: 343    lr: 0.0001    evaluation reward: 4.92
episode: 2192  score: 4.0    memory length: 471842    epsilon:

```


0.26375086000935977 steps: 238 lr: 0.0001 evaluation reward: 4.91
 episode: 2193 score: 4.0 memory length: 472084 epsilon:
 0.26327170000935673 steps: 242 lr: 0.0001 evaluation reward: 4.91
 episode: 2194 score: 3.0 memory length: 472294 epsilon: 0.2628559000093541
 steps: 210 lr: 0.0001 evaluation reward: 4.91
 episode: 2195 score: 2.0 memory length: 472494 epsilon: 0.2624599000093516
 steps: 200 lr: 0.0001 evaluation reward: 4.89
 episode: 2196 score: 3.0 memory length: 472722 epsilon:
 0.26200846000934874 steps: 228 lr: 0.0001 evaluation reward: 4.89
 episode: 2197 score: 4.0 memory length: 472963 epsilon: 0.2615312800093457
 steps: 241 lr: 0.0001 evaluation reward: 4.87
 episode: 2198 score: 4.0 memory length: 473242 epsilon:
 0.26097886000934223 steps: 279 lr: 0.0001 evaluation reward: 4.85
 episode: 2199 score: 5.0 memory length: 473567 epsilon:
 0.26033536000933816 steps: 325 lr: 0.0001 evaluation reward: 4.85
 episode: 2200 score: 11.0 memory length: 474030 epsilon:
 0.25941862000933236 steps: 463 lr: 0.0001 evaluation reward: 4.9
 episode: 2201 score: 3.0 memory length: 474241 epsilon: 0.2590008400093297
 steps: 211 lr: 0.0001 evaluation reward: 4.89
 episode: 2202 score: 4.0 memory length: 474501 epsilon:
 0.25848604000932646 steps: 260 lr: 0.0001 evaluation reward: 4.89
 episode: 2203 score: 6.0 memory length: 474838 epsilon:
 0.25781878000932223 steps: 337 lr: 0.0001 evaluation reward: 4.91
 episode: 2204 score: 11.0 memory length: 475260 epsilon:
 0.25698322000931695 steps: 422 lr: 0.0001 evaluation reward: 4.98
 episode: 2205 score: 5.0 memory length: 475581 epsilon: 0.2563476400093129
 steps: 321 lr: 0.0001 evaluation reward: 4.98
 episode: 2206 score: 4.0 memory length: 475820 epsilon:
 0.25587442000930993 steps: 239 lr: 0.0001 evaluation reward: 4.99
 episode: 2207 score: 4.0 memory length: 476099 epsilon:
 0.2553220000930644 steps: 279 lr: 0.0001 evaluation reward: 4.97
 episode: 2208 score: 3.0 memory length: 476328 epsilon:
 0.25486858000930357 steps: 229 lr: 0.0001 evaluation reward: 4.93
 episode: 2209 score: 4.0 memory length: 476588 epsilon: 0.2543537800093003
 steps: 260 lr: 0.0001 evaluation reward: 4.94
 episode: 2210 score: 3.0 memory length: 476815 epsilon:
 0.25390432000929747 steps: 227 lr: 0.0001 evaluation reward: 4.94
 episode: 2211 score: 3.0 memory length: 477027 epsilon: 0.2534845600092948
 steps: 212 lr: 0.0001 evaluation reward: 4.94
 episode: 2212 score: 6.0 memory length: 477383 epsilon:
 0.25277968000929035 steps: 356 lr: 0.0001 evaluation reward: 4.94
 episode: 2213 score: 6.0 memory length: 477723 epsilon: 0.2521064800092861
 steps: 340 lr: 0.0001 evaluation reward: 4.97
 episode: 2214 score: 5.0 memory length: 478053 epsilon:
 0.25145308000928196 steps: 330 lr: 0.0001 evaluation reward: 4.99
 episode: 2215 score: 6.0 memory length: 478428 epsilon:
 0.25071058000927726 steps: 375 lr: 0.0001 evaluation reward: 5.02
 episode: 2216 score: 5.0 memory length: 478750 epsilon:

0.25007302000927323 steps: 322 lr: 0.0001 evaluation reward: 5.02
 episode: 2217 score: 4.0 memory length: 478993 epsilon:
 0.24959188000927018 steps: 243 lr: 0.0001 evaluation reward: 5.02
 episode: 2218 score: 2.0 memory length: 479174 epsilon:
 0.24923350000926792 steps: 181 lr: 0.0001 evaluation reward: 4.94
 episode: 2219 score: 7.0 memory length: 479577 epsilon:
 0.24843556000926287 steps: 403 lr: 0.0001 evaluation reward: 4.96
 episode: 2220 score: 4.0 memory length: 479833 epsilon:
 0.24792868000925966 steps: 256 lr: 0.0001 evaluation reward: 4.95
 episode: 2221 score: 4.0 memory length: 480128 epsilon:
 0.24734458000925597 steps: 295 lr: 0.0001 evaluation reward: 4.96
 episode: 2222 score: 6.0 memory length: 480466 epsilon:
 0.24667534000925173 steps: 338 lr: 0.0001 evaluation reward: 4.97
 episode: 2223 score: 3.0 memory length: 480675 epsilon: 0.2462615200092491
 steps: 209 lr: 0.0001 evaluation reward: 4.96
 episode: 2224 score: 6.0 memory length: 480978 epsilon:
 0.24566158000924532 steps: 303 lr: 0.0001 evaluation reward: 4.96
 episode: 2225 score: 7.0 memory length: 481361 epsilon:
 0.24490324000924052 steps: 383 lr: 0.0001 evaluation reward: 4.96
 episode: 2226 score: 3.0 memory length: 481574 epsilon:
 0.24448150000923785 steps: 213 lr: 0.0001 evaluation reward: 4.94
 episode: 2227 score: 8.0 memory length: 482011 epsilon:
 0.24361624000923238 steps: 437 lr: 0.0001 evaluation reward: 4.96
 episode: 2228 score: 7.0 memory length: 482391 epsilon:
 0.24286384000922762 steps: 380 lr: 0.0001 evaluation reward: 4.96
 episode: 2229 score: 4.0 memory length: 482687 epsilon: 0.2422777600092239
 steps: 296 lr: 0.0001 evaluation reward: 4.95
 episode: 2230 score: 9.0 memory length: 483172 epsilon:
 0.24131746000921783 steps: 485 lr: 0.0001 evaluation reward: 4.97
 episode: 2231 score: 7.0 memory length: 483576 epsilon:
 0.24051754000921277 steps: 404 lr: 0.0001 evaluation reward: 4.99
 episode: 2232 score: 5.0 memory length: 483881 epsilon:
 0.23991364000920895 steps: 305 lr: 0.0001 evaluation reward: 4.99
 episode: 2233 score: 8.0 memory length: 484333 epsilon: 0.2390186800092033
 steps: 452 lr: 0.0001 evaluation reward: 5.04
 episode: 2234 score: 4.0 memory length: 484589 epsilon:
 0.23851180000920008 steps: 256 lr: 0.0001 evaluation reward: 5.02
 episode: 2235 score: 7.0 memory length: 484980 epsilon:
 0.23773762000919518 steps: 391 lr: 0.0001 evaluation reward: 5.04
 episode: 2236 score: 2.0 memory length: 485161 epsilon:
 0.23737924000919292 steps: 181 lr: 0.0001 evaluation reward: 5.03
 episode: 2237 score: 6.0 memory length: 485500 epsilon:
 0.23670802000918867 steps: 339 lr: 0.0001 evaluation reward: 4.99
 episode: 2238 score: 6.0 memory length: 485876 epsilon:
 0.23596354000918396 steps: 376 lr: 0.0001 evaluation reward: 4.98
 episode: 2239 score: 5.0 memory length: 486198 epsilon:
 0.23532598000917992 steps: 322 lr: 0.0001 evaluation reward: 4.98
 episode: 2240 score: 4.0 memory length: 486475 epsilon:

0.23477752000917645 steps: 277 lr: 0.0001 evaluation reward: 4.96
 episode: 2241 score: 8.0 memory length: 486906 epsilon:
 0.23392414000917106 steps: 431 lr: 0.0001 evaluation reward: 5.01
 episode: 2242 score: 3.0 memory length: 487135 epsilon: 0.2334707200091682
 steps: 229 lr: 0.0001 evaluation reward: 4.99
 episode: 2243 score: 3.0 memory length: 487344 epsilon:
 0.23305690000916557 steps: 209 lr: 0.0001 evaluation reward: 4.95
 episode: 2244 score: 9.0 memory length: 487796 epsilon: 0.2321619400091599
 steps: 452 lr: 0.0001 evaluation reward: 4.99
 episode: 2245 score: 7.0 memory length: 488182 epsilon:
 0.23139766000915507 steps: 386 lr: 0.0001 evaluation reward: 5.01
 episode: 2246 score: 4.0 memory length: 488461 epsilon:
 0.23084524000915158 steps: 279 lr: 0.0001 evaluation reward: 5.0
 episode: 2247 score: 4.0 memory length: 488741 epsilon:
 0.23029084000914807 steps: 280 lr: 0.0001 evaluation reward: 5.0
 episode: 2248 score: 3.0 memory length: 488954 epsilon: 0.2298691000091454
 steps: 213 lr: 0.0001 evaluation reward: 4.98
 episode: 2249 score: 7.0 memory length: 489357 epsilon:
 0.22907116000914035 steps: 403 lr: 0.0001 evaluation reward: 5.02
 episode: 2250 score: 6.0 memory length: 489711 epsilon:
 0.22837024000913592 steps: 354 lr: 0.0001 evaluation reward: 5.03
 episode: 2251 score: 3.0 memory length: 489921 epsilon:
 0.22795444000913329 steps: 210 lr: 0.0001 evaluation reward: 4.97
 episode: 2252 score: 5.0 memory length: 490249 epsilon:
 0.22730500000912918 steps: 328 lr: 0.0001 evaluation reward: 4.91
 episode: 2253 score: 4.0 memory length: 490515 epsilon:
 0.22677832000912584 steps: 266 lr: 0.0001 evaluation reward: 4.9
 episode: 2254 score: 4.0 memory length: 490775 epsilon: 0.2262635200091226
 steps: 260 lr: 0.0001 evaluation reward: 4.88
 episode: 2255 score: 5.0 memory length: 491089 epsilon:
 0.22564180000911865 steps: 314 lr: 0.0001 evaluation reward: 4.87
 episode: 2256 score: 5.0 memory length: 491435 epsilon:
 0.22495672000911432 steps: 346 lr: 0.0001 evaluation reward: 4.88
 episode: 2257 score: 5.0 memory length: 491744 epsilon:
 0.22434490000911045 steps: 309 lr: 0.0001 evaluation reward: 4.89
 episode: 2258 score: 6.0 memory length: 492099 epsilon: 0.223642000009106
 steps: 355 lr: 0.0001 evaluation reward: 4.92
 episode: 2259 score: 5.0 memory length: 492389 epsilon:
 0.22306780000910237 steps: 290 lr: 0.0001 evaluation reward: 4.93
 episode: 2260 score: 5.0 memory length: 492719 epsilon:
 0.22241440000909823 steps: 330 lr: 0.0001 evaluation reward: 4.94
 episode: 2261 score: 6.0 memory length: 493074 epsilon: 0.2217115000090938
 steps: 355 lr: 0.0001 evaluation reward: 4.96
 episode: 2262 score: 6.0 memory length: 493427 epsilon:
 0.22101256000908936 steps: 353 lr: 0.0001 evaluation reward: 4.96
 episode: 2263 score: 3.0 memory length: 493639 epsilon: 0.2205928000090867
 steps: 212 lr: 0.0001 evaluation reward: 4.96
 episode: 2264 score: 3.0 memory length: 493852 epsilon:

0.22017106000908404 steps: 213 lr: 0.0001 evaluation reward: 4.95
 episode: 2265 score: 11.0 memory length: 494381 epsilon:
 0.21912364000907741 steps: 529 lr: 0.0001 evaluation reward: 5.02
 episode: 2266 score: 7.0 memory length: 494780 epsilon:
 0.21833362000907242 steps: 399 lr: 0.0001 evaluation reward: 5.02
 episode: 2267 score: 8.0 memory length: 495258 epsilon:
 0.21738718000906643 steps: 478 lr: 0.0001 evaluation reward: 5.06
 episode: 2268 score: 5.0 memory length: 495590 epsilon:
 0.21672982000906227 steps: 332 lr: 0.0001 evaluation reward: 5.05
 episode: 2269 score: 5.0 memory length: 495880 epsilon:
 0.21615562000905864 steps: 290 lr: 0.0001 evaluation reward: 5.06
 episode: 2270 score: 6.0 memory length: 496258 epsilon: 0.2154071800090539
 steps: 378 lr: 0.0001 evaluation reward: 5.07
 episode: 2271 score: 5.0 memory length: 496548 epsilon:
 0.21483298000905027 steps: 290 lr: 0.0001 evaluation reward: 5.09
 episode: 2272 score: 3.0 memory length: 496761 epsilon: 0.2144112400090476
 steps: 213 lr: 0.0001 evaluation reward: 5.09
 episode: 2273 score: 9.0 memory length: 497208 epsilon: 0.213526180009042
 steps: 447 lr: 0.0001 evaluation reward: 5.14
 episode: 2274 score: 4.0 memory length: 497487 epsilon: 0.2129737600090385
 steps: 279 lr: 0.0001 evaluation reward: 5.11
 episode: 2275 score: 3.0 memory length: 497700 epsilon:
 0.21255202000903584 steps: 213 lr: 0.0001 evaluation reward: 5.1
 episode: 2276 score: 7.0 memory length: 498087 epsilon: 0.211785760009031
 steps: 387 lr: 0.0001 evaluation reward: 5.14
 episode: 2277 score: 6.0 memory length: 498441 epsilon:
 0.21108484000902655 steps: 354 lr: 0.0001 evaluation reward: 5.13
 episode: 2278 score: 6.0 memory length: 498784 epsilon:
 0.21040570000902226 steps: 343 lr: 0.0001 evaluation reward: 5.15
 episode: 2279 score: 5.0 memory length: 499095 epsilon:
 0.20978992000901836 steps: 311 lr: 0.0001 evaluation reward: 5.15
 episode: 2280 score: 7.0 memory length: 499460 epsilon: 0.2090672200090138
 steps: 365 lr: 0.0001 evaluation reward: 5.19
 episode: 2281 score: 6.0 memory length: 499803 epsilon: 0.2083880800090095
 steps: 343 lr: 0.0001 evaluation reward: 5.2
 episode: 2282 score: 8.0 memory length: 500241 epsilon: 0.207520840009004
 steps: 438 lr: 0.0001 evaluation reward: 5.22
 episode: 2283 score: 3.0 memory length: 500451 epsilon:
 0.20710504000900137 steps: 210 lr: 0.0001 evaluation reward: 5.21
 episode: 2284 score: 6.0 memory length: 500790 epsilon:
 0.20643382000899713 steps: 339 lr: 0.0001 evaluation reward: 5.14
 episode: 2285 score: 5.0 memory length: 501138 epsilon:
 0.20574478000899277 steps: 348 lr: 0.0001 evaluation reward: 5.15
 episode: 2286 score: 5.0 memory length: 501462 epsilon: 0.2051032600089887
 steps: 324 lr: 0.0001 evaluation reward: 5.16
 episode: 2287 score: 3.0 memory length: 501693 epsilon:
 0.20464588000898581 steps: 231 lr: 0.0001 evaluation reward: 5.13
 episode: 2288 score: 3.0 memory length: 501906 epsilon:

0.20422414000898315 steps: 213 lr: 0.0001 evaluation reward: 5.12
 episode: 2289 score: 4.0 memory length: 502168 epsilon:
 0.20370538000897986 steps: 262 lr: 0.0001 evaluation reward: 5.14
 episode: 2290 score: 3.0 memory length: 502381 epsilon: 0.2032836400089772
 steps: 213 lr: 0.0001 evaluation reward: 5.14
 episode: 2291 score: 6.0 memory length: 502699 epsilon: 0.2026540000089732
 steps: 318 lr: 0.0001 evaluation reward: 5.14
 episode: 2292 score: 7.0 memory length: 503074 epsilon:
 0.20191150000896851 steps: 375 lr: 0.0001 evaluation reward: 5.17
 episode: 2293 score: 7.0 memory length: 503461 epsilon:
 0.20114524000896367 steps: 387 lr: 0.0001 evaluation reward: 5.2
 episode: 2294 score: 5.0 memory length: 503787 epsilon:
 0.20049976000895958 steps: 326 lr: 0.0001 evaluation reward: 5.22
 episode: 2295 score: 5.0 memory length: 504091 epsilon:
 0.19989784000895577 steps: 304 lr: 0.0001 evaluation reward: 5.25
 episode: 2296 score: 6.0 memory length: 504449 epsilon: 0.1991890000089513
 steps: 358 lr: 0.0001 evaluation reward: 5.28
 episode: 2297 score: 8.0 memory length: 504884 epsilon:
 0.19832770000894584 steps: 435 lr: 0.0001 evaluation reward: 5.32
 episode: 2298 score: 4.0 memory length: 505126 epsilon: 0.1978485400089428
 steps: 242 lr: 0.0001 evaluation reward: 5.32
 episode: 2299 score: 8.0 memory length: 505523 epsilon:
 0.19706248000893783 steps: 397 lr: 0.0001 evaluation reward: 5.35
 episode: 2300 score: 5.0 memory length: 505829 epsilon: 0.196456600008934
 steps: 306 lr: 0.0001 evaluation reward: 5.29
 episode: 2301 score: 6.0 memory length: 506142 epsilon:
 0.19583686000893008 steps: 313 lr: 0.0001 evaluation reward: 5.32
 episode: 2302 score: 7.0 memory length: 506567 epsilon:
 0.19499536000892476 steps: 425 lr: 0.0001 evaluation reward: 5.35
 episode: 2303 score: 6.0 memory length: 506904 epsilon:
 0.19432810000892053 steps: 337 lr: 0.0001 evaluation reward: 5.35
 episode: 2304 score: 5.0 memory length: 507226 epsilon: 0.1936905400089165
 steps: 322 lr: 0.0001 evaluation reward: 5.29
 episode: 2305 score: 6.0 memory length: 507574 epsilon:
 0.19300150000891214 steps: 348 lr: 0.0001 evaluation reward: 5.3
 episode: 2306 score: 5.0 memory length: 507894 epsilon:
 0.19236790000890813 steps: 320 lr: 0.0001 evaluation reward: 5.31
 episode: 2307 score: 5.0 memory length: 508184 epsilon: 0.1917937000089045
 steps: 290 lr: 0.0001 evaluation reward: 5.32
 episode: 2308 score: 6.0 memory length: 508505 epsilon:
 0.19115812000890048 steps: 321 lr: 0.0001 evaluation reward: 5.35
 episode: 2309 score: 5.0 memory length: 508795 epsilon:
 0.19058392000889685 steps: 290 lr: 0.0001 evaluation reward: 5.36
 episode: 2310 score: 7.0 memory length: 509043 epsilon:
 0.19009288000889374 steps: 248 lr: 0.0001 evaluation reward: 5.4
 episode: 2311 score: 4.0 memory length: 509320 epsilon:
 0.18954442000889027 steps: 277 lr: 0.0001 evaluation reward: 5.41
 episode: 2312 score: 4.0 memory length: 509578 epsilon:

0.18903358000888704 steps: 258 lr: 0.0001 evaluation reward: 5.39
 episode: 2313 score: 3.0 memory length: 509788 epsilon: 0.1886177800088844
 steps: 210 lr: 0.0001 evaluation reward: 5.36
 episode: 2314 score: 5.0 memory length: 510095 epsilon:
 0.18800992000888056 steps: 307 lr: 0.0001 evaluation reward: 5.36
 episode: 2315 score: 8.0 memory length: 510545 epsilon:
 0.18711892000887492 steps: 450 lr: 0.0001 evaluation reward: 5.38
 episode: 2316 score: 3.0 memory length: 510758 epsilon:
 0.18669718000887225 steps: 213 lr: 0.0001 evaluation reward: 5.36
 episode: 2317 score: 2.0 memory length: 510956 epsilon:
 0.18630514000886977 steps: 198 lr: 0.0001 evaluation reward: 5.34
 episode: 2318 score: 6.0 memory length: 511332 epsilon:
 0.18556066000886506 steps: 376 lr: 0.0001 evaluation reward: 5.38
 episode: 2319 score: 5.0 memory length: 511655 epsilon:
 0.18492112000886102 steps: 323 lr: 0.0001 evaluation reward: 5.36
 episode: 2320 score: 6.0 memory length: 511975 epsilon: 0.184287520008857
 steps: 320 lr: 0.0001 evaluation reward: 5.38
 episode: 2321 score: 3.0 memory length: 512186 epsilon:
 0.18386974000885437 steps: 211 lr: 0.0001 evaluation reward: 5.37
 episode: 2322 score: 6.0 memory length: 512561 epsilon:
 0.18312724000884967 steps: 375 lr: 0.0001 evaluation reward: 5.37
 episode: 2323 score: 8.0 memory length: 512978 epsilon:
 0.18230158000884444 steps: 417 lr: 0.0001 evaluation reward: 5.42
 episode: 2324 score: 3.0 memory length: 513190 epsilon: 0.1818818200088418
 steps: 212 lr: 0.0001 evaluation reward: 5.39
 episode: 2325 score: 6.0 memory length: 513563 epsilon:
 0.18114328000883712 steps: 373 lr: 0.0001 evaluation reward: 5.38
 episode: 2326 score: 5.0 memory length: 513875 epsilon: 0.1805255200088332
 steps: 312 lr: 0.0001 evaluation reward: 5.4
 episode: 2327 score: 7.0 memory length: 514262 epsilon:
 0.17975926000882836 steps: 387 lr: 0.0001 evaluation reward: 5.39
 episode: 2328 score: 4.0 memory length: 514503 epsilon:
 0.17928208000882534 steps: 241 lr: 0.0001 evaluation reward: 5.36
 episode: 2329 score: 5.0 memory length: 514791 epsilon:
 0.17871184000882173 steps: 288 lr: 0.0001 evaluation reward: 5.37
 episode: 2330 score: 9.0 memory length: 515265 epsilon: 0.1777733200088158
 steps: 474 lr: 0.0001 evaluation reward: 5.37
 episode: 2331 score: 6.0 memory length: 515599 epsilon: 0.1771120000088116
 steps: 334 lr: 0.0001 evaluation reward: 5.36
 episode: 2332 score: 6.0 memory length: 515955 epsilon:
 0.17640712000880715 steps: 356 lr: 0.0001 evaluation reward: 5.37
 episode: 2333 score: 6.0 memory length: 516287 epsilon: 0.175749760008803
 steps: 332 lr: 0.0001 evaluation reward: 5.35
 episode: 2334 score: 13.0 memory length: 516921 epsilon:
 0.17449444000879505 steps: 634 lr: 0.0001 evaluation reward: 5.44
 episode: 2335 score: 4.0 memory length: 517200 epsilon:
 0.17394202000879155 steps: 279 lr: 0.0001 evaluation reward: 5.41
 episode: 2336 score: 10.0 memory length: 517707 epsilon:

0.1729381600087852 steps: 507 lr: 0.0001 evaluation reward: 5.49
 episode: 2337 score: 6.0 memory length: 518050 epsilon: 0.1722590200087809
 steps: 343 lr: 0.0001 evaluation reward: 5.49
 episode: 2338 score: 3.0 memory length: 518263 epsilon:
 0.17183728000877824 steps: 213 lr: 0.0001 evaluation reward: 5.46
 episode: 2339 score: 2.0 memory length: 518443 epsilon:
 0.17148088000877598 steps: 180 lr: 0.0001 evaluation reward: 5.43
 episode: 2340 score: 3.0 memory length: 518656 epsilon: 0.1710591400087733
 steps: 213 lr: 0.0001 evaluation reward: 5.42
 episode: 2341 score: 6.0 memory length: 519010 epsilon:
 0.17035822000876888 steps: 354 lr: 0.0001 evaluation reward: 5.4
 episode: 2342 score: 6.0 memory length: 519369 epsilon:
 0.16964740000876438 steps: 359 lr: 0.0001 evaluation reward: 5.43
 episode: 2343 score: 10.0 memory length: 519852 epsilon:
 0.16869106000875833 steps: 483 lr: 0.0001 evaluation reward: 5.5
 episode: 2344 score: 5.0 memory length: 520177 epsilon:
 0.16804756000875426 steps: 325 lr: 0.0001 evaluation reward: 5.46
 episode: 2345 score: 7.0 memory length: 520573 epsilon: 0.1672634800087493
 steps: 396 lr: 0.0001 evaluation reward: 5.46
 episode: 2346 score: 8.0 memory length: 521024 epsilon:
 0.16637050000874365 steps: 451 lr: 0.0001 evaluation reward: 5.5
 episode: 2347 score: 7.0 memory length: 521411 epsilon: 0.1656042400087388
 steps: 387 lr: 0.0001 evaluation reward: 5.53
 episode: 2348 score: 6.0 memory length: 521765 epsilon:
 0.16490332000873437 steps: 354 lr: 0.0001 evaluation reward: 5.56
 episode: 2349 score: 4.0 memory length: 522044 epsilon:
 0.16435090000873087 steps: 279 lr: 0.0001 evaluation reward: 5.53
 episode: 2350 score: 5.0 memory length: 522349 epsilon:
 0.16374700000872705 steps: 305 lr: 0.0001 evaluation reward: 5.52
 episode: 2351 score: 7.0 memory length: 522715 epsilon:
 0.16302232000872247 steps: 366 lr: 0.0001 evaluation reward: 5.56
 episode: 2352 score: 5.0 memory length: 523041 epsilon:
 0.16237684000871838 steps: 326 lr: 0.0001 evaluation reward: 5.56
 episode: 2353 score: 4.0 memory length: 523282 epsilon:
 0.16189966000871536 steps: 241 lr: 0.0001 evaluation reward: 5.56
 episode: 2354 score: 5.0 memory length: 523568 epsilon:
 0.16133338000871178 steps: 286 lr: 0.0001 evaluation reward: 5.57
 episode: 2355 score: 8.0 memory length: 524002 epsilon:
 0.16047406000870634 steps: 434 lr: 0.0001 evaluation reward: 5.6
 episode: 2356 score: 8.0 memory length: 524423 epsilon:
 0.15964048000870107 steps: 421 lr: 0.0001 evaluation reward: 5.63
 episode: 2357 score: 10.0 memory length: 524936 epsilon:
 0.15862474000869464 steps: 513 lr: 0.0001 evaluation reward: 5.68
 episode: 2358 score: 5.0 memory length: 525225 epsilon:
 0.15805252000869102 steps: 289 lr: 0.0001 evaluation reward: 5.67
 episode: 2359 score: 9.0 memory length: 525682 epsilon: 0.1571476600086853
 steps: 457 lr: 0.0001 evaluation reward: 5.71
 episode: 2360 score: 7.0 memory length: 526089 epsilon: 0.1563418000086802

```

steps: 407    lr: 0.0001    evaluation reward: 5.73
episode: 2361  score: 9.0    memory length: 526592    epsilon: 0.1553458600086739
steps: 503    lr: 0.0001    evaluation reward: 5.76
episode: 2362  score: 11.0   memory length: 527138    epsilon:
0.15426478000866706    steps: 546    lr: 0.0001    evaluation reward: 5.81
episode: 2363  score: 11.0   memory length: 527722    epsilon:
0.15310846000865974    steps: 584    lr: 0.0001    evaluation reward: 5.89
episode: 2364  score: 5.0    memory length: 528031    epsilon:
0.15249664000865587    steps: 309    lr: 0.0001    evaluation reward: 5.91
episode: 2365  score: 6.0    memory length: 528367    epsilon:
0.15183136000865166    steps: 336    lr: 0.0001    evaluation reward: 5.86
episode: 2366  score: 7.0    memory length: 528747    epsilon: 0.1510789600086469
steps: 380    lr: 0.0001    evaluation reward: 5.86
episode: 2367  score: 8.0    memory length: 529040    epsilon:
0.15049882000864323    steps: 293    lr: 0.0001    evaluation reward: 5.86
episode: 2368  score: 6.0    memory length: 529409    epsilon: 0.1497682000086386
steps: 369    lr: 0.0001    evaluation reward: 5.87
episode: 2369  score: 8.0    memory length: 529797    epsilon:
0.14899996000863375    steps: 388    lr: 0.0001    evaluation reward: 5.9
episode: 2370  score: 3.0    memory length: 530007    epsilon:
0.14858416000863112    steps: 210    lr: 0.0001    evaluation reward: 5.87
episode: 2371  score: 10.0   memory length: 530506    epsilon:
0.14759614000862487    steps: 499    lr: 0.0001    evaluation reward: 5.92
episode: 2372  score: 5.0    memory length: 530779    epsilon:
0.14705560000862145    steps: 273    lr: 0.0001    evaluation reward: 5.94
episode: 2373  score: 9.0    memory length: 531272    epsilon:
0.14607946000861527    steps: 493    lr: 0.0001    evaluation reward: 5.94
episode: 2374  score: 5.0    memory length: 531579    epsilon:
0.14547160000861142    steps: 307    lr: 0.0001    evaluation reward: 5.95
episode: 2375  score: 4.0    memory length: 531853    epsilon: 0.144929080008608
steps: 274    lr: 0.0001    evaluation reward: 5.96
episode: 2376  score: 9.0    memory length: 532277    epsilon:
0.14408956000860268    steps: 424    lr: 0.0001    evaluation reward: 5.98
episode: 2377  score: 7.0    memory length: 532664    epsilon:
0.14332330000859783    steps: 387    lr: 0.0001    evaluation reward: 5.99
episode: 2378  score: 5.0    memory length: 532971    epsilon:
0.14271544000859399    steps: 307    lr: 0.0001    evaluation reward: 5.98
episode: 2379  score: 7.0    memory length: 533394    epsilon: 0.1418779000085887
steps: 423    lr: 0.0001    evaluation reward: 6.0
episode: 2380  score: 6.0    memory length: 533743    epsilon:
0.14118688000858431    steps: 349    lr: 0.0001    evaluation reward: 5.99
episode: 2381  score: 3.0    memory length: 533989    epsilon:
0.14069980000858123    steps: 246    lr: 0.0001    evaluation reward: 5.96
episode: 2382  score: 9.0    memory length: 534458    epsilon:
0.13977118000857536    steps: 469    lr: 0.0001    evaluation reward: 5.97
episode: 2383  score: 9.0    memory length: 534932    epsilon:
0.13883266000856942    steps: 474    lr: 0.0001    evaluation reward: 6.03
episode: 2384  score: 3.0    memory length: 535158    epsilon: 0.1383851800085666

```



```

steps: 226    lr: 0.0001    evaluation reward: 6.0
episode: 2385  score: 8.0    memory length: 535565    epsilon: 0.1375793200085615
steps: 407    lr: 0.0001    evaluation reward: 6.03
episode: 2386  score: 5.0    memory length: 535852    epsilon: 0.1370110600085579
steps: 287    lr: 0.0001    evaluation reward: 6.03
episode: 2387  score: 13.0   memory length: 536363    epsilon:
0.1359992800085515    steps: 511    lr: 0.0001    evaluation reward: 6.13
episode: 2388  score: 5.0    memory length: 536677    epsilon:
0.13537756000854756    steps: 314    lr: 0.0001    evaluation reward: 6.15
episode: 2389  score: 7.0    memory length: 537052    epsilon:
0.13463506000854286    steps: 375    lr: 0.0001    evaluation reward: 6.18
episode: 2390  score: 5.0    memory length: 537360    epsilon: 0.134025220008539
steps: 308    lr: 0.0001    evaluation reward: 6.2
episode: 2391  score: 10.0   memory length: 537706    epsilon:
0.13334014000853467    steps: 346    lr: 0.0001    evaluation reward: 6.24
episode: 2392  score: 8.0    memory length: 538144    epsilon:
0.13247290000852918    steps: 438    lr: 0.0001    evaluation reward: 6.25
episode: 2393  score: 5.0    memory length: 538416    epsilon:
0.13193434000852577    steps: 272    lr: 0.0001    evaluation reward: 6.23
episode: 2394  score: 9.0    memory length: 538861    epsilon: 0.1310532400085202
steps: 445    lr: 0.0001    evaluation reward: 6.27
episode: 2395  score: 7.0    memory length: 539266    epsilon:
0.13025134000851513    steps: 405    lr: 0.0001    evaluation reward: 6.29
episode: 2396  score: 8.0    memory length: 539702    epsilon:
0.12938806000850966    steps: 436    lr: 0.0001    evaluation reward: 6.31
episode: 2397  score: 9.0    memory length: 540158    epsilon:
0.12848518000850395    steps: 456    lr: 0.0001    evaluation reward: 6.32
episode: 2398  score: 5.0    memory length: 540443    epsilon:
0.12792088000850038    steps: 285    lr: 0.0001    evaluation reward: 6.33
episode: 2399  score: 6.0    memory length: 540781    epsilon:
0.12725164000849615    steps: 338    lr: 0.0001    evaluation reward: 6.31
episode: 2400  score: 4.0    memory length: 541057    epsilon: 0.1267051600084927
steps: 276    lr: 0.0001    evaluation reward: 6.3
episode: 2401  score: 5.0    memory length: 541347    epsilon:
0.12613096000848906    steps: 290    lr: 0.0001    evaluation reward: 6.29
episode: 2402  score: 7.0    memory length: 541741    epsilon:
0.12535084000848412    steps: 394    lr: 0.0001    evaluation reward: 6.29
episode: 2403  score: 5.0    memory length: 542049    epsilon:
0.12474100000848208    steps: 308    lr: 0.0001    evaluation reward: 6.28
episode: 2404  score: 8.0    memory length: 542482    epsilon:
0.12388366000848267    steps: 433    lr: 0.0001    evaluation reward: 6.31
episode: 2405  score: 8.0    memory length: 542922    epsilon:
0.12301246000848326    steps: 440    lr: 0.0001    evaluation reward: 6.33
episode: 2406  score: 4.0    memory length: 543199    epsilon:
0.12246400000848363    steps: 277    lr: 0.0001    evaluation reward: 6.32
episode: 2407  score: 3.0    memory length: 543429    epsilon:
0.12200860000848394    steps: 230    lr: 0.0001    evaluation reward: 6.3
episode: 2408  score: 6.0    memory length: 543767    epsilon: 0.1213393600084844

```

```

steps: 338    lr: 0.0001    evaluation reward: 6.3
episode: 2409    score: 5.0    memory length: 544073    epsilon:
0.12073348000848481    steps: 306    lr: 0.0001    evaluation reward: 6.3
episode: 2410    score: 6.0    memory length: 544411    epsilon:
0.12006424000848527    steps: 338    lr: 0.0001    evaluation reward: 6.29
episode: 2411    score: 8.0    memory length: 544839    epsilon:
0.11921680000848585    steps: 428    lr: 0.0001    evaluation reward: 6.33
episode: 2412    score: 10.0    memory length: 545312    epsilon:
0.11828026000848649    steps: 473    lr: 0.0001    evaluation reward: 6.39
episode: 2413    score: 5.0    memory length: 545599    epsilon:
0.11771200000848688    steps: 287    lr: 0.0001    evaluation reward: 6.41
episode: 2414    score: 5.0    memory length: 545928    epsilon:
0.11706058000848732    steps: 329    lr: 0.0001    evaluation reward: 6.41
episode: 2415    score: 5.0    memory length: 546240    epsilon:
0.11644282000848774    steps: 312    lr: 0.0001    evaluation reward: 6.38
episode: 2416    score: 10.0    memory length: 546735    epsilon:
0.11546272000848841    steps: 495    lr: 0.0001    evaluation reward: 6.45
episode: 2417    score: 4.0    memory length: 546993    epsilon:
0.11495188000848876    steps: 258    lr: 0.0001    evaluation reward: 6.47
episode: 2418    score: 8.0    memory length: 547445    epsilon:
0.11405692000848937    steps: 452    lr: 0.0001    evaluation reward: 6.49
episode: 2419    score: 4.0    memory length: 547704    epsilon:
0.11354410000848972    steps: 259    lr: 0.0001    evaluation reward: 6.48
episode: 2420    score: 8.0    memory length: 548119    epsilon:
0.11272240000849028    steps: 415    lr: 0.0001    evaluation reward: 6.5
episode: 2421    score: 5.0    memory length: 548405    epsilon:
0.11215612000849066    steps: 286    lr: 0.0001    evaluation reward: 6.52
episode: 2422    score: 8.0    memory length: 548836    epsilon:
0.11130274000849125    steps: 431    lr: 0.0001    evaluation reward: 6.54
episode: 2423    score: 4.0    memory length: 549096    epsilon: 0.1107879400084916
steps: 260    lr: 0.0001    evaluation reward: 6.5
episode: 2424    score: 7.0    memory length: 549448    epsilon:
0.11009098000849207    steps: 352    lr: 0.0001    evaluation reward: 6.54
episode: 2425    score: 5.0    memory length: 549755    epsilon:
0.10948312000849249    steps: 307    lr: 0.0001    evaluation reward: 6.53
episode: 2426    score: 8.0    memory length: 550071    epsilon:
0.10885744000849291    steps: 316    lr: 0.0001    evaluation reward: 6.56
episode: 2427    score: 10.0    memory length: 550555    epsilon:
0.10789912000849357    steps: 484    lr: 0.0001    evaluation reward: 6.59
episode: 2428    score: 7.0    memory length: 550959    epsilon:
0.10709920000849411    steps: 404    lr: 0.0001    evaluation reward: 6.62
episode: 2429    score: 7.0    memory length: 551344    epsilon:
0.10633690000849463    steps: 385    lr: 0.0001    evaluation reward: 6.64
episode: 2430    score: 8.0    memory length: 551787    epsilon:
0.10545976000849523    steps: 443    lr: 0.0001    evaluation reward: 6.63
episode: 2431    score: 8.0    memory length: 552214    epsilon:
0.10461430000849581    steps: 427    lr: 0.0001    evaluation reward: 6.65
episode: 2432    score: 9.0    memory length: 552700    epsilon:

```

0.10365202000849646	steps: 486	lr: 0.0001	evaluation reward: 6.68
episode: 2433	score: 5.0	memory length: 553006	epsilon:
0.10304614000849688	steps: 306	lr: 0.0001	evaluation reward: 6.67
episode: 2434	score: 8.0	memory length: 553455	epsilon:
0.10215712000849748	steps: 449	lr: 0.0001	evaluation reward: 6.62
episode: 2435	score: 9.0	memory length: 553940	epsilon:
0.10119682000849814	steps: 485	lr: 0.0001	evaluation reward: 6.67
episode: 2436	score: 7.0	memory length: 554322	epsilon:
0.10044046000849866	steps: 382	lr: 0.0001	evaluation reward: 6.64
episode: 2437	score: 7.0	memory length: 554706	epsilon:
0.09968014000849917	steps: 384	lr: 0.0001	evaluation reward: 6.65
episode: 2438	score: 4.0	memory length: 554948	epsilon: 0.0992009800084995
steps: 242	lr: 0.0001	evaluation reward: 6.66	
episode: 2439	score: 6.0	memory length: 555306	epsilon:
0.09849214000849998	steps: 358	lr: 0.0001	evaluation reward: 6.7
episode: 2440	score: 8.0	memory length: 555742	epsilon:
0.09762886000850057	steps: 436	lr: 0.0001	evaluation reward: 6.75
episode: 2441	score: 3.0	memory length: 555953	epsilon:
0.09721108000850086	steps: 211	lr: 0.0001	evaluation reward: 6.72
episode: 2442	score: 9.0	memory length: 556406	epsilon:
0.09631414000850147	steps: 453	lr: 0.0001	evaluation reward: 6.75
episode: 2443	score: 8.0	memory length: 556845	epsilon:
0.09544492000850206	steps: 439	lr: 0.0001	evaluation reward: 6.73
episode: 2444	score: 7.0	memory length: 557254	epsilon:
0.09463510000850262	steps: 409	lr: 0.0001	evaluation reward: 6.75
episode: 2445	score: 13.0	memory length: 557880	epsilon:
0.09339562000850346	steps: 626	lr: 0.0001	evaluation reward: 6.81
episode: 2446	score: 8.0	memory length: 558331	epsilon:
0.09250264000850407	steps: 451	lr: 0.0001	evaluation reward: 6.81
episode: 2447	score: 8.0	memory length: 558759	epsilon:
0.09165520000850465	steps: 428	lr: 0.0001	evaluation reward: 6.82
episode: 2448	score: 7.0	memory length: 559126	epsilon:
0.09092854000850514	steps: 367	lr: 0.0001	evaluation reward: 6.83
episode: 2449	score: 3.0	memory length: 559338	epsilon:
0.09050878000850543	steps: 212	lr: 0.0001	evaluation reward: 6.82
episode: 2450	score: 7.0	memory length: 559745	epsilon:
0.08970292000850598	steps: 407	lr: 0.0001	evaluation reward: 6.84
episode: 2451	score: 5.0	memory length: 560073	epsilon:
0.08905348000850642	steps: 328	lr: 0.0001	evaluation reward: 6.82
episode: 2452	score: 6.0	memory length: 560444	epsilon:
0.08831890000850692	steps: 371	lr: 0.0001	evaluation reward: 6.83
episode: 2453	score: 6.0	memory length: 560799	epsilon: 0.0876160000085074
steps: 355	lr: 0.0001	evaluation reward: 6.85	
episode: 2454	score: 9.0	memory length: 561330	epsilon:
0.08656462000850812	steps: 531	lr: 0.0001	evaluation reward: 6.89
episode: 2455	score: 4.0	memory length: 561572	epsilon:
0.08608546000850845	steps: 242	lr: 0.0001	evaluation reward: 6.85
episode: 2456	score: 4.0	memory length: 561829	epsilon: 0.0855766000085088

```

steps: 257    lr: 0.0001    evaluation reward: 6.81
episode: 2457    score: 11.0    memory length: 562353    epsilon:
0.0845390800085095    steps: 524    lr: 0.0001    evaluation reward: 6.82
episode: 2458    score: 5.0    memory length: 562660    epsilon:
0.08393122000850992    steps: 307    lr: 0.0001    evaluation reward: 6.82
episode: 2459    score: 4.0    memory length: 562921    epsilon:
0.08341444000851027    steps: 261    lr: 0.0001    evaluation reward: 6.77
episode: 2460    score: 20.0    memory length: 563595    epsilon:
0.08207992000851118    steps: 674    lr: 0.0001    evaluation reward: 6.9
episode: 2461    score: 6.0    memory length: 563971    epsilon:
0.08133544000851169    steps: 376    lr: 0.0001    evaluation reward: 6.87
episode: 2462    score: 11.0    memory length: 564478    epsilon:
0.08033158000851237    steps: 507    lr: 0.0001    evaluation reward: 6.87
episode: 2463    score: 10.0    memory length: 564969    epsilon:
0.07935940000851303    steps: 491    lr: 0.0001    evaluation reward: 6.86
episode: 2464    score: 11.0    memory length: 565408    epsilon:
0.07849018000851363    steps: 439    lr: 0.0001    evaluation reward: 6.92
episode: 2465    score: 6.0    memory length: 565739    epsilon:
0.07783480000851407    steps: 331    lr: 0.0001    evaluation reward: 6.92
episode: 2466    score: 6.0    memory length: 566112    epsilon:
0.07709626000851458    steps: 373    lr: 0.0001    evaluation reward: 6.91
episode: 2467    score: 6.0    memory length: 566467    epsilon:
0.07639336000851506    steps: 355    lr: 0.0001    evaluation reward: 6.89
episode: 2468    score: 6.0    memory length: 566820    epsilon:
0.07569442000851553    steps: 353    lr: 0.0001    evaluation reward: 6.89
episode: 2469    score: 10.0    memory length: 567306    epsilon:
0.07473214000851619    steps: 486    lr: 0.0001    evaluation reward: 6.91
episode: 2470    score: 14.0    memory length: 567804    epsilon:
0.07374610000851686    steps: 498    lr: 0.0001    evaluation reward: 7.02
episode: 2471    score: 5.0    memory length: 568091    epsilon:
0.07317784000851725    steps: 287    lr: 0.0001    evaluation reward: 6.97
episode: 2472    score: 6.0    memory length: 568428    epsilon: 0.0725105800085177
steps: 337    lr: 0.0001    evaluation reward: 6.98
episode: 2473    score: 4.0    memory length: 568668    epsilon:
0.07203538000851803    steps: 240    lr: 0.0001    evaluation reward: 6.93
episode: 2474    score: 3.0    memory length: 568898    epsilon:
0.07157998000851834    steps: 230    lr: 0.0001    evaluation reward: 6.91
episode: 2475    score: 4.0    memory length: 569159    epsilon:
0.07106320000851869    steps: 261    lr: 0.0001    evaluation reward: 6.91
episode: 2476    score: 8.0    memory length: 569629    epsilon:
0.07013260000851933    steps: 470    lr: 0.0001    evaluation reward: 6.9
episode: 2477    score: 9.0    memory length: 570096    epsilon:
0.06920794000851996    steps: 467    lr: 0.0001    evaluation reward: 6.92
episode: 2478    score: 13.0    memory length: 570613    epsilon:
0.06818428000852066    steps: 517    lr: 0.0001    evaluation reward: 7.0
episode: 2479    score: 5.0    memory length: 570904    epsilon:
0.06760810000852105    steps: 291    lr: 0.0001    evaluation reward: 6.98
episode: 2480    score: 5.0    memory length: 571192    epsilon:

```

0.06703786000852144 steps: 288 lr: 0.0001 evaluation reward: 6.97
 episode: 2481 score: 6.0 memory length: 571548 epsilon:
 0.06633298000852192 steps: 356 lr: 0.0001 evaluation reward: 7.0
 episode: 2482 score: 5.0 memory length: 571854 epsilon:
 0.06572710000852233 steps: 306 lr: 0.0001 evaluation reward: 6.96
 episode: 2483 score: 2.0 memory length: 572054 epsilon: 0.0653311000085226
 steps: 200 lr: 0.0001 evaluation reward: 6.89
 episode: 2484 score: 8.0 memory length: 572505 epsilon:
 0.06443812000852321 steps: 451 lr: 0.0001 evaluation reward: 6.94
 episode: 2485 score: 7.0 memory length: 572911 epsilon:
 0.06363424000852376 steps: 406 lr: 0.0001 evaluation reward: 6.93
 episode: 2486 score: 3.0 memory length: 573119 epsilon:
 0.06322240000852404 steps: 208 lr: 0.0001 evaluation reward: 6.91
 episode: 2487 score: 9.0 memory length: 573590 epsilon:
 0.06228982000852468 steps: 471 lr: 0.0001 evaluation reward: 6.87
 episode: 2488 score: 8.0 memory length: 573984 epsilon:
 0.06150970000852521 steps: 394 lr: 0.0001 evaluation reward: 6.9
 episode: 2489 score: 7.0 memory length: 574367 epsilon:
 0.060751360008525726 steps: 383 lr: 0.0001 evaluation reward: 6.9
 episode: 2490 score: 4.0 memory length: 574625 epsilon:
 0.060240520008526074 steps: 258 lr: 0.0001 evaluation reward: 6.89
 episode: 2491 score: 6.0 memory length: 574941 epsilon: 0.0596148400085265
 steps: 316 lr: 0.0001 evaluation reward: 6.85
 episode: 2492 score: 5.0 memory length: 575236 epsilon: 0.0590307400085269
 steps: 295 lr: 0.0001 evaluation reward: 6.82
 episode: 2493 score: 5.0 memory length: 575541 epsilon:
 0.05842684000852731 steps: 305 lr: 0.0001 evaluation reward: 6.82
 episode: 2494 score: 9.0 memory length: 575981 epsilon:
 0.057555640008527906 steps: 440 lr: 0.0001 evaluation reward: 6.82
 episode: 2495 score: 7.0 memory length: 576387 epsilon:
 0.056751760008528454 steps: 406 lr: 0.0001 evaluation reward: 6.82
 episode: 2496 score: 4.0 memory length: 576666 epsilon:
 0.05619934000852883 steps: 279 lr: 0.0001 evaluation reward: 6.78
 episode: 2497 score: 9.0 memory length: 577084 epsilon:
 0.055371700008529395 steps: 418 lr: 0.0001 evaluation reward: 6.78
 episode: 2498 score: 7.0 memory length: 577508 epsilon:
 0.05453218000852997 steps: 424 lr: 0.0001 evaluation reward: 6.8
 episode: 2499 score: 4.0 memory length: 577768 epsilon:
 0.05401738000853032 steps: 260 lr: 0.0001 evaluation reward: 6.78
 episode: 2500 score: 9.0 memory length: 578217 epsilon:
 0.053128360008530925 steps: 449 lr: 0.0001 evaluation reward: 6.83
 episode: 2501 score: 6.0 memory length: 578539 epsilon:
 0.05249080000853136 steps: 322 lr: 0.0001 evaluation reward: 6.84
 episode: 2502 score: 5.0 memory length: 578846 epsilon:
 0.051882940008531775 steps: 307 lr: 0.0001 evaluation reward: 6.82
 episode: 2503 score: 8.0 memory length: 579298 epsilon:
 0.050987980008532385 steps: 452 lr: 0.0001 evaluation reward: 6.85
 episode: 2504 score: 5.0 memory length: 579620 epsilon:

0.05035042000853282	steps: 322	lr: 0.0001	evaluation reward: 6.82
episode: 2505	score: 9.0	memory length: 580089	epsilon:
0.049421800008533454	steps: 469	lr: 0.0001	evaluation reward: 6.83
episode: 2506	score: 5.0	memory length: 580396	epsilon:
0.04881394000853387	steps: 307	lr: 0.0001	evaluation reward: 6.84
episode: 2507	score: 6.0	memory length: 580734	epsilon:
0.048144700008534325	steps: 338	lr: 0.0001	evaluation reward: 6.87
episode: 2508	score: 7.0	memory length: 581118	epsilon:
0.04738438000853484	steps: 384	lr: 0.0001	evaluation reward: 6.88
episode: 2509	score: 7.0	memory length: 581485	epsilon:
0.04665772000853534	steps: 367	lr: 0.0001	evaluation reward: 6.9
episode: 2510	score: 8.0	memory length: 581942	epsilon:
0.045752860008535956	steps: 457	lr: 0.0001	evaluation reward: 6.92
episode: 2511	score: 4.0	memory length: 582220	epsilon:
0.04520242000853633	steps: 278	lr: 0.0001	evaluation reward: 6.88
episode: 2512	score: 12.0	memory length: 582648	epsilon:
0.04435498000853691	steps: 428	lr: 0.0001	evaluation reward: 6.9
episode: 2513	score: 6.0	memory length: 583022	epsilon:
0.043614460008537415	steps: 374	lr: 0.0001	evaluation reward: 6.91
episode: 2514	score: 10.0	memory length: 583526	epsilon:
0.042616540008538095	steps: 504	lr: 0.0001	evaluation reward: 6.96
episode: 2515	score: 8.0	memory length: 583997	epsilon:
0.04168396000853873	steps: 471	lr: 0.0001	evaluation reward: 6.99
episode: 2516	score: 8.0	memory length: 584382	epsilon:
0.04092166000853925	steps: 385	lr: 0.0001	evaluation reward: 6.97
episode: 2517	score: 4.0	memory length: 584656	epsilon:
0.04037914000853962	steps: 274	lr: 0.0001	evaluation reward: 6.97
episode: 2518	score: 7.0	memory length: 585058	epsilon:
0.039583180008540164	steps: 402	lr: 0.0001	evaluation reward: 6.96
episode: 2519	score: 6.0	memory length: 585378	epsilon:
0.038949580008540596	steps: 320	lr: 0.0001	evaluation reward: 6.98
episode: 2520	score: 6.0	memory length: 585732	epsilon:
0.038248660008541074	steps: 354	lr: 0.0001	evaluation reward: 6.96
episode: 2521	score: 4.0	memory length: 585988	epsilon:
0.03774178000854142	steps: 256	lr: 0.0001	evaluation reward: 6.95
episode: 2522	score: 5.0	memory length: 586274	epsilon:
0.037175500008541806	steps: 286	lr: 0.0001	evaluation reward: 6.92
episode: 2523	score: 8.0	memory length: 586682	epsilon:
0.03636766000854236	steps: 408	lr: 0.0001	evaluation reward: 6.96
episode: 2524	score: 15.0	memory length: 587220	epsilon:
0.035302420008543084	steps: 538	lr: 0.0001	evaluation reward: 7.04
episode: 2525	score: 3.0	memory length: 587448	epsilon:
0.03485098000854339	steps: 228	lr: 0.0001	evaluation reward: 7.02
episode: 2526	score: 9.0	memory length: 587935	epsilon:
0.03388672000854405	steps: 487	lr: 0.0001	evaluation reward: 7.03
episode: 2527	score: 4.0	memory length: 588179	epsilon:
0.03340360000854438	steps: 244	lr: 0.0001	evaluation reward: 6.97
episode: 2528	score: 5.0	memory length: 588482	epsilon:

0.03280366000854479 steps: 303 lr: 0.0001 evaluation reward: 6.95
 episode: 2529 score: 4.0 memory length: 588739 epsilon:
 0.032294800008545135 steps: 257 lr: 0.0001 evaluation reward: 6.92
 episode: 2530 score: 14.0 memory length: 589258 epsilon:
 0.031267180008545836 steps: 519 lr: 0.0001 evaluation reward: 6.98
 episode: 2531 score: 6.0 memory length: 589603 epsilon:
 0.030584080008546302 steps: 345 lr: 0.0001 evaluation reward: 6.96
 episode: 2532 score: 7.0 memory length: 590028 epsilon:
 0.029742580008546876 steps: 425 lr: 0.0001 evaluation reward: 6.94
 episode: 2533 score: 5.0 memory length: 590355 epsilon:
 0.029095120008547318 steps: 327 lr: 0.0001 evaluation reward: 6.94
 episode: 2534 score: 6.0 memory length: 590709 epsilon:
 0.028394200008547796 steps: 354 lr: 0.0001 evaluation reward: 6.92
 episode: 2535 score: 7.0 memory length: 591061 epsilon:
 0.02769724000854827 steps: 352 lr: 0.0001 evaluation reward: 6.9
 episode: 2536 score: 7.0 memory length: 591461 epsilon:
 0.02690524000854881 steps: 400 lr: 0.0001 evaluation reward: 6.9
 episode: 2537 score: 5.0 memory length: 591750 epsilon: 0.0263330200085492
 steps: 289 lr: 0.0001 evaluation reward: 6.88
 episode: 2538 score: 6.0 memory length: 592104 epsilon:
 0.02563210000854968 steps: 354 lr: 0.0001 evaluation reward: 6.9
 episode: 2539 score: 4.0 memory length: 592343 epsilon:
 0.025158880008550003 steps: 239 lr: 0.0001 evaluation reward: 6.88
 episode: 2540 score: 6.0 memory length: 592720 epsilon:
 0.02441242000855051 steps: 377 lr: 0.0001 evaluation reward: 6.86
 episode: 2541 score: 8.0 memory length: 593168 epsilon:
 0.023525380008551117 steps: 448 lr: 0.0001 evaluation reward: 6.91
 episode: 2542 score: 8.0 memory length: 593565 epsilon:
 0.022739320008551653 steps: 397 lr: 0.0001 evaluation reward: 6.9
 episode: 2543 score: 9.0 memory length: 594014 epsilon:
 0.02185030000855226 steps: 449 lr: 0.0001 evaluation reward: 6.91
 episode: 2544 score: 6.0 memory length: 594370 epsilon:
 0.02114542000855274 steps: 356 lr: 0.0001 evaluation reward: 6.9
 episode: 2545 score: 4.0 memory length: 594631 epsilon:
 0.020628640008553092 steps: 261 lr: 0.0001 evaluation reward: 6.81
 episode: 2546 score: 5.0 memory length: 594919 epsilon:
 0.02005840000855348 steps: 288 lr: 0.0001 evaluation reward: 6.78
 episode: 2547 score: 5.0 memory length: 595245 epsilon:
 0.01941292000855392 steps: 326 lr: 0.0001 evaluation reward: 6.75
 episode: 2548 score: 8.0 memory length: 595661 epsilon:
 0.018589240008554483 steps: 416 lr: 0.0001 evaluation reward: 6.76
 episode: 2549 score: 5.0 memory length: 595969 epsilon: 0.0179794000085549
 steps: 308 lr: 0.0001 evaluation reward: 6.78
 episode: 2550 score: 8.0 memory length: 596426 epsilon:
 0.017074540008555517 steps: 457 lr: 0.0001 evaluation reward: 6.79
 episode: 2551 score: 5.0 memory length: 596697 epsilon:
 0.016537960008555883 steps: 271 lr: 0.0001 evaluation reward: 6.79
 episode: 2552 score: 5.0 memory length: 596969 epsilon:

0.01599940000855625	steps: 272	lr: 0.0001	evaluation reward: 6.78
episode: 2553	score: 6.0	memory length: 597286	epsilon:
0.015371740008556456	steps: 317	lr: 0.0001	evaluation reward: 6.78
episode: 2554	score: 8.0	memory length: 597706	epsilon:
0.014540140008556295	steps: 420	lr: 0.0001	evaluation reward: 6.77
episode: 2555	score: 5.0	memory length: 598006	epsilon:
0.01394614000855618	steps: 300	lr: 0.0001	evaluation reward: 6.78
episode: 2556	score: 7.0	memory length: 598438	epsilon:
0.013090780008556013	steps: 432	lr: 0.0001	evaluation reward: 6.81
episode: 2557	score: 3.0	memory length: 598647	epsilon:
0.012676960008555933	steps: 209	lr: 0.0001	evaluation reward: 6.73
episode: 2558	score: 5.0	memory length: 598954	epsilon:
0.012069100008555815	steps: 307	lr: 0.0001	evaluation reward: 6.73
episode: 2559	score: 12.0	memory length: 599403	epsilon:
0.011180080008555643	steps: 449	lr: 0.0001	evaluation reward: 6.81
episode: 2560	score: 11.0	memory length: 599776	epsilon:
0.0104415400085555	steps: 373	lr: 0.0001	evaluation reward: 6.72
episode: 2561	score: 9.0	memory length: 600117	epsilon:
0.009998020008555413	steps: 341	lr: 0.0001	evaluation reward: 6.75
episode: 2562	score: 8.0	memory length: 600525	epsilon:
0.009998020008555413	steps: 408	lr: 0.0001	evaluation reward: 6.72
episode: 2563	score: 7.0	memory length: 600858	epsilon:
0.009998020008555413	steps: 333	lr: 0.0001	evaluation reward: 6.69
episode: 2564	score: 8.0	memory length: 601296	epsilon:
0.009998020008555413	steps: 438	lr: 0.0001	evaluation reward: 6.66
episode: 2565	score: 10.0	memory length: 601796	epsilon:
0.009998020008555413	steps: 500	lr: 0.0001	evaluation reward: 6.7
episode: 2566	score: 5.0	memory length: 602101	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 6.69
episode: 2567	score: 9.0	memory length: 602551	epsilon:
0.009998020008555413	steps: 450	lr: 0.0001	evaluation reward: 6.72
episode: 2568	score: 5.0	memory length: 602860	epsilon:
0.009998020008555413	steps: 309	lr: 0.0001	evaluation reward: 6.71
episode: 2569	score: 9.0	memory length: 603312	epsilon:
0.009998020008555413	steps: 452	lr: 0.0001	evaluation reward: 6.7
episode: 2570	score: 7.0	memory length: 603659	epsilon:
0.009998020008555413	steps: 347	lr: 0.0001	evaluation reward: 6.63
episode: 2571	score: 7.0	memory length: 604028	epsilon:
0.009998020008555413	steps: 369	lr: 0.0001	evaluation reward: 6.65
episode: 2572	score: 7.0	memory length: 604434	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 6.66
episode: 2573	score: 6.0	memory length: 604769	epsilon:
0.009998020008555413	steps: 335	lr: 0.0001	evaluation reward: 6.68
episode: 2574	score: 10.0	memory length: 605228	epsilon:
0.009998020008555413	steps: 459	lr: 0.0001	evaluation reward: 6.75
episode: 2575	score: 9.0	memory length: 605657	epsilon:
0.009998020008555413	steps: 429	lr: 0.0001	evaluation reward: 6.8
episode: 2576	score: 5.0	memory length: 605960	epsilon:

0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 6.77
episode: 2577	score: 8.0	memory length: 606404	epsilon:
0.009998020008555413	steps: 444	lr: 0.0001	evaluation reward: 6.76
episode: 2578	score: 10.0	memory length: 606960	epsilon:
0.009998020008555413	steps: 556	lr: 0.0001	evaluation reward: 6.73
episode: 2579	score: 9.0	memory length: 607405	epsilon:
0.009998020008555413	steps: 445	lr: 0.0001	evaluation reward: 6.77
episode: 2580	score: 7.0	memory length: 607792	epsilon:
0.009998020008555413	steps: 387	lr: 0.0001	evaluation reward: 6.79
episode: 2581	score: 8.0	memory length: 608196	epsilon:
0.009998020008555413	steps: 404	lr: 0.0001	evaluation reward: 6.81
episode: 2582	score: 11.0	memory length: 608705	epsilon:
0.009998020008555413	steps: 509	lr: 0.0001	evaluation reward: 6.87
episode: 2583	score: 5.0	memory length: 609029	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 6.9
episode: 2584	score: 10.0	memory length: 609513	epsilon:
0.009998020008555413	steps: 484	lr: 0.0001	evaluation reward: 6.92
episode: 2585	score: 10.0	memory length: 610026	epsilon:
0.009998020008555413	steps: 513	lr: 0.0001	evaluation reward: 6.95
episode: 2586	score: 6.0	memory length: 610383	epsilon:
0.009998020008555413	steps: 357	lr: 0.0001	evaluation reward: 6.98
episode: 2587	score: 7.0	memory length: 610797	epsilon:
0.009998020008555413	steps: 414	lr: 0.0001	evaluation reward: 6.96
episode: 2588	score: 5.0	memory length: 611083	epsilon:
0.009998020008555413	steps: 286	lr: 0.0001	evaluation reward: 6.93
episode: 2589	score: 8.0	memory length: 611516	epsilon:
0.009998020008555413	steps: 433	lr: 0.0001	evaluation reward: 6.94
episode: 2590	score: 10.0	memory length: 612045	epsilon:
0.009998020008555413	steps: 529	lr: 0.0001	evaluation reward: 7.0
episode: 2591	score: 8.0	memory length: 612485	epsilon:
0.009998020008555413	steps: 440	lr: 0.0001	evaluation reward: 7.02
episode: 2592	score: 9.0	memory length: 612914	epsilon:
0.009998020008555413	steps: 429	lr: 0.0001	evaluation reward: 7.06
episode: 2593	score: 9.0	memory length: 613397	epsilon:
0.009998020008555413	steps: 483	lr: 0.0001	evaluation reward: 7.1
episode: 2594	score: 8.0	memory length: 613798	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 7.09
episode: 2595	score: 9.0	memory length: 614270	epsilon:
0.009998020008555413	steps: 472	lr: 0.0001	evaluation reward: 7.11
episode: 2596	score: 7.0	memory length: 614677	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 7.14
episode: 2597	score: 7.0	memory length: 615049	epsilon:
0.009998020008555413	steps: 372	lr: 0.0001	evaluation reward: 7.12
episode: 2598	score: 4.0	memory length: 615329	epsilon:
0.009998020008555413	steps: 280	lr: 0.0001	evaluation reward: 7.09
episode: 2599	score: 4.0	memory length: 615604	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 7.09
episode: 2600	score: 9.0	memory length: 616110	epsilon:

0.009998020008555413	steps: 506	lr: 0.0001	evaluation reward: 7.09
episode: 2601	score: 12.0	memory length: 616677	epsilon:
0.009998020008555413	steps: 567	lr: 0.0001	evaluation reward: 7.15
episode: 2602	score: 5.0	memory length: 617004	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 7.15
episode: 2603	score: 12.0	memory length: 617561	epsilon:
0.009998020008555413	steps: 557	lr: 0.0001	evaluation reward: 7.19
episode: 2604	score: 9.0	memory length: 618007	epsilon:
0.009998020008555413	steps: 446	lr: 0.0001	evaluation reward: 7.23
episode: 2605	score: 8.0	memory length: 618426	epsilon:
0.009998020008555413	steps: 419	lr: 0.0001	evaluation reward: 7.22
episode: 2606	score: 14.0	memory length: 618809	epsilon:
0.009998020008555413	steps: 383	lr: 0.0001	evaluation reward: 7.31
episode: 2607	score: 4.0	memory length: 619051	epsilon:
0.009998020008555413	steps: 242	lr: 0.0001	evaluation reward: 7.29
episode: 2608	score: 6.0	memory length: 619410	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 7.28
episode: 2609	score: 8.0	memory length: 619864	epsilon:
0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 7.29
episode: 2610	score: 7.0	memory length: 620201	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 7.28
episode: 2611	score: 7.0	memory length: 620576	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 7.31
episode: 2612	score: 9.0	memory length: 621040	epsilon:
0.009998020008555413	steps: 464	lr: 0.0001	evaluation reward: 7.28
episode: 2613	score: 4.0	memory length: 621280	epsilon:
0.009998020008555413	steps: 240	lr: 0.0001	evaluation reward: 7.26
episode: 2614	score: 12.0	memory length: 621696	epsilon:
0.009998020008555413	steps: 416	lr: 0.0001	evaluation reward: 7.28
episode: 2615	score: 5.0	memory length: 621985	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 7.25
episode: 2616	score: 6.0	memory length: 622306	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 7.23
episode: 2617	score: 5.0	memory length: 622599	epsilon:
0.009998020008555413	steps: 293	lr: 0.0001	evaluation reward: 7.24
episode: 2618	score: 12.0	memory length: 623084	epsilon:
0.009998020008555413	steps: 485	lr: 0.0001	evaluation reward: 7.29
episode: 2619	score: 10.0	memory length: 623548	epsilon:
0.009998020008555413	steps: 464	lr: 0.0001	evaluation reward: 7.33
episode: 2620	score: 7.0	memory length: 623915	epsilon:
0.009998020008555413	steps: 367	lr: 0.0001	evaluation reward: 7.34
episode: 2621	score: 9.0	memory length: 624354	epsilon:
0.009998020008555413	steps: 439	lr: 0.0001	evaluation reward: 7.39
episode: 2622	score: 11.0	memory length: 624887	epsilon:
0.009998020008555413	steps: 533	lr: 0.0001	evaluation reward: 7.45
episode: 2623	score: 6.0	memory length: 625203	epsilon:
0.009998020008555413	steps: 316	lr: 0.0001	evaluation reward: 7.43
episode: 2624	score: 7.0	memory length: 625586	epsilon:

0.009998020008555413	steps: 383	lr: 0.0001	evaluation reward: 7.35
episode: 2625	score: 5.0	memory length: 625906	epsilon:
0.009998020008555413	steps: 320	lr: 0.0001	evaluation reward: 7.37
episode: 2626	score: 7.0	memory length: 626297	epsilon:
0.009998020008555413	steps: 391	lr: 0.0001	evaluation reward: 7.35
episode: 2627	score: 7.0	memory length: 626703	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 7.38
episode: 2628	score: 5.0	memory length: 626998	epsilon:
0.009998020008555413	steps: 295	lr: 0.0001	evaluation reward: 7.38
episode: 2629	score: 4.0	memory length: 627255	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 7.38
episode: 2630	score: 8.0	memory length: 627667	epsilon:
0.009998020008555413	steps: 412	lr: 0.0001	evaluation reward: 7.32
episode: 2631	score: 17.0	memory length: 628172	epsilon:
0.009998020008555413	steps: 505	lr: 0.0001	evaluation reward: 7.43
episode: 2632	score: 7.0	memory length: 628534	epsilon:
0.009998020008555413	steps: 362	lr: 0.0001	evaluation reward: 7.43
episode: 2633	score: 4.0	memory length: 628772	epsilon:
0.009998020008555413	steps: 238	lr: 0.0001	evaluation reward: 7.42
episode: 2634	score: 8.0	memory length: 629231	epsilon:
0.009998020008555413	steps: 459	lr: 0.0001	evaluation reward: 7.44
episode: 2635	score: 7.0	memory length: 629640	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 7.44
episode: 2636	score: 7.0	memory length: 630019	epsilon:
0.009998020008555413	steps: 379	lr: 0.0001	evaluation reward: 7.44
episode: 2637	score: 8.0	memory length: 630436	epsilon:
0.009998020008555413	steps: 417	lr: 0.0001	evaluation reward: 7.47
episode: 2638	score: 10.0	memory length: 630946	epsilon:
0.009998020008555413	steps: 510	lr: 0.0001	evaluation reward: 7.51
episode: 2639	score: 12.0	memory length: 631502	epsilon:
0.009998020008555413	steps: 556	lr: 0.0001	evaluation reward: 7.59
episode: 2640	score: 5.0	memory length: 631846	epsilon:
0.009998020008555413	steps: 344	lr: 0.0001	evaluation reward: 7.58
episode: 2641	score: 6.0	memory length: 632190	epsilon:
0.009998020008555413	steps: 344	lr: 0.0001	evaluation reward: 7.56
episode: 2642	score: 8.0	memory length: 632644	epsilon:
0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 7.56
episode: 2643	score: 6.0	memory length: 632998	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 7.53
episode: 2644	score: 7.0	memory length: 633346	epsilon:
0.009998020008555413	steps: 348	lr: 0.0001	evaluation reward: 7.54
episode: 2645	score: 7.0	memory length: 633735	epsilon:
0.009998020008555413	steps: 389	lr: 0.0001	evaluation reward: 7.57
episode: 2646	score: 5.0	memory length: 634020	epsilon:
0.009998020008555413	steps: 285	lr: 0.0001	evaluation reward: 7.57
episode: 2647	score: 7.0	memory length: 634420	epsilon:
0.009998020008555413	steps: 400	lr: 0.0001	evaluation reward: 7.59
episode: 2648	score: 7.0	memory length: 634808	epsilon:

0.009998020008555413	steps: 388	lr: 0.0001	evaluation reward: 7.58
episode: 2649	score: 12.0	memory length: 635213	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 7.65
episode: 2650	score: 8.0	memory length: 635649	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 7.65
episode: 2651	score: 6.0	memory length: 636025	epsilon:
0.009998020008555413	steps: 376	lr: 0.0001	evaluation reward: 7.66
episode: 2652	score: 4.0	memory length: 636282	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 7.65
episode: 2653	score: 5.0	memory length: 636553	epsilon:
0.009998020008555413	steps: 271	lr: 0.0001	evaluation reward: 7.64
episode: 2654	score: 11.0	memory length: 637067	epsilon:
0.009998020008555413	steps: 514	lr: 0.0001	evaluation reward: 7.67
episode: 2655	score: 11.0	memory length: 637599	epsilon:
0.009998020008555413	steps: 532	lr: 0.0001	evaluation reward: 7.73
episode: 2656	score: 10.0	memory length: 638097	epsilon:
0.009998020008555413	steps: 498	lr: 0.0001	evaluation reward: 7.76
episode: 2657	score: 8.0	memory length: 638529	epsilon:
0.009998020008555413	steps: 432	lr: 0.0001	evaluation reward: 7.81
episode: 2658	score: 7.0	memory length: 638911	epsilon:
0.009998020008555413	steps: 382	lr: 0.0001	evaluation reward: 7.83
episode: 2659	score: 17.0	memory length: 639459	epsilon:
0.009998020008555413	steps: 548	lr: 0.0001	evaluation reward: 7.88
episode: 2660	score: 7.0	memory length: 639862	epsilon:
0.009998020008555413	steps: 403	lr: 0.0001	evaluation reward: 7.84
episode: 2661	score: 17.0	memory length: 640363	epsilon:
0.009998020008555413	steps: 501	lr: 0.0001	evaluation reward: 7.92
episode: 2662	score: 7.0	memory length: 640752	epsilon:
0.009998020008555413	steps: 389	lr: 0.0001	evaluation reward: 7.91
episode: 2663	score: 8.0	memory length: 641159	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 7.92
episode: 2664	score: 9.0	memory length: 641661	epsilon:
0.009998020008555413	steps: 502	lr: 0.0001	evaluation reward: 7.93
episode: 2665	score: 8.0	memory length: 642083	epsilon:
0.009998020008555413	steps: 422	lr: 0.0001	evaluation reward: 7.91
episode: 2666	score: 11.0	memory length: 642666	epsilon:
0.009998020008555413	steps: 583	lr: 0.0001	evaluation reward: 7.97
episode: 2667	score: 8.0	memory length: 643121	epsilon:
0.009998020008555413	steps: 455	lr: 0.0001	evaluation reward: 7.96
episode: 2668	score: 8.0	memory length: 643554	epsilon:
0.009998020008555413	steps: 433	lr: 0.0001	evaluation reward: 7.99
episode: 2669	score: 13.0	memory length: 644054	epsilon:
0.009998020008555413	steps: 500	lr: 0.0001	evaluation reward: 8.03
episode: 2670	score: 8.0	memory length: 644479	epsilon:
0.009998020008555413	steps: 425	lr: 0.0001	evaluation reward: 8.04
episode: 2671	score: 6.0	memory length: 644830	epsilon:
0.009998020008555413	steps: 351	lr: 0.0001	evaluation reward: 8.03
episode: 2672	score: 5.0	memory length: 645118	epsilon:

0.009998020008555413	steps: 288	lr: 0.0001	evaluation reward: 8.01
episode: 2673	score: 7.0	memory length: 645504	epsilon:
0.009998020008555413	steps: 386	lr: 0.0001	evaluation reward: 8.02
episode: 2674	score: 3.0	memory length: 645716	epsilon:
0.009998020008555413	steps: 212	lr: 0.0001	evaluation reward: 7.95
episode: 2675	score: 7.0	memory length: 646071	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 7.93
episode: 2676	score: 7.0	memory length: 646452	epsilon:
0.009998020008555413	steps: 381	lr: 0.0001	evaluation reward: 7.95
episode: 2677	score: 8.0	memory length: 646896	epsilon:
0.009998020008555413	steps: 444	lr: 0.0001	evaluation reward: 7.95
episode: 2678	score: 10.0	memory length: 647394	epsilon:
0.009998020008555413	steps: 498	lr: 0.0001	evaluation reward: 7.95
episode: 2679	score: 9.0	memory length: 647737	epsilon:
0.009998020008555413	steps: 343	lr: 0.0001	evaluation reward: 7.95
episode: 2680	score: 5.0	memory length: 648066	epsilon:
0.009998020008555413	steps: 329	lr: 0.0001	evaluation reward: 7.93
episode: 2681	score: 5.0	memory length: 648395	epsilon:
0.009998020008555413	steps: 329	lr: 0.0001	evaluation reward: 7.9
episode: 2682	score: 9.0	memory length: 648865	epsilon:
0.009998020008555413	steps: 470	lr: 0.0001	evaluation reward: 7.88
episode: 2683	score: 7.0	memory length: 649238	epsilon:
0.009998020008555413	steps: 373	lr: 0.0001	evaluation reward: 7.9
episode: 2684	score: 4.0	memory length: 649495	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 7.84
episode: 2685	score: 6.0	memory length: 649847	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 7.8
episode: 2686	score: 8.0	memory length: 650281	epsilon:
0.009998020008555413	steps: 434	lr: 0.0001	evaluation reward: 7.82
episode: 2687	score: 13.0	memory length: 650760	epsilon:
0.009998020008555413	steps: 479	lr: 0.0001	evaluation reward: 7.88
episode: 2688	score: 6.0	memory length: 651060	epsilon:
0.009998020008555413	steps: 300	lr: 0.0001	evaluation reward: 7.89
episode: 2689	score: 14.0	memory length: 651636	epsilon:
0.009998020008555413	steps: 576	lr: 0.0001	evaluation reward: 7.95
episode: 2690	score: 7.0	memory length: 652081	epsilon:
0.009998020008555413	steps: 445	lr: 0.0001	evaluation reward: 7.92
episode: 2691	score: 5.0	memory length: 652388	epsilon:
0.009998020008555413	steps: 307	lr: 0.0001	evaluation reward: 7.89
episode: 2692	score: 5.0	memory length: 652696	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 7.85
episode: 2693	score: 5.0	memory length: 653000	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 7.81
episode: 2694	score: 7.0	memory length: 653419	epsilon:
0.009998020008555413	steps: 419	lr: 0.0001	evaluation reward: 7.8
episode: 2695	score: 4.0	memory length: 653679	epsilon:
0.009998020008555413	steps: 260	lr: 0.0001	evaluation reward: 7.75
episode: 2696	score: 13.0	memory length: 654194	epsilon:

0.009998020008555413	steps: 515	lr: 0.0001	evaluation reward: 7.81
episode: 2697	score: 6.0	memory length: 654569	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 7.8
episode: 2698	score: 6.0	memory length: 654924	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 7.82
episode: 2699	score: 5.0	memory length: 655226	epsilon:
0.009998020008555413	steps: 302	lr: 0.0001	evaluation reward: 7.83
episode: 2700	score: 5.0	memory length: 655548	epsilon:
0.009998020008555413	steps: 322	lr: 0.0001	evaluation reward: 7.79
episode: 2701	score: 7.0	memory length: 655942	epsilon:
0.009998020008555413	steps: 394	lr: 0.0001	evaluation reward: 7.74
episode: 2702	score: 5.0	memory length: 656244	epsilon:
0.009998020008555413	steps: 302	lr: 0.0001	evaluation reward: 7.74
episode: 2703	score: 17.0	memory length: 656732	epsilon:
0.009998020008555413	steps: 488	lr: 0.0001	evaluation reward: 7.79
episode: 2704	score: 11.0	memory length: 657270	epsilon:
0.009998020008555413	steps: 538	lr: 0.0001	evaluation reward: 7.81
episode: 2705	score: 5.0	memory length: 657575	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 7.78
episode: 2706	score: 5.0	memory length: 657900	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 7.69
episode: 2707	score: 7.0	memory length: 658303	epsilon:
0.009998020008555413	steps: 403	lr: 0.0001	evaluation reward: 7.72
episode: 2708	score: 5.0	memory length: 658630	epsilon:
0.009998020008555413	steps: 327	lr: 0.0001	evaluation reward: 7.71
episode: 2709	score: 6.0	memory length: 658948	epsilon:
0.009998020008555413	steps: 318	lr: 0.0001	evaluation reward: 7.69
episode: 2710	score: 7.0	memory length: 659334	epsilon:
0.009998020008555413	steps: 386	lr: 0.0001	evaluation reward: 7.69
episode: 2711	score: 12.0	memory length: 659884	epsilon:
0.009998020008555413	steps: 550	lr: 0.0001	evaluation reward: 7.74
episode: 2712	score: 5.0	memory length: 660171	epsilon:
0.009998020008555413	steps: 287	lr: 0.0001	evaluation reward: 7.7
episode: 2713	score: 7.0	memory length: 660591	epsilon:
0.009998020008555413	steps: 420	lr: 0.0001	evaluation reward: 7.73
episode: 2714	score: 4.0	memory length: 660831	epsilon:
0.009998020008555413	steps: 240	lr: 0.0001	evaluation reward: 7.65
episode: 2715	score: 4.0	memory length: 661070	epsilon:
0.009998020008555413	steps: 239	lr: 0.0001	evaluation reward: 7.64
episode: 2716	score: 8.0	memory length: 661517	epsilon:
0.009998020008555413	steps: 447	lr: 0.0001	evaluation reward: 7.66
episode: 2717	score: 6.0	memory length: 661821	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 7.67
episode: 2718	score: 3.0	memory length: 662049	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 7.58
episode: 2719	score: 5.0	memory length: 662335	epsilon:
0.009998020008555413	steps: 286	lr: 0.0001	evaluation reward: 7.53
episode: 2720	score: 7.0	memory length: 662693	epsilon:

0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 7.53
episode: 2721	score: 8.0	memory length: 663109	epsilon:
0.009998020008555413	steps: 416	lr: 0.0001	evaluation reward: 7.52
episode: 2722	score: 11.0	memory length: 663547	epsilon:
0.009998020008555413	steps: 438	lr: 0.0001	evaluation reward: 7.52
episode: 2723	score: 7.0	memory length: 663931	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 7.53
episode: 2724	score: 9.0	memory length: 664388	epsilon:
0.009998020008555413	steps: 457	lr: 0.0001	evaluation reward: 7.55
episode: 2725	score: 7.0	memory length: 664797	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 7.57
episode: 2726	score: 4.0	memory length: 665072	epsilon:
0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 7.54
episode: 2727	score: 8.0	memory length: 665499	epsilon:
0.009998020008555413	steps: 427	lr: 0.0001	evaluation reward: 7.55
episode: 2728	score: 9.0	memory length: 665934	epsilon:
0.009998020008555413	steps: 435	lr: 0.0001	evaluation reward: 7.59
episode: 2729	score: 9.0	memory length: 666277	epsilon:
0.009998020008555413	steps: 343	lr: 0.0001	evaluation reward: 7.64
episode: 2730	score: 5.0	memory length: 666598	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 7.61
episode: 2731	score: 16.0	memory length: 667152	epsilon:
0.009998020008555413	steps: 554	lr: 0.0001	evaluation reward: 7.6
episode: 2732	score: 6.0	memory length: 667495	epsilon:
0.009998020008555413	steps: 343	lr: 0.0001	evaluation reward: 7.59
episode: 2733	score: 6.0	memory length: 667851	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 7.61
episode: 2734	score: 9.0	memory length: 668351	epsilon:
0.009998020008555413	steps: 500	lr: 0.0001	evaluation reward: 7.62
episode: 2735	score: 9.0	memory length: 668820	epsilon:
0.009998020008555413	steps: 469	lr: 0.0001	evaluation reward: 7.64
episode: 2736	score: 8.0	memory length: 669224	epsilon:
0.009998020008555413	steps: 404	lr: 0.0001	evaluation reward: 7.65
episode: 2737	score: 6.0	memory length: 669557	epsilon:
0.009998020008555413	steps: 333	lr: 0.0001	evaluation reward: 7.63
episode: 2738	score: 11.0	memory length: 670120	epsilon:
0.009998020008555413	steps: 563	lr: 0.0001	evaluation reward: 7.64
episode: 2739	score: 4.0	memory length: 670418	epsilon:
0.009998020008555413	steps: 298	lr: 0.0001	evaluation reward: 7.56
episode: 2740	score: 18.0	memory length: 670847	epsilon:
0.009998020008555413	steps: 429	lr: 0.0001	evaluation reward: 7.69
episode: 2741	score: 9.0	memory length: 671190	epsilon:
0.009998020008555413	steps: 343	lr: 0.0001	evaluation reward: 7.72
episode: 2742	score: 9.0	memory length: 671663	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 7.73
episode: 2743	score: 5.0	memory length: 671971	epsilon:
0.009998020008555413	steps: 308	lr: 0.0001	evaluation reward: 7.72
episode: 2744	score: 8.0	memory length: 672395	epsilon:

0.009998020008555413	steps: 424	lr: 0.0001	evaluation reward: 7.73
episode: 2745	score: 6.0	memory length: 672752	epsilon:
0.009998020008555413	steps: 357	lr: 0.0001	evaluation reward: 7.72
episode: 2746	score: 7.0	memory length: 673106	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 7.74
episode: 2747	score: 7.0	memory length: 673485	epsilon:
0.009998020008555413	steps: 379	lr: 0.0001	evaluation reward: 7.74
episode: 2748	score: 10.0	memory length: 673970	epsilon:
0.009998020008555413	steps: 485	lr: 0.0001	evaluation reward: 7.77
episode: 2749	score: 7.0	memory length: 674358	epsilon:
0.009998020008555413	steps: 388	lr: 0.0001	evaluation reward: 7.72
episode: 2750	score: 6.0	memory length: 674712	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 7.7
episode: 2751	score: 6.0	memory length: 675045	epsilon:
0.009998020008555413	steps: 333	lr: 0.0001	evaluation reward: 7.7
episode: 2752	score: 11.0	memory length: 675435	epsilon:
0.009998020008555413	steps: 390	lr: 0.0001	evaluation reward: 7.77
episode: 2753	score: 7.0	memory length: 675816	epsilon:
0.009998020008555413	steps: 381	lr: 0.0001	evaluation reward: 7.79
episode: 2754	score: 9.0	memory length: 676255	epsilon:
0.009998020008555413	steps: 439	lr: 0.0001	evaluation reward: 7.77
episode: 2755	score: 6.0	memory length: 676618	epsilon:
0.009998020008555413	steps: 363	lr: 0.0001	evaluation reward: 7.72
episode: 2756	score: 7.0	memory length: 677003	epsilon:
0.009998020008555413	steps: 385	lr: 0.0001	evaluation reward: 7.69
episode: 2757	score: 13.0	memory length: 677480	epsilon:
0.009998020008555413	steps: 477	lr: 0.0001	evaluation reward: 7.74
episode: 2758	score: 6.0	memory length: 677837	epsilon:
0.009998020008555413	steps: 357	lr: 0.0001	evaluation reward: 7.73
episode: 2759	score: 10.0	memory length: 678333	epsilon:
0.009998020008555413	steps: 496	lr: 0.0001	evaluation reward: 7.66
episode: 2760	score: 9.0	memory length: 678826	epsilon:
0.009998020008555413	steps: 493	lr: 0.0001	evaluation reward: 7.68
episode: 2761	score: 7.0	memory length: 679070	epsilon:
0.009998020008555413	steps: 244	lr: 0.0001	evaluation reward: 7.58
episode: 2762	score: 6.0	memory length: 679407	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 7.57
episode: 2763	score: 7.0	memory length: 679781	epsilon:
0.009998020008555413	steps: 374	lr: 0.0001	evaluation reward: 7.56
episode: 2764	score: 6.0	memory length: 680120	epsilon:
0.009998020008555413	steps: 339	lr: 0.0001	evaluation reward: 7.53
episode: 2765	score: 7.0	memory length: 680486	epsilon:
0.009998020008555413	steps: 366	lr: 0.0001	evaluation reward: 7.52
episode: 2766	score: 9.0	memory length: 680923	epsilon:
0.009998020008555413	steps: 437	lr: 0.0001	evaluation reward: 7.5
episode: 2767	score: 7.0	memory length: 681347	epsilon:
0.009998020008555413	steps: 424	lr: 0.0001	evaluation reward: 7.49
episode: 2768	score: 8.0	memory length: 681761	epsilon:

0.009998020008555413	steps: 414	lr: 0.0001	evaluation reward: 7.49
episode: 2769	score: 6.0	memory length: 682101	epsilon:
0.009998020008555413	steps: 340	lr: 0.0001	evaluation reward: 7.42
episode: 2770	score: 9.0	memory length: 682594	epsilon:
0.009998020008555413	steps: 493	lr: 0.0001	evaluation reward: 7.43
episode: 2771	score: 5.0	memory length: 682885	epsilon:
0.009998020008555413	steps: 291	lr: 0.0001	evaluation reward: 7.42
episode: 2772	score: 3.0	memory length: 683113	epsilon:
0.009998020008555413	steps: 228	lr: 0.0001	evaluation reward: 7.4
episode: 2773	score: 5.0	memory length: 683401	epsilon:
0.009998020008555413	steps: 288	lr: 0.0001	evaluation reward: 7.38
episode: 2774	score: 8.0	memory length: 683818	epsilon:
0.009998020008555413	steps: 417	lr: 0.0001	evaluation reward: 7.43
episode: 2775	score: 7.0	memory length: 684173	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 7.43
episode: 2776	score: 6.0	memory length: 684507	epsilon:
0.009998020008555413	steps: 334	lr: 0.0001	evaluation reward: 7.42
episode: 2777	score: 9.0	memory length: 684952	epsilon:
0.009998020008555413	steps: 445	lr: 0.0001	evaluation reward: 7.43
episode: 2778	score: 11.0	memory length: 685472	epsilon:
0.009998020008555413	steps: 520	lr: 0.0001	evaluation reward: 7.44
episode: 2779	score: 16.0	memory length: 686110	epsilon:
0.009998020008555413	steps: 638	lr: 0.0001	evaluation reward: 7.51
episode: 2780	score: 8.0	memory length: 686578	epsilon:
0.009998020008555413	steps: 468	lr: 0.0001	evaluation reward: 7.54
episode: 2781	score: 5.0	memory length: 686867	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 7.54
episode: 2782	score: 11.0	memory length: 687254	epsilon:
0.009998020008555413	steps: 387	lr: 0.0001	evaluation reward: 7.56
episode: 2783	score: 6.0	memory length: 687609	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 7.55
episode: 2784	score: 6.0	memory length: 687948	epsilon:
0.009998020008555413	steps: 339	lr: 0.0001	evaluation reward: 7.57
episode: 2785	score: 9.0	memory length: 688413	epsilon:
0.009998020008555413	steps: 465	lr: 0.0001	evaluation reward: 7.6
episode: 2786	score: 5.0	memory length: 688698	epsilon:
0.009998020008555413	steps: 285	lr: 0.0001	evaluation reward: 7.57
episode: 2787	score: 11.0	memory length: 689103	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 7.55
episode: 2788	score: 9.0	memory length: 689561	epsilon:
0.009998020008555413	steps: 458	lr: 0.0001	evaluation reward: 7.58
episode: 2789	score: 7.0	memory length: 689966	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 7.51
episode: 2790	score: 7.0	memory length: 690389	epsilon:
0.009998020008555413	steps: 423	lr: 0.0001	evaluation reward: 7.51
episode: 2791	score: 12.0	memory length: 690974	epsilon:
0.009998020008555413	steps: 585	lr: 0.0001	evaluation reward: 7.58
episode: 2792	score: 4.0	memory length: 691249	epsilon:

0.009998020008555413	steps: 275	lr: 0.0001	evaluation reward: 7.57
episode: 2793	score: 3.0	memory length: 691462	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 7.55
episode: 2794	score: 10.0	memory length: 691975	epsilon:
0.009998020008555413	steps: 513	lr: 0.0001	evaluation reward: 7.58
episode: 2795	score: 10.0	memory length: 692448	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 7.64
episode: 2796	score: 4.0	memory length: 692707	epsilon:
0.009998020008555413	steps: 259	lr: 0.0001	evaluation reward: 7.55
episode: 2797	score: 8.0	memory length: 693180	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 7.57
episode: 2798	score: 5.0	memory length: 693486	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 7.56
episode: 2799	score: 5.0	memory length: 693792	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 7.56
episode: 2800	score: 7.0	memory length: 694213	epsilon:
0.009998020008555413	steps: 421	lr: 0.0001	evaluation reward: 7.58
episode: 2801	score: 8.0	memory length: 694625	epsilon:
0.009998020008555413	steps: 412	lr: 0.0001	evaluation reward: 7.59
episode: 2802	score: 7.0	memory length: 695039	epsilon:
0.009998020008555413	steps: 414	lr: 0.0001	evaluation reward: 7.61
episode: 2803	score: 10.0	memory length: 695531	epsilon:
0.009998020008555413	steps: 492	lr: 0.0001	evaluation reward: 7.54
episode: 2804	score: 6.0	memory length: 695869	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 7.49
episode: 2805	score: 4.0	memory length: 696131	epsilon:
0.009998020008555413	steps: 262	lr: 0.0001	evaluation reward: 7.48
episode: 2806	score: 6.0	memory length: 696486	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 7.49
episode: 2807	score: 11.0	memory length: 696960	epsilon:
0.009998020008555413	steps: 474	lr: 0.0001	evaluation reward: 7.53
episode: 2808	score: 6.0	memory length: 697315	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 7.54
episode: 2809	score: 8.0	memory length: 697742	epsilon:
0.009998020008555413	steps: 427	lr: 0.0001	evaluation reward: 7.56
episode: 2810	score: 11.0	memory length: 698249	epsilon:
0.009998020008555413	steps: 507	lr: 0.0001	evaluation reward: 7.6
episode: 2811	score: 9.0	memory length: 698716	epsilon:
0.009998020008555413	steps: 467	lr: 0.0001	evaluation reward: 7.57
episode: 2812	score: 5.0	memory length: 699006	epsilon:
0.009998020008555413	steps: 290	lr: 0.0001	evaluation reward: 7.57
episode: 2813	score: 10.0	memory length: 699509	epsilon:
0.009998020008555413	steps: 503	lr: 0.0001	evaluation reward: 7.6
episode: 2814	score: 3.0	memory length: 699719	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 7.59
episode: 2815	score: 18.0	memory length: 700368	epsilon:
0.009998020008555413	steps: 649	lr: 0.0001	evaluation reward: 7.73
episode: 2816	score: 5.0	memory length: 700691	epsilon:

0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 7.7
episode: 2817	score: 10.0	memory length: 701212	epsilon:
0.009998020008555413	steps: 521	lr: 0.0001	evaluation reward: 7.74
episode: 2818	score: 10.0	memory length: 701681	epsilon:
0.009998020008555413	steps: 469	lr: 0.0001	evaluation reward: 7.81
episode: 2819	score: 6.0	memory length: 702015	epsilon:
0.009998020008555413	steps: 334	lr: 0.0001	evaluation reward: 7.82
episode: 2820	score: 6.0	memory length: 702369	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 7.81
episode: 2821	score: 12.0	memory length: 702930	epsilon:
0.009998020008555413	steps: 561	lr: 0.0001	evaluation reward: 7.85
episode: 2822	score: 8.0	memory length: 703335	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 7.82
episode: 2823	score: 12.0	memory length: 703829	epsilon:
0.009998020008555413	steps: 494	lr: 0.0001	evaluation reward: 7.87
episode: 2824	score: 8.0	memory length: 704304	epsilon:
0.009998020008555413	steps: 475	lr: 0.0001	evaluation reward: 7.86
episode: 2825	score: 11.0	memory length: 704748	epsilon:
0.009998020008555413	steps: 444	lr: 0.0001	evaluation reward: 7.9
episode: 2826	score: 7.0	memory length: 705171	epsilon:
0.009998020008555413	steps: 423	lr: 0.0001	evaluation reward: 7.93
episode: 2827	score: 5.0	memory length: 705480	epsilon:
0.009998020008555413	steps: 309	lr: 0.0001	evaluation reward: 7.9
episode: 2828	score: 9.0	memory length: 705952	epsilon:
0.009998020008555413	steps: 472	lr: 0.0001	evaluation reward: 7.9
episode: 2829	score: 8.0	memory length: 706383	epsilon:
0.009998020008555413	steps: 431	lr: 0.0001	evaluation reward: 7.89
episode: 2830	score: 6.0	memory length: 706722	epsilon:
0.009998020008555413	steps: 339	lr: 0.0001	evaluation reward: 7.9
episode: 2831	score: 6.0	memory length: 707075	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 7.8
episode: 2832	score: 8.0	memory length: 707527	epsilon:
0.009998020008555413	steps: 452	lr: 0.0001	evaluation reward: 7.82
episode: 2833	score: 11.0	memory length: 708053	epsilon:
0.009998020008555413	steps: 526	lr: 0.0001	evaluation reward: 7.87
episode: 2834	score: 11.0	memory length: 708424	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 7.89
episode: 2835	score: 7.0	memory length: 708805	epsilon:
0.009998020008555413	steps: 381	lr: 0.0001	evaluation reward: 7.87
episode: 2836	score: 12.0	memory length: 709418	epsilon:
0.009998020008555413	steps: 613	lr: 0.0001	evaluation reward: 7.91
episode: 2837	score: 6.0	memory length: 709770	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 7.91
episode: 2838	score: 10.0	memory length: 710257	epsilon:
0.009998020008555413	steps: 487	lr: 0.0001	evaluation reward: 7.9
episode: 2839	score: 11.0	memory length: 710682	epsilon:
0.009998020008555413	steps: 425	lr: 0.0001	evaluation reward: 7.97
episode: 2840	score: 13.0	memory length: 711220	epsilon:

0.009998020008555413	steps: 538	lr: 0.0001	evaluation reward: 7.92
episode: 2841	score: 5.0	memory length: 711526	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 7.88
episode: 2842	score: 7.0	memory length: 711898	epsilon:
0.009998020008555413	steps: 372	lr: 0.0001	evaluation reward: 7.86
episode: 2843	score: 8.0	memory length: 712330	epsilon:
0.009998020008555413	steps: 432	lr: 0.0001	evaluation reward: 7.89
episode: 2844	score: 4.0	memory length: 712574	epsilon:
0.009998020008555413	steps: 244	lr: 0.0001	evaluation reward: 7.85
episode: 2845	score: 5.0	memory length: 712863	epsilon:
0.009998020008555413	steps: 289	lr: 0.0001	evaluation reward: 7.84
episode: 2846	score: 9.0	memory length: 713300	epsilon:
0.009998020008555413	steps: 437	lr: 0.0001	evaluation reward: 7.86
episode: 2847	score: 7.0	memory length: 713702	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 7.86
episode: 2848	score: 8.0	memory length: 714117	epsilon:
0.009998020008555413	steps: 415	lr: 0.0001	evaluation reward: 7.84
episode: 2849	score: 7.0	memory length: 714500	epsilon:
0.009998020008555413	steps: 383	lr: 0.0001	evaluation reward: 7.84
episode: 2850	score: 6.0	memory length: 714879	epsilon:
0.009998020008555413	steps: 379	lr: 0.0001	evaluation reward: 7.84
episode: 2851	score: 6.0	memory length: 715232	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 7.84
episode: 2852	score: 6.0	memory length: 715568	epsilon:
0.009998020008555413	steps: 336	lr: 0.0001	evaluation reward: 7.79
episode: 2853	score: 5.0	memory length: 715855	epsilon:
0.009998020008555413	steps: 287	lr: 0.0001	evaluation reward: 7.77
episode: 2854	score: 6.0	memory length: 716230	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 7.74
episode: 2855	score: 7.0	memory length: 716633	epsilon:
0.009998020008555413	steps: 403	lr: 0.0001	evaluation reward: 7.75
episode: 2856	score: 6.0	memory length: 716970	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 7.74
episode: 2857	score: 8.0	memory length: 717383	epsilon:
0.009998020008555413	steps: 413	lr: 0.0001	evaluation reward: 7.69
episode: 2858	score: 6.0	memory length: 717738	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 7.69
episode: 2859	score: 7.0	memory length: 718147	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 7.66
episode: 2860	score: 12.0	memory length: 718675	epsilon:
0.009998020008555413	steps: 528	lr: 0.0001	evaluation reward: 7.69
episode: 2861	score: 7.0	memory length: 719058	epsilon:
0.009998020008555413	steps: 383	lr: 0.0001	evaluation reward: 7.69
episode: 2862	score: 6.0	memory length: 719396	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 7.69
episode: 2863	score: 4.0	memory length: 719635	epsilon:
0.009998020008555413	steps: 239	lr: 0.0001	evaluation reward: 7.66
episode: 2864	score: 9.0	memory length: 720078	epsilon:

0.009998020008555413	steps: 443	lr: 0.0001	evaluation reward: 7.69
episode: 2865	score: 12.0	memory length: 720582	epsilon:
0.009998020008555413	steps: 504	lr: 0.0001	evaluation reward: 7.74
episode: 2866	score: 8.0	memory length: 721039	epsilon:
0.009998020008555413	steps: 457	lr: 0.0001	evaluation reward: 7.73
episode: 2867	score: 6.0	memory length: 721414	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 7.72
episode: 2868	score: 10.0	memory length: 721937	epsilon:
0.009998020008555413	steps: 523	lr: 0.0001	evaluation reward: 7.74
episode: 2869	score: 10.0	memory length: 722441	epsilon:
0.009998020008555413	steps: 504	lr: 0.0001	evaluation reward: 7.78
episode: 2870	score: 11.0	memory length: 722953	epsilon:
0.009998020008555413	steps: 512	lr: 0.0001	evaluation reward: 7.8
episode: 2871	score: 3.0	memory length: 723166	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 7.78
episode: 2872	score: 3.0	memory length: 723379	epsilon:
0.009998020008555413	steps: 213	lr: 0.0001	evaluation reward: 7.78
episode: 2873	score: 6.0	memory length: 723750	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 7.79
episode: 2874	score: 13.0	memory length: 724226	epsilon:
0.009998020008555413	steps: 476	lr: 0.0001	evaluation reward: 7.84
episode: 2875	score: 10.0	memory length: 724741	epsilon:
0.009998020008555413	steps: 515	lr: 0.0001	evaluation reward: 7.87
episode: 2876	score: 6.0	memory length: 725074	epsilon:
0.009998020008555413	steps: 333	lr: 0.0001	evaluation reward: 7.87
episode: 2877	score: 6.0	memory length: 725416	epsilon:
0.009998020008555413	steps: 342	lr: 0.0001	evaluation reward: 7.84
episode: 2878	score: 9.0	memory length: 725923	epsilon:
0.009998020008555413	steps: 507	lr: 0.0001	evaluation reward: 7.82
episode: 2879	score: 8.0	memory length: 726357	epsilon:
0.009998020008555413	steps: 434	lr: 0.0001	evaluation reward: 7.74
episode: 2880	score: 3.0	memory length: 726567	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 7.69
episode: 2881	score: 8.0	memory length: 726966	epsilon:
0.009998020008555413	steps: 399	lr: 0.0001	evaluation reward: 7.72
episode: 2882	score: 4.0	memory length: 727210	epsilon:
0.009998020008555413	steps: 244	lr: 0.0001	evaluation reward: 7.65
episode: 2883	score: 8.0	memory length: 727650	epsilon:
0.009998020008555413	steps: 440	lr: 0.0001	evaluation reward: 7.67
episode: 2884	score: 7.0	memory length: 728059	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 7.68
episode: 2885	score: 6.0	memory length: 728399	epsilon:
0.009998020008555413	steps: 340	lr: 0.0001	evaluation reward: 7.65
episode: 2886	score: 6.0	memory length: 728774	epsilon:
0.009998020008555413	steps: 375	lr: 0.0001	evaluation reward: 7.66
episode: 2887	score: 7.0	memory length: 729181	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 7.62
episode: 2888	score: 7.0	memory length: 729541	epsilon:

0.009998020008555413	steps: 360	lr: 0.0001	evaluation reward: 7.6
episode: 2889	score: 7.0	memory length: 729941	epsilon:
0.009998020008555413	steps: 400	lr: 0.0001	evaluation reward: 7.6
episode: 2890	score: 8.0	memory length: 730377	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 7.61
episode: 2891	score: 5.0	memory length: 730680	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 7.54
episode: 2892	score: 6.0	memory length: 731049	epsilon:
0.009998020008555413	steps: 369	lr: 0.0001	evaluation reward: 7.56
episode: 2893	score: 12.0	memory length: 731672	epsilon:
0.009998020008555413	steps: 623	lr: 0.0001	evaluation reward: 7.65
episode: 2894	score: 8.0	memory length: 732113	epsilon:
0.009998020008555413	steps: 441	lr: 0.0001	evaluation reward: 7.63
episode: 2895	score: 12.0	memory length: 732584	epsilon:
0.009998020008555413	steps: 471	lr: 0.0001	evaluation reward: 7.65
episode: 2896	score: 5.0	memory length: 732908	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 7.66
episode: 2897	score: 5.0	memory length: 733213	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 7.63
episode: 2898	score: 9.0	memory length: 733672	epsilon:
0.009998020008555413	steps: 459	lr: 0.0001	evaluation reward: 7.67
episode: 2899	score: 7.0	memory length: 734050	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 7.69
episode: 2900	score: 8.0	memory length: 734506	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 7.7
episode: 2901	score: 11.0	memory length: 735053	epsilon:
0.009998020008555413	steps: 547	lr: 0.0001	evaluation reward: 7.73
episode: 2902	score: 7.0	memory length: 735437	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 7.73
episode: 2903	score: 7.0	memory length: 735805	epsilon:
0.009998020008555413	steps: 368	lr: 0.0001	evaluation reward: 7.7
episode: 2904	score: 8.0	memory length: 736225	epsilon:
0.009998020008555413	steps: 420	lr: 0.0001	evaluation reward: 7.72
episode: 2905	score: 10.0	memory length: 736722	epsilon:
0.009998020008555413	steps: 497	lr: 0.0001	evaluation reward: 7.78
episode: 2906	score: 5.0	memory length: 736995	epsilon:
0.009998020008555413	steps: 273	lr: 0.0001	evaluation reward: 7.77
episode: 2907	score: 7.0	memory length: 737400	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 7.73
episode: 2908	score: 9.0	memory length: 737850	epsilon:
0.009998020008555413	steps: 450	lr: 0.0001	evaluation reward: 7.76
episode: 2909	score: 4.0	memory length: 738129	epsilon:
0.009998020008555413	steps: 279	lr: 0.0001	evaluation reward: 7.72
episode: 2910	score: 7.0	memory length: 738530	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 7.68
episode: 2911	score: 10.0	memory length: 739046	epsilon:
0.009998020008555413	steps: 516	lr: 0.0001	evaluation reward: 7.69
episode: 2912	score: 12.0	memory length: 739629	epsilon:

0.009998020008555413	steps: 583	lr: 0.0001	evaluation reward: 7.76
episode: 2913	score: 9.0	memory length: 740075	epsilon:
0.009998020008555413	steps: 446	lr: 0.0001	evaluation reward: 7.75
episode: 2914	score: 6.0	memory length: 740428	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 7.78
episode: 2915	score: 15.0	memory length: 740839	epsilon:
0.009998020008555413	steps: 411	lr: 0.0001	evaluation reward: 7.75
episode: 2916	score: 5.0	memory length: 741124	epsilon:
0.009998020008555413	steps: 285	lr: 0.0001	evaluation reward: 7.75
episode: 2917	score: 9.0	memory length: 741606	epsilon:
0.009998020008555413	steps: 482	lr: 0.0001	evaluation reward: 7.74
episode: 2918	score: 6.0	memory length: 741921	epsilon:
0.009998020008555413	steps: 315	lr: 0.0001	evaluation reward: 7.7
episode: 2919	score: 5.0	memory length: 742228	epsilon:
0.009998020008555413	steps: 307	lr: 0.0001	evaluation reward: 7.69
episode: 2920	score: 7.0	memory length: 742656	epsilon:
0.009998020008555413	steps: 428	lr: 0.0001	evaluation reward: 7.7
episode: 2921	score: 6.0	memory length: 742971	epsilon:
0.009998020008555413	steps: 315	lr: 0.0001	evaluation reward: 7.64
episode: 2922	score: 6.0	memory length: 743323	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 7.62
episode: 2923	score: 5.0	memory length: 743627	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 7.55
episode: 2924	score: 8.0	memory length: 744042	epsilon:
0.009998020008555413	steps: 415	lr: 0.0001	evaluation reward: 7.55
episode: 2925	score: 6.0	memory length: 744358	epsilon:
0.009998020008555413	steps: 316	lr: 0.0001	evaluation reward: 7.5
episode: 2926	score: 9.0	memory length: 744832	epsilon:
0.009998020008555413	steps: 474	lr: 0.0001	evaluation reward: 7.52
episode: 2927	score: 4.0	memory length: 745086	epsilon:
0.009998020008555413	steps: 254	lr: 0.0001	evaluation reward: 7.51
episode: 2928	score: 8.0	memory length: 745540	epsilon:
0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 7.5
episode: 2929	score: 7.0	memory length: 745960	epsilon:
0.009998020008555413	steps: 420	lr: 0.0001	evaluation reward: 7.49
episode: 2930	score: 6.0	memory length: 746298	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 7.49
episode: 2931	score: 6.0	memory length: 746613	epsilon:
0.009998020008555413	steps: 315	lr: 0.0001	evaluation reward: 7.49
episode: 2932	score: 7.0	memory length: 746967	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 7.48
episode: 2933	score: 3.0	memory length: 747179	epsilon:
0.009998020008555413	steps: 212	lr: 0.0001	evaluation reward: 7.4
episode: 2934	score: 5.0	memory length: 747470	epsilon:
0.009998020008555413	steps: 291	lr: 0.0001	evaluation reward: 7.34
episode: 2935	score: 8.0	memory length: 747908	epsilon:
0.009998020008555413	steps: 438	lr: 0.0001	evaluation reward: 7.35
episode: 2936	score: 9.0	memory length: 748416	epsilon:

0.009998020008555413	steps: 508	lr: 0.0001	evaluation reward: 7.32
episode: 2937	score: 11.0	memory length: 748958	epsilon:
0.009998020008555413	steps: 542	lr: 0.0001	evaluation reward: 7.37
episode: 2938	score: 10.0	memory length: 749440	epsilon:
0.009998020008555413	steps: 482	lr: 0.0001	evaluation reward: 7.37
episode: 2939	score: 6.0	memory length: 749796	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 7.32
episode: 2940	score: 6.0	memory length: 750175	epsilon:
0.009998020008555413	steps: 379	lr: 0.0001	evaluation reward: 7.25
episode: 2941	score: 10.0	memory length: 750730	epsilon:
0.009998020008555413	steps: 555	lr: 0.0001	evaluation reward: 7.3
episode: 2942	score: 11.0	memory length: 751133	epsilon:
0.009998020008555413	steps: 403	lr: 0.0001	evaluation reward: 7.34
episode: 2943	score: 6.0	memory length: 751451	epsilon:
0.009998020008555413	steps: 318	lr: 0.0001	evaluation reward: 7.32
episode: 2944	score: 7.0	memory length: 751872	epsilon:
0.009998020008555413	steps: 421	lr: 0.0001	evaluation reward: 7.35
episode: 2945	score: 5.0	memory length: 752173	epsilon:
0.009998020008555413	steps: 301	lr: 0.0001	evaluation reward: 7.35
episode: 2946	score: 5.0	memory length: 752464	epsilon:
0.009998020008555413	steps: 291	lr: 0.0001	evaluation reward: 7.31
episode: 2947	score: 8.0	memory length: 752868	epsilon:
0.009998020008555413	steps: 404	lr: 0.0001	evaluation reward: 7.32
episode: 2948	score: 7.0	memory length: 753289	epsilon:
0.009998020008555413	steps: 421	lr: 0.0001	evaluation reward: 7.31
episode: 2949	score: 10.0	memory length: 753777	epsilon:
0.009998020008555413	steps: 488	lr: 0.0001	evaluation reward: 7.34
episode: 2950	score: 10.0	memory length: 754323	epsilon:
0.009998020008555413	steps: 546	lr: 0.0001	evaluation reward: 7.38
episode: 2951	score: 12.0	memory length: 754847	epsilon:
0.009998020008555413	steps: 524	lr: 0.0001	evaluation reward: 7.44
episode: 2952	score: 5.0	memory length: 755172	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 7.43
episode: 2953	score: 8.0	memory length: 755628	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 7.46
episode: 2954	score: 10.0	memory length: 756121	epsilon:
0.009998020008555413	steps: 493	lr: 0.0001	evaluation reward: 7.5
episode: 2955	score: 15.0	memory length: 756576	epsilon:
0.009998020008555413	steps: 455	lr: 0.0001	evaluation reward: 7.58
episode: 2956	score: 11.0	memory length: 757066	epsilon:
0.009998020008555413	steps: 490	lr: 0.0001	evaluation reward: 7.63
episode: 2957	score: 10.0	memory length: 757575	epsilon:
0.009998020008555413	steps: 509	lr: 0.0001	evaluation reward: 7.65
episode: 2958	score: 8.0	memory length: 757891	epsilon:
0.009998020008555413	steps: 316	lr: 0.0001	evaluation reward: 7.67
episode: 2959	score: 7.0	memory length: 758272	epsilon:
0.009998020008555413	steps: 381	lr: 0.0001	evaluation reward: 7.67
episode: 2960	score: 5.0	memory length: 758595	epsilon:

0.009998020008555413	steps: 323	lr: 0.0001	evaluation reward: 7.6
episode: 2961	score: 10.0	memory length: 759096	epsilon:
0.009998020008555413	steps: 501	lr: 0.0001	evaluation reward: 7.63
episode: 2962	score: 6.0	memory length: 759431	epsilon:
0.009998020008555413	steps: 335	lr: 0.0001	evaluation reward: 7.63
episode: 2963	score: 8.0	memory length: 759882	epsilon:
0.009998020008555413	steps: 451	lr: 0.0001	evaluation reward: 7.67
episode: 2964	score: 7.0	memory length: 760254	epsilon:
0.009998020008555413	steps: 372	lr: 0.0001	evaluation reward: 7.65
episode: 2965	score: 12.0	memory length: 760688	epsilon:
0.009998020008555413	steps: 434	lr: 0.0001	evaluation reward: 7.65
episode: 2966	score: 8.0	memory length: 761121	epsilon:
0.009998020008555413	steps: 433	lr: 0.0001	evaluation reward: 7.65
episode: 2967	score: 10.0	memory length: 761607	epsilon:
0.009998020008555413	steps: 486	lr: 0.0001	evaluation reward: 7.69
episode: 2968	score: 6.0	memory length: 761944	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 7.65
episode: 2969	score: 8.0	memory length: 762359	epsilon:
0.009998020008555413	steps: 415	lr: 0.0001	evaluation reward: 7.63
episode: 2970	score: 9.0	memory length: 762830	epsilon:
0.009998020008555413	steps: 471	lr: 0.0001	evaluation reward: 7.61
episode: 2971	score: 11.0	memory length: 763339	epsilon:
0.009998020008555413	steps: 509	lr: 0.0001	evaluation reward: 7.69
episode: 2972	score: 4.0	memory length: 763596	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 7.7
episode: 2973	score: 9.0	memory length: 764118	epsilon:
0.009998020008555413	steps: 522	lr: 0.0001	evaluation reward: 7.73
episode: 2974	score: 6.0	memory length: 764474	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 7.66
episode: 2975	score: 5.0	memory length: 764760	epsilon:
0.009998020008555413	steps: 286	lr: 0.0001	evaluation reward: 7.61
episode: 2976	score: 2.0	memory length: 764941	epsilon:
0.009998020008555413	steps: 181	lr: 0.0001	evaluation reward: 7.57
episode: 2977	score: 3.0	memory length: 765151	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 7.54
episode: 2978	score: 8.0	memory length: 765571	epsilon:
0.009998020008555413	steps: 420	lr: 0.0001	evaluation reward: 7.53
episode: 2979	score: 10.0	memory length: 766102	epsilon:
0.009998020008555413	steps: 531	lr: 0.0001	evaluation reward: 7.55
episode: 2980	score: 5.0	memory length: 766389	epsilon:
0.009998020008555413	steps: 287	lr: 0.0001	evaluation reward: 7.57
episode: 2981	score: 6.0	memory length: 766722	epsilon:
0.009998020008555413	steps: 333	lr: 0.0001	evaluation reward: 7.55
episode: 2982	score: 11.0	memory length: 767160	epsilon:
0.009998020008555413	steps: 438	lr: 0.0001	evaluation reward: 7.62
episode: 2983	score: 5.0	memory length: 767466	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 7.59
episode: 2984	score: 9.0	memory length: 767957	epsilon:

0.009998020008555413	steps: 491	lr: 0.0001	evaluation reward: 7.61
episode: 2985	score: 16.0	memory length: 768437	epsilon:
0.009998020008555413	steps: 480	lr: 0.0001	evaluation reward: 7.71
episode: 2986	score: 8.0	memory length: 768892	epsilon:
0.009998020008555413	steps: 455	lr: 0.0001	evaluation reward: 7.73
episode: 2987	score: 7.0	memory length: 769263	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 7.73
episode: 2988	score: 3.0	memory length: 769473	epsilon:
0.009998020008555413	steps: 210	lr: 0.0001	evaluation reward: 7.69
episode: 2989	score: 10.0	memory length: 769948	epsilon:
0.009998020008555413	steps: 475	lr: 0.0001	evaluation reward: 7.72
episode: 2990	score: 7.0	memory length: 770359	epsilon:
0.009998020008555413	steps: 411	lr: 0.0001	evaluation reward: 7.71
episode: 2991	score: 11.0	memory length: 770891	epsilon:
0.009998020008555413	steps: 532	lr: 0.0001	evaluation reward: 7.77
episode: 2992	score: 3.0	memory length: 771117	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 7.74
episode: 2993	score: 8.0	memory length: 771546	epsilon:
0.009998020008555413	steps: 429	lr: 0.0001	evaluation reward: 7.7
episode: 2994	score: 7.0	memory length: 771916	epsilon:
0.009998020008555413	steps: 370	lr: 0.0001	evaluation reward: 7.69
episode: 2995	score: 7.0	memory length: 772286	epsilon:
0.009998020008555413	steps: 370	lr: 0.0001	evaluation reward: 7.64
episode: 2996	score: 10.0	memory length: 772835	epsilon:
0.009998020008555413	steps: 549	lr: 0.0001	evaluation reward: 7.69
episode: 2997	score: 7.0	memory length: 773242	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 7.71
episode: 2998	score: 7.0	memory length: 773633	epsilon:
0.009998020008555413	steps: 391	lr: 0.0001	evaluation reward: 7.69
episode: 2999	score: 6.0	memory length: 773992	epsilon:
0.009998020008555413	steps: 359	lr: 0.0001	evaluation reward: 7.68
episode: 3000	score: 7.0	memory length: 774336	epsilon:
0.009998020008555413	steps: 344	lr: 0.0001	evaluation reward: 7.67
episode: 3001	score: 7.0	memory length: 774739	epsilon:
0.009998020008555413	steps: 403	lr: 0.0001	evaluation reward: 7.63
episode: 3002	score: 8.0	memory length: 775176	epsilon:
0.009998020008555413	steps: 437	lr: 0.0001	evaluation reward: 7.64
episode: 3003	score: 6.0	memory length: 775512	epsilon:
0.009998020008555413	steps: 336	lr: 0.0001	evaluation reward: 7.63
episode: 3004	score: 9.0	memory length: 775965	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 7.64
episode: 3005	score: 10.0	memory length: 776477	epsilon:
0.009998020008555413	steps: 512	lr: 0.0001	evaluation reward: 7.64
episode: 3006	score: 7.0	memory length: 776843	epsilon:
0.009998020008555413	steps: 366	lr: 0.0001	evaluation reward: 7.66
episode: 3007	score: 7.0	memory length: 777225	epsilon:
0.009998020008555413	steps: 382	lr: 0.0001	evaluation reward: 7.66
episode: 3008	score: 9.0	memory length: 777731	epsilon:

0.009998020008555413	steps: 506	lr: 0.0001	evaluation reward: 7.66
episode: 3009	score: 5.0	memory length: 778018	epsilon:
0.009998020008555413	steps: 287	lr: 0.0001	evaluation reward: 7.67
episode: 3010	score: 8.0	memory length: 778491	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 7.68
episode: 3011	score: 11.0	memory length: 778997	epsilon:
0.009998020008555413	steps: 506	lr: 0.0001	evaluation reward: 7.69
episode: 3012	score: 8.0	memory length: 779461	epsilon:
0.009998020008555413	steps: 464	lr: 0.0001	evaluation reward: 7.65
episode: 3013	score: 11.0	memory length: 779915	epsilon:
0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 7.67
episode: 3014	score: 13.0	memory length: 780419	epsilon:
0.009998020008555413	steps: 504	lr: 0.0001	evaluation reward: 7.74
episode: 3015	score: 5.0	memory length: 780745	epsilon:
0.009998020008555413	steps: 326	lr: 0.0001	evaluation reward: 7.64
episode: 3016	score: 10.0	memory length: 781197	epsilon:
0.009998020008555413	steps: 452	lr: 0.0001	evaluation reward: 7.69
episode: 3017	score: 11.0	memory length: 781760	epsilon:
0.009998020008555413	steps: 563	lr: 0.0001	evaluation reward: 7.71
episode: 3018	score: 5.0	memory length: 782063	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 7.7
episode: 3019	score: 8.0	memory length: 782493	epsilon:
0.009998020008555413	steps: 430	lr: 0.0001	evaluation reward: 7.73
episode: 3020	score: 6.0	memory length: 782862	epsilon:
0.009998020008555413	steps: 369	lr: 0.0001	evaluation reward: 7.72
episode: 3021	score: 8.0	memory length: 783315	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 7.74
episode: 3022	score: 10.0	memory length: 783868	epsilon:
0.009998020008555413	steps: 553	lr: 0.0001	evaluation reward: 7.78
episode: 3023	score: 11.0	memory length: 784365	epsilon:
0.009998020008555413	steps: 497	lr: 0.0001	evaluation reward: 7.84
episode: 3024	score: 7.0	memory length: 784758	epsilon:
0.009998020008555413	steps: 393	lr: 0.0001	evaluation reward: 7.83
episode: 3025	score: 7.0	memory length: 785179	epsilon:
0.009998020008555413	steps: 421	lr: 0.0001	evaluation reward: 7.84
episode: 3026	score: 9.0	memory length: 785630	epsilon:
0.009998020008555413	steps: 451	lr: 0.0001	evaluation reward: 7.84
episode: 3027	score: 9.0	memory length: 786103	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 7.89
episode: 3028	score: 5.0	memory length: 786427	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 7.86
episode: 3029	score: 9.0	memory length: 786776	epsilon:
0.009998020008555413	steps: 349	lr: 0.0001	evaluation reward: 7.88
episode: 3030	score: 4.0	memory length: 787033	epsilon:
0.009998020008555413	steps: 257	lr: 0.0001	evaluation reward: 7.86
episode: 3031	score: 10.0	memory length: 787492	epsilon:
0.009998020008555413	steps: 459	lr: 0.0001	evaluation reward: 7.9
episode: 3032	score: 11.0	memory length: 787952	epsilon:

0.009998020008555413	steps: 460	lr: 0.0001	evaluation reward: 7.94
episode: 3033	score: 6.0	memory length: 788303	epsilon:
0.009998020008555413	steps: 351	lr: 0.0001	evaluation reward: 7.97
episode: 3034	score: 9.0	memory length: 788805	epsilon:
0.009998020008555413	steps: 502	lr: 0.0001	evaluation reward: 8.01
episode: 3035	score: 8.0	memory length: 789220	epsilon:
0.009998020008555413	steps: 415	lr: 0.0001	evaluation reward: 8.01
episode: 3036	score: 11.0	memory length: 789742	epsilon:
0.009998020008555413	steps: 522	lr: 0.0001	evaluation reward: 8.03
episode: 3037	score: 6.0	memory length: 790103	epsilon:
0.009998020008555413	steps: 361	lr: 0.0001	evaluation reward: 7.98
episode: 3038	score: 13.0	memory length: 790593	epsilon:
0.009998020008555413	steps: 490	lr: 0.0001	evaluation reward: 8.01
episode: 3039	score: 10.0	memory length: 791134	epsilon:
0.009998020008555413	steps: 541	lr: 0.0001	evaluation reward: 8.05
episode: 3040	score: 10.0	memory length: 791603	epsilon:
0.009998020008555413	steps: 469	lr: 0.0001	evaluation reward: 8.09
episode: 3041	score: 6.0	memory length: 791956	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 8.05
episode: 3042	score: 8.0	memory length: 792378	epsilon:
0.009998020008555413	steps: 422	lr: 0.0001	evaluation reward: 8.02
episode: 3043	score: 7.0	memory length: 792774	epsilon:
0.009998020008555413	steps: 396	lr: 0.0001	evaluation reward: 8.03
episode: 3044	score: 11.0	memory length: 793311	epsilon:
0.009998020008555413	steps: 537	lr: 0.0001	evaluation reward: 8.07
episode: 3045	score: 13.0	memory length: 793797	epsilon:
0.009998020008555413	steps: 486	lr: 0.0001	evaluation reward: 8.15
episode: 3046	score: 8.0	memory length: 794203	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 8.18
episode: 3047	score: 6.0	memory length: 794544	epsilon:
0.009998020008555413	steps: 341	lr: 0.0001	evaluation reward: 8.16
episode: 3048	score: 11.0	memory length: 795027	epsilon:
0.009998020008555413	steps: 483	lr: 0.0001	evaluation reward: 8.2
episode: 3049	score: 10.0	memory length: 795512	epsilon:
0.009998020008555413	steps: 485	lr: 0.0001	evaluation reward: 8.2
episode: 3050	score: 5.0	memory length: 795814	epsilon:
0.009998020008555413	steps: 302	lr: 0.0001	evaluation reward: 8.15
episode: 3051	score: 11.0	memory length: 796177	epsilon:
0.009998020008555413	steps: 363	lr: 0.0001	evaluation reward: 8.14
episode: 3052	score: 10.0	memory length: 796686	epsilon:
0.009998020008555413	steps: 509	lr: 0.0001	evaluation reward: 8.19
episode: 3053	score: 10.0	memory length: 797206	epsilon:
0.009998020008555413	steps: 520	lr: 0.0001	evaluation reward: 8.21
episode: 3054	score: 12.0	memory length: 797785	epsilon:
0.009998020008555413	steps: 579	lr: 0.0001	evaluation reward: 8.23
episode: 3055	score: 4.0	memory length: 798059	epsilon:
0.009998020008555413	steps: 274	lr: 0.0001	evaluation reward: 8.12
episode: 3056	score: 5.0	memory length: 798350	epsilon:

0.009998020008555413	steps: 291	lr: 0.0001	evaluation reward: 8.06
episode: 3057	score: 10.0	memory length: 798906	epsilon:
0.009998020008555413	steps: 556	lr: 0.0001	evaluation reward: 8.06
episode: 3058	score: 7.0	memory length: 799304	epsilon:
0.009998020008555413	steps: 398	lr: 0.0001	evaluation reward: 8.05
episode: 3059	score: 10.0	memory length: 799840	epsilon:
0.009998020008555413	steps: 536	lr: 0.0001	evaluation reward: 8.08
episode: 3060	score: 9.0	memory length: 800280	epsilon:
0.009998020008555413	steps: 440	lr: 0.0001	evaluation reward: 8.12
episode: 3061	score: 8.0	memory length: 800722	epsilon:
0.009998020008555413	steps: 442	lr: 0.0001	evaluation reward: 8.1
episode: 3062	score: 8.0	memory length: 801148	epsilon:
0.009998020008555413	steps: 426	lr: 0.0001	evaluation reward: 8.12
episode: 3063	score: 6.0	memory length: 801493	epsilon:
0.009998020008555413	steps: 345	lr: 0.0001	evaluation reward: 8.1
episode: 3064	score: 8.0	memory length: 801924	epsilon:
0.009998020008555413	steps: 431	lr: 0.0001	evaluation reward: 8.11
episode: 3065	score: 11.0	memory length: 802476	epsilon:
0.009998020008555413	steps: 552	lr: 0.0001	evaluation reward: 8.1
episode: 3066	score: 6.0	memory length: 802830	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 8.08
episode: 3067	score: 9.0	memory length: 803280	epsilon:
0.009998020008555413	steps: 450	lr: 0.0001	evaluation reward: 8.07
episode: 3068	score: 5.0	memory length: 803585	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 8.06
episode: 3069	score: 9.0	memory length: 804053	epsilon:
0.009998020008555413	steps: 468	lr: 0.0001	evaluation reward: 8.07
episode: 3070	score: 11.0	memory length: 804540	epsilon:
0.009998020008555413	steps: 487	lr: 0.0001	evaluation reward: 8.09
episode: 3071	score: 9.0	memory length: 805039	epsilon:
0.009998020008555413	steps: 499	lr: 0.0001	evaluation reward: 8.07
episode: 3072	score: 5.0	memory length: 805349	epsilon:
0.009998020008555413	steps: 310	lr: 0.0001	evaluation reward: 8.08
episode: 3073	score: 10.0	memory length: 805811	epsilon:
0.009998020008555413	steps: 462	lr: 0.0001	evaluation reward: 8.09
episode: 3074	score: 8.0	memory length: 806216	epsilon:
0.009998020008555413	steps: 405	lr: 0.0001	evaluation reward: 8.11
episode: 3075	score: 11.0	memory length: 806744	epsilon:
0.009998020008555413	steps: 528	lr: 0.0001	evaluation reward: 8.17
episode: 3076	score: 19.0	memory length: 807404	epsilon:
0.009998020008555413	steps: 660	lr: 0.0001	evaluation reward: 8.34
episode: 3077	score: 7.0	memory length: 807775	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 8.38
episode: 3078	score: 7.0	memory length: 808177	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 8.37
episode: 3079	score: 8.0	memory length: 808649	epsilon:
0.009998020008555413	steps: 472	lr: 0.0001	evaluation reward: 8.35
episode: 3080	score: 8.0	memory length: 809063	epsilon:

0.009998020008555413	steps: 414	lr: 0.0001	evaluation reward: 8.38
episode: 3081	score: 7.0	memory length: 809435	epsilon:
0.009998020008555413	steps: 372	lr: 0.0001	evaluation reward: 8.39
episode: 3082	score: 10.0	memory length: 809928	epsilon:
0.009998020008555413	steps: 493	lr: 0.0001	evaluation reward: 8.38
episode: 3083	score: 10.0	memory length: 810412	epsilon:
0.009998020008555413	steps: 484	lr: 0.0001	evaluation reward: 8.43
episode: 3084	score: 9.0	memory length: 810896	epsilon:
0.009998020008555413	steps: 484	lr: 0.0001	evaluation reward: 8.43
episode: 3085	score: 13.0	memory length: 811379	epsilon:
0.009998020008555413	steps: 483	lr: 0.0001	evaluation reward: 8.4
episode: 3086	score: 10.0	memory length: 811914	epsilon:
0.009998020008555413	steps: 535	lr: 0.0001	evaluation reward: 8.42
episode: 3087	score: 8.0	memory length: 812370	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 8.43
episode: 3088	score: 9.0	memory length: 812840	epsilon:
0.009998020008555413	steps: 470	lr: 0.0001	evaluation reward: 8.49
episode: 3089	score: 8.0	memory length: 813296	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 8.47
episode: 3090	score: 6.0	memory length: 813652	epsilon:
0.009998020008555413	steps: 356	lr: 0.0001	evaluation reward: 8.46
episode: 3091	score: 7.0	memory length: 814054	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 8.42
episode: 3092	score: 9.0	memory length: 814483	epsilon:
0.009998020008555413	steps: 429	lr: 0.0001	evaluation reward: 8.48
episode: 3093	score: 7.0	memory length: 814894	epsilon:
0.009998020008555413	steps: 411	lr: 0.0001	evaluation reward: 8.47
episode: 3094	score: 6.0	memory length: 815215	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 8.46
episode: 3095	score: 9.0	memory length: 815681	epsilon:
0.009998020008555413	steps: 466	lr: 0.0001	evaluation reward: 8.48
episode: 3096	score: 4.0	memory length: 815954	epsilon:
0.009998020008555413	steps: 273	lr: 0.0001	evaluation reward: 8.42
episode: 3097	score: 6.0	memory length: 816292	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 8.41
episode: 3098	score: 10.0	memory length: 816861	epsilon:
0.009998020008555413	steps: 569	lr: 0.0001	evaluation reward: 8.44
episode: 3099	score: 9.0	memory length: 817361	epsilon:
0.009998020008555413	steps: 500	lr: 0.0001	evaluation reward: 8.47
episode: 3100	score: 8.0	memory length: 817798	epsilon:
0.009998020008555413	steps: 437	lr: 0.0001	evaluation reward: 8.48
episode: 3101	score: 10.0	memory length: 818294	epsilon:
0.009998020008555413	steps: 496	lr: 0.0001	evaluation reward: 8.51
episode: 3102	score: 6.0	memory length: 818599	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 8.49
episode: 3103	score: 6.0	memory length: 818916	epsilon:
0.009998020008555413	steps: 317	lr: 0.0001	evaluation reward: 8.49
episode: 3104	score: 17.0	memory length: 819546	epsilon:

0.009998020008555413	steps: 630	lr: 0.0001	evaluation reward: 8.57
episode: 3105	score: 12.0	memory length: 820096	epsilon:
0.009998020008555413	steps: 550	lr: 0.0001	evaluation reward: 8.59
episode: 3106	score: 8.0	memory length: 820505	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 8.6
episode: 3107	score: 6.0	memory length: 820858	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 8.59
episode: 3108	score: 10.0	memory length: 821227	epsilon:
0.009998020008555413	steps: 369	lr: 0.0001	evaluation reward: 8.6
episode: 3109	score: 5.0	memory length: 821515	epsilon:
0.009998020008555413	steps: 288	lr: 0.0001	evaluation reward: 8.6
episode: 3110	score: 5.0	memory length: 821799	epsilon:
0.009998020008555413	steps: 284	lr: 0.0001	evaluation reward: 8.57
episode: 3111	score: 7.0	memory length: 822183	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 8.53
episode: 3112	score: 8.0	memory length: 822652	epsilon:
0.009998020008555413	steps: 469	lr: 0.0001	evaluation reward: 8.53
episode: 3113	score: 7.0	memory length: 823045	epsilon:
0.009998020008555413	steps: 393	lr: 0.0001	evaluation reward: 8.49
episode: 3114	score: 6.0	memory length: 823382	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 8.42
episode: 3115	score: 11.0	memory length: 823740	epsilon:
0.009998020008555413	steps: 358	lr: 0.0001	evaluation reward: 8.48
episode: 3116	score: 9.0	memory length: 824180	epsilon:
0.009998020008555413	steps: 440	lr: 0.0001	evaluation reward: 8.47
episode: 3117	score: 3.0	memory length: 824405	epsilon:
0.009998020008555413	steps: 225	lr: 0.0001	evaluation reward: 8.39
episode: 3118	score: 10.0	memory length: 824924	epsilon:
0.009998020008555413	steps: 519	lr: 0.0001	evaluation reward: 8.44
episode: 3119	score: 15.0	memory length: 825486	epsilon:
0.009998020008555413	steps: 562	lr: 0.0001	evaluation reward: 8.51
episode: 3120	score: 8.0	memory length: 825939	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 8.53
episode: 3121	score: 12.0	memory length: 826532	epsilon:
0.009998020008555413	steps: 593	lr: 0.0001	evaluation reward: 8.57
episode: 3122	score: 9.0	memory length: 826960	epsilon:
0.009998020008555413	steps: 428	lr: 0.0001	evaluation reward: 8.56
episode: 3123	score: 8.0	memory length: 827355	epsilon:
0.009998020008555413	steps: 395	lr: 0.0001	evaluation reward: 8.53
episode: 3124	score: 10.0	memory length: 827823	epsilon:
0.009998020008555413	steps: 468	lr: 0.0001	evaluation reward: 8.56
episode: 3125	score: 10.0	memory length: 828361	epsilon:
0.009998020008555413	steps: 538	lr: 0.0001	evaluation reward: 8.59
episode: 3126	score: 3.0	memory length: 828587	epsilon:
0.009998020008555413	steps: 226	lr: 0.0001	evaluation reward: 8.53
episode: 3127	score: 11.0	memory length: 829109	epsilon:
0.009998020008555413	steps: 522	lr: 0.0001	evaluation reward: 8.55
episode: 3128	score: 5.0	memory length: 829380	epsilon:

0.009998020008555413	steps: 271	lr: 0.0001	evaluation reward: 8.55
episode: 3129	score: 14.0	memory length: 829888	epsilon:
0.009998020008555413	steps: 508	lr: 0.0001	evaluation reward: 8.6
episode: 3130	score: 9.0	memory length: 830324	epsilon:
0.009998020008555413	steps: 436	lr: 0.0001	evaluation reward: 8.65
episode: 3131	score: 5.0	memory length: 830627	epsilon:
0.009998020008555413	steps: 303	lr: 0.0001	evaluation reward: 8.6
episode: 3132	score: 10.0	memory length: 831139	epsilon:
0.009998020008555413	steps: 512	lr: 0.0001	evaluation reward: 8.59
episode: 3133	score: 6.0	memory length: 831492	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 8.59
episode: 3134	score: 6.0	memory length: 831841	epsilon:
0.009998020008555413	steps: 349	lr: 0.0001	evaluation reward: 8.56
episode: 3135	score: 9.0	memory length: 832272	epsilon:
0.009998020008555413	steps: 431	lr: 0.0001	evaluation reward: 8.57
episode: 3136	score: 11.0	memory length: 832773	epsilon:
0.009998020008555413	steps: 501	lr: 0.0001	evaluation reward: 8.57
episode: 3137	score: 9.0	memory length: 833294	epsilon:
0.009998020008555413	steps: 521	lr: 0.0001	evaluation reward: 8.6
episode: 3138	score: 5.0	memory length: 833581	epsilon:
0.009998020008555413	steps: 287	lr: 0.0001	evaluation reward: 8.52
episode: 3139	score: 9.0	memory length: 834018	epsilon:
0.009998020008555413	steps: 437	lr: 0.0001	evaluation reward: 8.51
episode: 3140	score: 7.0	memory length: 834402	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 8.48
episode: 3141	score: 7.0	memory length: 834765	epsilon:
0.009998020008555413	steps: 363	lr: 0.0001	evaluation reward: 8.49
episode: 3142	score: 9.0	memory length: 835209	epsilon:
0.009998020008555413	steps: 444	lr: 0.0001	evaluation reward: 8.5
episode: 3143	score: 9.0	memory length: 835674	epsilon:
0.009998020008555413	steps: 465	lr: 0.0001	evaluation reward: 8.52
episode: 3144	score: 9.0	memory length: 836128	epsilon:
0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 8.5
episode: 3145	score: 12.0	memory length: 836699	epsilon:
0.009998020008555413	steps: 571	lr: 0.0001	evaluation reward: 8.49
episode: 3146	score: 10.0	memory length: 837173	epsilon:
0.009998020008555413	steps: 474	lr: 0.0001	evaluation reward: 8.51
episode: 3147	score: 5.0	memory length: 837458	epsilon:
0.009998020008555413	steps: 285	lr: 0.0001	evaluation reward: 8.5
episode: 3148	score: 13.0	memory length: 837922	epsilon:
0.009998020008555413	steps: 464	lr: 0.0001	evaluation reward: 8.52
episode: 3149	score: 10.0	memory length: 838394	epsilon:
0.009998020008555413	steps: 472	lr: 0.0001	evaluation reward: 8.52
episode: 3150	score: 10.0	memory length: 838890	epsilon:
0.009998020008555413	steps: 496	lr: 0.0001	evaluation reward: 8.57
episode: 3151	score: 10.0	memory length: 839388	epsilon:
0.009998020008555413	steps: 498	lr: 0.0001	evaluation reward: 8.56
episode: 3152	score: 4.0	memory length: 839634	epsilon:

0.009998020008555413	steps: 246	lr: 0.0001	evaluation reward: 8.5
episode: 3153	score: 9.0	memory length: 840104	epsilon:
0.009998020008555413	steps: 470	lr: 0.0001	evaluation reward: 8.49
episode: 3154	score: 9.0	memory length: 840549	epsilon:
0.009998020008555413	steps: 445	lr: 0.0001	evaluation reward: 8.46
episode: 3155	score: 6.0	memory length: 840879	epsilon:
0.009998020008555413	steps: 330	lr: 0.0001	evaluation reward: 8.48
episode: 3156	score: 8.0	memory length: 841322	epsilon:
0.009998020008555413	steps: 443	lr: 0.0001	evaluation reward: 8.51
episode: 3157	score: 12.0	memory length: 841921	epsilon:
0.009998020008555413	steps: 599	lr: 0.0001	evaluation reward: 8.53
episode: 3158	score: 6.0	memory length: 842258	epsilon:
0.009998020008555413	steps: 337	lr: 0.0001	evaluation reward: 8.52
episode: 3159	score: 9.0	memory length: 842728	epsilon:
0.009998020008555413	steps: 470	lr: 0.0001	evaluation reward: 8.51
episode: 3160	score: 8.0	memory length: 843129	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 8.5
episode: 3161	score: 5.0	memory length: 843473	epsilon:
0.009998020008555413	steps: 344	lr: 0.0001	evaluation reward: 8.47
episode: 3162	score: 17.0	memory length: 844095	epsilon:
0.009998020008555413	steps: 622	lr: 0.0001	evaluation reward: 8.56
episode: 3163	score: 10.0	memory length: 844593	epsilon:
0.009998020008555413	steps: 498	lr: 0.0001	evaluation reward: 8.6
episode: 3164	score: 14.0	memory length: 845109	epsilon:
0.009998020008555413	steps: 516	lr: 0.0001	evaluation reward: 8.66
episode: 3165	score: 9.0	memory length: 845551	epsilon:
0.009998020008555413	steps: 442	lr: 0.0001	evaluation reward: 8.64
episode: 3166	score: 6.0	memory length: 845886	epsilon:
0.009998020008555413	steps: 335	lr: 0.0001	evaluation reward: 8.64
episode: 3167	score: 7.0	memory length: 846271	epsilon:
0.009998020008555413	steps: 385	lr: 0.0001	evaluation reward: 8.62
episode: 3168	score: 8.0	memory length: 846656	epsilon:
0.009998020008555413	steps: 385	lr: 0.0001	evaluation reward: 8.65
episode: 3169	score: 7.0	memory length: 847039	epsilon:
0.009998020008555413	steps: 383	lr: 0.0001	evaluation reward: 8.63
episode: 3170	score: 7.0	memory length: 847402	epsilon:
0.009998020008555413	steps: 363	lr: 0.0001	evaluation reward: 8.59
episode: 3171	score: 5.0	memory length: 847748	epsilon:
0.009998020008555413	steps: 346	lr: 0.0001	evaluation reward: 8.55
episode: 3172	score: 7.0	memory length: 848088	epsilon:
0.009998020008555413	steps: 340	lr: 0.0001	evaluation reward: 8.57
episode: 3173	score: 10.0	memory length: 848530	epsilon:
0.009998020008555413	steps: 442	lr: 0.0001	evaluation reward: 8.57
episode: 3174	score: 14.0	memory length: 848882	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 8.63
episode: 3175	score: 6.0	memory length: 849203	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 8.58
episode: 3176	score: 10.0	memory length: 849675	epsilon:

0.009998020008555413	steps: 472	lr: 0.0001	evaluation reward: 8.49
episode: 3177	score: 9.0	memory length: 850110	epsilon:
0.009998020008555413	steps: 435	lr: 0.0001	evaluation reward: 8.51
episode: 3178	score: 6.0	memory length: 850429	epsilon:
0.009998020008555413	steps: 319	lr: 0.0001	evaluation reward: 8.5
episode: 3179	score: 10.0	memory length: 850919	epsilon:
0.009998020008555413	steps: 490	lr: 0.0001	evaluation reward: 8.52
episode: 3180	score: 13.0	memory length: 851302	epsilon:
0.009998020008555413	steps: 383	lr: 0.0001	evaluation reward: 8.57
episode: 3181	score: 15.0	memory length: 851994	epsilon:
0.009998020008555413	steps: 692	lr: 0.0001	evaluation reward: 8.65
episode: 3182	score: 7.0	memory length: 852359	epsilon:
0.009998020008555413	steps: 365	lr: 0.0001	evaluation reward: 8.62
episode: 3183	score: 16.0	memory length: 852866	epsilon:
0.009998020008555413	steps: 507	lr: 0.0001	evaluation reward: 8.68
episode: 3184	score: 10.0	memory length: 853355	epsilon:
0.009998020008555413	steps: 489	lr: 0.0001	evaluation reward: 8.69
episode: 3185	score: 10.0	memory length: 853832	epsilon:
0.009998020008555413	steps: 477	lr: 0.0001	evaluation reward: 8.66
episode: 3186	score: 8.0	memory length: 854239	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 8.64
episode: 3187	score: 16.0	memory length: 854780	epsilon:
0.009998020008555413	steps: 541	lr: 0.0001	evaluation reward: 8.72
episode: 3188	score: 12.0	memory length: 855222	epsilon:
0.009998020008555413	steps: 442	lr: 0.0001	evaluation reward: 8.75
episode: 3189	score: 7.0	memory length: 855621	epsilon:
0.009998020008555413	steps: 399	lr: 0.0001	evaluation reward: 8.74
episode: 3190	score: 12.0	memory length: 856132	epsilon:
0.009998020008555413	steps: 511	lr: 0.0001	evaluation reward: 8.8
episode: 3191	score: 14.0	memory length: 856650	epsilon:
0.009998020008555413	steps: 518	lr: 0.0001	evaluation reward: 8.87
episode: 3192	score: 7.0	memory length: 857057	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 8.85
episode: 3193	score: 11.0	memory length: 857557	epsilon:
0.009998020008555413	steps: 500	lr: 0.0001	evaluation reward: 8.89
episode: 3194	score: 5.0	memory length: 857881	epsilon:
0.009998020008555413	steps: 324	lr: 0.0001	evaluation reward: 8.88
episode: 3195	score: 9.0	memory length: 858337	epsilon:
0.009998020008555413	steps: 456	lr: 0.0001	evaluation reward: 8.88
episode: 3196	score: 12.0	memory length: 858768	epsilon:
0.009998020008555413	steps: 431	lr: 0.0001	evaluation reward: 8.96
episode: 3197	score: 7.0	memory length: 859175	epsilon:
0.009998020008555413	steps: 407	lr: 0.0001	evaluation reward: 8.97
episode: 3198	score: 22.0	memory length: 859861	epsilon:
0.009998020008555413	steps: 686	lr: 0.0001	evaluation reward: 9.09
episode: 3199	score: 11.0	memory length: 860428	epsilon:
0.009998020008555413	steps: 567	lr: 0.0001	evaluation reward: 9.11
episode: 3200	score: 15.0	memory length: 861137	epsilon:

0.009998020008555413	steps: 709	lr: 0.0001	evaluation reward: 9.18
episode: 3201	score: 13.0	memory length: 861674	epsilon:
0.009998020008555413	steps: 537	lr: 0.0001	evaluation reward: 9.21
episode: 3202	score: 10.0	memory length: 862192	epsilon:
0.009998020008555413	steps: 518	lr: 0.0001	evaluation reward: 9.25
episode: 3203	score: 6.0	memory length: 862544	epsilon:
0.009998020008555413	steps: 352	lr: 0.0001	evaluation reward: 9.25
episode: 3204	score: 7.0	memory length: 862948	epsilon:
0.009998020008555413	steps: 404	lr: 0.0001	evaluation reward: 9.15
episode: 3205	score: 8.0	memory length: 863350	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 9.11
episode: 3206	score: 10.0	memory length: 863850	epsilon:
0.009998020008555413	steps: 500	lr: 0.0001	evaluation reward: 9.13
episode: 3207	score: 7.0	memory length: 864265	epsilon:
0.009998020008555413	steps: 415	lr: 0.0001	evaluation reward: 9.14
episode: 3208	score: 6.0	memory length: 864586	epsilon:
0.009998020008555413	steps: 321	lr: 0.0001	evaluation reward: 9.1
episode: 3209	score: 9.0	memory length: 865084	epsilon:
0.009998020008555413	steps: 498	lr: 0.0001	evaluation reward: 9.14
episode: 3210	score: 7.0	memory length: 865468	epsilon:
0.009998020008555413	steps: 384	lr: 0.0001	evaluation reward: 9.16
episode: 3211	score: 4.0	memory length: 865709	epsilon:
0.009998020008555413	steps: 241	lr: 0.0001	evaluation reward: 9.13
episode: 3212	score: 8.0	memory length: 866125	epsilon:
0.009998020008555413	steps: 416	lr: 0.0001	evaluation reward: 9.13
episode: 3213	score: 6.0	memory length: 866478	epsilon:
0.009998020008555413	steps: 353	lr: 0.0001	evaluation reward: 9.12
episode: 3214	score: 15.0	memory length: 866974	epsilon:
0.009998020008555413	steps: 496	lr: 0.0001	evaluation reward: 9.21
episode: 3215	score: 14.0	memory length: 867539	epsilon:
0.009998020008555413	steps: 565	lr: 0.0001	evaluation reward: 9.24
episode: 3216	score: 6.0	memory length: 867877	epsilon:
0.009998020008555413	steps: 338	lr: 0.0001	evaluation reward: 9.21
episode: 3217	score: 11.0	memory length: 868315	epsilon:
0.009998020008555413	steps: 438	lr: 0.0001	evaluation reward: 9.29
episode: 3218	score: 7.0	memory length: 868703	epsilon:
0.009998020008555413	steps: 388	lr: 0.0001	evaluation reward: 9.26
episode: 3219	score: 5.0	memory length: 869028	epsilon:
0.009998020008555413	steps: 325	lr: 0.0001	evaluation reward: 9.16
episode: 3220	score: 10.0	memory length: 869542	epsilon:
0.009998020008555413	steps: 514	lr: 0.0001	evaluation reward: 9.18
episode: 3221	score: 8.0	memory length: 869965	epsilon:
0.009998020008555413	steps: 423	lr: 0.0001	evaluation reward: 9.14
episode: 3222	score: 9.0	memory length: 870438	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 9.14
episode: 3223	score: 6.0	memory length: 870756	epsilon:
0.009998020008555413	steps: 318	lr: 0.0001	evaluation reward: 9.12
episode: 3224	score: 14.0	memory length: 871275	epsilon:

0.009998020008555413	steps: 519	lr: 0.0001	evaluation reward: 9.16
episode: 3225	score: 7.0	memory length: 871676	epsilon:
0.009998020008555413	steps: 401	lr: 0.0001	evaluation reward: 9.13
episode: 3226	score: 7.0	memory length: 872078	epsilon:
0.009998020008555413	steps: 402	lr: 0.0001	evaluation reward: 9.17
episode: 3227	score: 5.0	memory length: 872384	epsilon:
0.009998020008555413	steps: 306	lr: 0.0001	evaluation reward: 9.11
episode: 3228	score: 10.0	memory length: 872870	epsilon:
0.009998020008555413	steps: 486	lr: 0.0001	evaluation reward: 9.16
episode: 3229	score: 6.0	memory length: 873189	epsilon:
0.009998020008555413	steps: 319	lr: 0.0001	evaluation reward: 9.08
episode: 3230	score: 7.0	memory length: 873578	epsilon:
0.009998020008555413	steps: 389	lr: 0.0001	evaluation reward: 9.06
episode: 3231	score: 6.0	memory length: 873952	epsilon:
0.009998020008555413	steps: 374	lr: 0.0001	evaluation reward: 9.07
episode: 3232	score: 8.0	memory length: 874344	epsilon:
0.009998020008555413	steps: 392	lr: 0.0001	evaluation reward: 9.05
episode: 3233	score: 7.0	memory length: 874731	epsilon:
0.009998020008555413	steps: 387	lr: 0.0001	evaluation reward: 9.06
episode: 3234	score: 11.0	memory length: 875204	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 9.11
episode: 3235	score: 10.0	memory length: 875580	epsilon:
0.009998020008555413	steps: 376	lr: 0.0001	evaluation reward: 9.12
episode: 3236	score: 11.0	memory length: 876062	epsilon:
0.009998020008555413	steps: 482	lr: 0.0001	evaluation reward: 9.12
episode: 3237	score: 9.0	memory length: 876561	epsilon:
0.009998020008555413	steps: 499	lr: 0.0001	evaluation reward: 9.12
episode: 3238	score: 7.0	memory length: 876985	epsilon:
0.009998020008555413	steps: 424	lr: 0.0001	evaluation reward: 9.14
episode: 3239	score: 9.0	memory length: 877450	epsilon:
0.009998020008555413	steps: 465	lr: 0.0001	evaluation reward: 9.14
episode: 3240	score: 8.0	memory length: 877870	epsilon:
0.009998020008555413	steps: 420	lr: 0.0001	evaluation reward: 9.15
episode: 3241	score: 8.0	memory length: 878281	epsilon:
0.009998020008555413	steps: 411	lr: 0.0001	evaluation reward: 9.16
episode: 3242	score: 7.0	memory length: 878687	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 9.14
episode: 3243	score: 9.0	memory length: 879154	epsilon:
0.009998020008555413	steps: 467	lr: 0.0001	evaluation reward: 9.14
episode: 3244	score: 4.0	memory length: 879430	epsilon:
0.009998020008555413	steps: 276	lr: 0.0001	evaluation reward: 9.09
episode: 3245	score: 7.0	memory length: 879829	epsilon:
0.009998020008555413	steps: 399	lr: 0.0001	evaluation reward: 9.04
episode: 3246	score: 8.0	memory length: 880238	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 9.02
episode: 3247	score: 4.0	memory length: 880494	epsilon:
0.009998020008555413	steps: 256	lr: 0.0001	evaluation reward: 9.01
episode: 3248	score: 8.0	memory length: 880948	epsilon:

0.009998020008555413	steps: 454	lr: 0.0001	evaluation reward: 8.96
episode: 3249	score: 5.0	memory length: 881250	epsilon:
0.009998020008555413	steps: 302	lr: 0.0001	evaluation reward: 8.91
episode: 3250	score: 8.0	memory length: 881682	epsilon:
0.009998020008555413	steps: 432	lr: 0.0001	evaluation reward: 8.89
episode: 3251	score: 7.0	memory length: 882049	epsilon:
0.009998020008555413	steps: 367	lr: 0.0001	evaluation reward: 8.86
episode: 3252	score: 7.0	memory length: 882431	epsilon:
0.009998020008555413	steps: 382	lr: 0.0001	evaluation reward: 8.89
episode: 3253	score: 9.0	memory length: 882881	epsilon:
0.009998020008555413	steps: 450	lr: 0.0001	evaluation reward: 8.89
episode: 3254	score: 9.0	memory length: 883334	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 8.89
episode: 3255	score: 11.0	memory length: 883915	epsilon:
0.009998020008555413	steps: 581	lr: 0.0001	evaluation reward: 8.94
episode: 3256	score: 6.0	memory length: 884231	epsilon:
0.009998020008555413	steps: 316	lr: 0.0001	evaluation reward: 8.92
episode: 3257	score: 5.0	memory length: 884540	epsilon:
0.009998020008555413	steps: 309	lr: 0.0001	evaluation reward: 8.85
episode: 3258	score: 8.0	memory length: 884977	epsilon:
0.009998020008555413	steps: 437	lr: 0.0001	evaluation reward: 8.87
episode: 3259	score: 13.0	memory length: 885450	epsilon:
0.009998020008555413	steps: 473	lr: 0.0001	evaluation reward: 8.91
episode: 3260	score: 9.0	memory length: 885882	epsilon:
0.009998020008555413	steps: 432	lr: 0.0001	evaluation reward: 8.92
episode: 3261	score: 8.0	memory length: 886339	epsilon:
0.009998020008555413	steps: 457	lr: 0.0001	evaluation reward: 8.95
episode: 3262	score: 5.0	memory length: 886626	epsilon:
0.009998020008555413	steps: 287	lr: 0.0001	evaluation reward: 8.83
episode: 3263	score: 9.0	memory length: 887111	epsilon:
0.009998020008555413	steps: 485	lr: 0.0001	evaluation reward: 8.82
episode: 3264	score: 15.0	memory length: 887764	epsilon:
0.009998020008555413	steps: 653	lr: 0.0001	evaluation reward: 8.83
episode: 3265	score: 5.0	memory length: 888068	epsilon:
0.009998020008555413	steps: 304	lr: 0.0001	evaluation reward: 8.79
episode: 3266	score: 7.0	memory length: 888458	epsilon:
0.009998020008555413	steps: 390	lr: 0.0001	evaluation reward: 8.8
episode: 3267	score: 13.0	memory length: 888948	epsilon:
0.009998020008555413	steps: 490	lr: 0.0001	evaluation reward: 8.86
episode: 3268	score: 17.0	memory length: 889578	epsilon:
0.009998020008555413	steps: 630	lr: 0.0001	evaluation reward: 8.95
episode: 3269	score: 13.0	memory length: 890127	epsilon:
0.009998020008555413	steps: 549	lr: 0.0001	evaluation reward: 9.01
episode: 3270	score: 8.0	memory length: 890548	epsilon:
0.009998020008555413	steps: 421	lr: 0.0001	evaluation reward: 9.02
episode: 3271	score: 11.0	memory length: 891051	epsilon:
0.009998020008555413	steps: 503	lr: 0.0001	evaluation reward: 9.08
episode: 3272	score: 9.0	memory length: 891527	epsilon:

0.009998020008555413	steps: 476	lr: 0.0001	evaluation reward: 9.1
episode: 3273	score: 17.0	memory length: 892152	epsilon:
0.009998020008555413	steps: 625	lr: 0.0001	evaluation reward: 9.17
episode: 3274	score: 8.0	memory length: 892604	epsilon:
0.009998020008555413	steps: 452	lr: 0.0001	evaluation reward: 9.11
episode: 3275	score: 8.0	memory length: 893030	epsilon:
0.009998020008555413	steps: 426	lr: 0.0001	evaluation reward: 9.13
episode: 3276	score: 9.0	memory length: 893442	epsilon:
0.009998020008555413	steps: 412	lr: 0.0001	evaluation reward: 9.12
episode: 3277	score: 8.0	memory length: 893861	epsilon:
0.009998020008555413	steps: 419	lr: 0.0001	evaluation reward: 9.11
episode: 3278	score: 7.0	memory length: 894225	epsilon:
0.009998020008555413	steps: 364	lr: 0.0001	evaluation reward: 9.12
episode: 3279	score: 6.0	memory length: 894524	epsilon:
0.009998020008555413	steps: 299	lr: 0.0001	evaluation reward: 9.08
episode: 3280	score: 8.0	memory length: 894964	epsilon:
0.009998020008555413	steps: 440	lr: 0.0001	evaluation reward: 9.03
episode: 3281	score: 6.0	memory length: 895282	epsilon:
0.009998020008555413	steps: 318	lr: 0.0001	evaluation reward: 8.94
episode: 3282	score: 8.0	memory length: 895703	epsilon:
0.009998020008555413	steps: 421	lr: 0.0001	evaluation reward: 8.95
episode: 3283	score: 6.0	memory length: 896058	epsilon:
0.009998020008555413	steps: 355	lr: 0.0001	evaluation reward: 8.85
episode: 3284	score: 7.0	memory length: 896454	epsilon:
0.009998020008555413	steps: 396	lr: 0.0001	evaluation reward: 8.82
episode: 3285	score: 11.0	memory length: 896929	epsilon:
0.009998020008555413	steps: 475	lr: 0.0001	evaluation reward: 8.83
episode: 3286	score: 12.0	memory length: 897518	epsilon:
0.009998020008555413	steps: 589	lr: 0.0001	evaluation reward: 8.87
episode: 3287	score: 10.0	memory length: 898058	epsilon:
0.009998020008555413	steps: 540	lr: 0.0001	evaluation reward: 8.81
episode: 3288	score: 9.0	memory length: 898548	epsilon:
0.009998020008555413	steps: 490	lr: 0.0001	evaluation reward: 8.78
episode: 3289	score: 10.0	memory length: 898926	epsilon:
0.009998020008555413	steps: 378	lr: 0.0001	evaluation reward: 8.81
episode: 3290	score: 7.0	memory length: 899324	epsilon:
0.009998020008555413	steps: 398	lr: 0.0001	evaluation reward: 8.76
episode: 3291	score: 8.0	memory length: 899730	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 8.7
episode: 3292	score: 8.0	memory length: 900168	epsilon:
0.009998020008555413	steps: 438	lr: 0.0001	evaluation reward: 8.71
episode: 3293	score: 9.0	memory length: 900610	epsilon:
0.009998020008555413	steps: 442	lr: 0.0001	evaluation reward: 8.69
episode: 3294	score: 9.0	memory length: 901086	epsilon:
0.009998020008555413	steps: 476	lr: 0.0001	evaluation reward: 8.73
episode: 3295	score: 9.0	memory length: 901525	epsilon:
0.009998020008555413	steps: 439	lr: 0.0001	evaluation reward: 8.73
episode: 3296	score: 7.0	memory length: 901891	epsilon:

0.009998020008555413	steps: 366	lr: 0.0001	evaluation reward: 8.68
episode: 3297	score: 10.0	memory length: 902340	epsilon:
0.009998020008555413	steps: 449	lr: 0.0001	evaluation reward: 8.71
episode: 3298	score: 11.0	memory length: 902746	epsilon:
0.009998020008555413	steps: 406	lr: 0.0001	evaluation reward: 8.6
episode: 3299	score: 9.0	memory length: 903199	epsilon:
0.009998020008555413	steps: 453	lr: 0.0001	evaluation reward: 8.58
episode: 3300	score: 8.0	memory length: 903621	epsilon:
0.009998020008555413	steps: 422	lr: 0.0001	evaluation reward: 8.51
episode: 3301	score: 10.0	memory length: 904124	epsilon:
0.009998020008555413	steps: 503	lr: 0.0001	evaluation reward: 8.48
episode: 3302	score: 5.0	memory length: 904429	epsilon:
0.009998020008555413	steps: 305	lr: 0.0001	evaluation reward: 8.43
episode: 3303	score: 11.0	memory length: 904979	epsilon:
0.009998020008555413	steps: 550	lr: 0.0001	evaluation reward: 8.48
episode: 3304	score: 8.0	memory length: 905391	epsilon:
0.009998020008555413	steps: 412	lr: 0.0001	evaluation reward: 8.49
episode: 3305	score: 8.0	memory length: 905809	epsilon:
0.009998020008555413	steps: 418	lr: 0.0001	evaluation reward: 8.49
episode: 3306	score: 6.0	memory length: 906180	epsilon:
0.009998020008555413	steps: 371	lr: 0.0001	evaluation reward: 8.45
episode: 3307	score: 6.0	memory length: 906574	epsilon:
0.009998020008555413	steps: 394	lr: 0.0001	evaluation reward: 8.44
episode: 3308	score: 16.0	memory length: 907035	epsilon:
0.009998020008555413	steps: 461	lr: 0.0001	evaluation reward: 8.54
episode: 3309	score: 4.0	memory length: 907273	epsilon:
0.009998020008555413	steps: 238	lr: 0.0001	evaluation reward: 8.49
episode: 3310	score: 11.0	memory length: 907814	epsilon:
0.009998020008555413	steps: 541	lr: 0.0001	evaluation reward: 8.53
episode: 3311	score: 9.0	memory length: 908223	epsilon:
0.009998020008555413	steps: 409	lr: 0.0001	evaluation reward: 8.58
episode: 3312	score: 10.0	memory length: 908715	epsilon:
0.009998020008555413	steps: 492	lr: 0.0001	evaluation reward: 8.6
episode: 3313	score: 4.0	memory length: 908971	epsilon:
0.009998020008555413	steps: 256	lr: 0.0001	evaluation reward: 8.58
episode: 3314	score: 8.0	memory length: 909339	epsilon:
0.009998020008555413	steps: 368	lr: 0.0001	evaluation reward: 8.51
episode: 3315	score: 10.0	memory length: 909815	epsilon:
0.009998020008555413	steps: 476	lr: 0.0001	evaluation reward: 8.47
episode: 3316	score: 7.0	memory length: 910164	epsilon:
0.009998020008555413	steps: 349	lr: 0.0001	evaluation reward: 8.48
episode: 3317	score: 6.0	memory length: 910518	epsilon:
0.009998020008555413	steps: 354	lr: 0.0001	evaluation reward: 8.43
episode: 3318	score: 14.0	memory length: 911022	epsilon:
0.009998020008555413	steps: 504	lr: 0.0001	evaluation reward: 8.5
episode: 3319	score: 13.0	memory length: 911516	epsilon:
0.009998020008555413	steps: 494	lr: 0.0001	evaluation reward: 8.58
episode: 3320	score: 9.0	memory length: 912015	epsilon:

```

0.009998020008555413    steps: 499    lr: 0.0001    evaluation reward: 8.57
episode: 3321    score: 9.0    memory length: 912507    epsilon:
0.009998020008555413    steps: 492    lr: 0.0001    evaluation reward: 8.58
episode: 3322    score: 9.0    memory length: 912961    epsilon:
0.009998020008555413    steps: 454    lr: 0.0001    evaluation reward: 8.58
episode: 3323    score: 9.0    memory length: 913454    epsilon:
0.009998020008555413    steps: 493    lr: 0.0001    evaluation reward: 8.61
episode: 3324    score: 11.0    memory length: 913951    epsilon:
0.009998020008555413    steps: 497    lr: 0.0001    evaluation reward: 8.58
episode: 3325    score: 8.0    memory length: 914367    epsilon:
0.009998020008555413    steps: 416    lr: 0.0001    evaluation reward: 8.59
episode: 3326    score: 8.0    memory length: 914789    epsilon:
0.009998020008555413    steps: 422    lr: 0.0001    evaluation reward: 8.6
episode: 3327    score: 7.0    memory length: 915192    epsilon:
0.009998020008555413    steps: 403    lr: 0.0001    evaluation reward: 8.62
episode: 3328    score: 5.0    memory length: 915479    epsilon:
0.009998020008555413    steps: 287    lr: 0.0001    evaluation reward: 8.57
episode: 3329    score: 10.0    memory length: 915996    epsilon:
0.009998020008555413    steps: 517    lr: 0.0001    evaluation reward: 8.61
episode: 3330    score: 7.0    memory length: 916389    epsilon:
0.009998020008555413    steps: 393    lr: 0.0001    evaluation reward: 8.61
episode: 3331    score: 9.0    memory length: 916842    epsilon:
0.009998020008555413    steps: 453    lr: 0.0001    evaluation reward: 8.64
episode: 3332    score: 5.0    memory length: 917125    epsilon:
0.009998020008555413    steps: 283    lr: 0.0001    evaluation reward: 8.61
episode: 3333    score: 8.0    memory length: 917546    epsilon:
0.009998020008555413    steps: 421    lr: 0.0001    evaluation reward: 8.62
episode: 3334    score: 6.0    memory length: 917888    epsilon:
0.009998020008555413    steps: 342    lr: 0.0001    evaluation reward: 8.57
episode: 3335    score: 10.0    memory length: 918455    epsilon:
0.009998020008555413    steps: 567    lr: 0.0001    evaluation reward: 8.57
episode: 3336    score: 11.0    memory length: 918962    epsilon:
0.009998020008555413    steps: 507    lr: 0.0001    evaluation reward: 8.57
episode: 3337    score: 7.0    memory length: 919333    epsilon:
0.009998020008555413    steps: 371    lr: 0.0001    evaluation reward: 8.55

```

```

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

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

```



```

life = number_lives

get_init_state(history, state)

#agent.scheduler.step()

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

    # Perform a fire action if ball is no longer on screen to continue onto
    ↪ next life
    if step > 1 and len(np.unique(next_state[:189] == state[:189])) < 2:
        action = 0
    else:
        action = agent.get_action(np.float32(history[:4, :, :]) / 255.)
    state = next_state
    next_state, reward, done, info = env.step(action + 1)

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

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

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

    if done:
        evaluation_reward.append(score)
        rewards.append(np.mean(evaluation_reward))
        episodes.append(e)
        pylab.plot(episodes, rewards, 'b')
        pylab.xlabel('Episodes')
        pylab.ylabel('Rewards')
        pylab.title('Episodes vs Reward')

```

```

        pylab.savefig("./save_graph/breakout_ddqn.png") # save graph for
→training visualization

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

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

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

```

```

episode: 0   score: 2.0   memory length: 197   epsilon: 1.0   steps: 197   lr:
0.0001      evaluation reward: 2.0
episode: 1   score: 0.0   memory length: 320   epsilon: 1.0   steps: 123   lr:
0.0001      evaluation reward: 1.0
episode: 2   score: 2.0   memory length: 517   epsilon: 1.0   steps: 197   lr:
0.0001      evaluation reward: 1.3333333333333333
episode: 3   score: 1.0   memory length: 688   epsilon: 1.0   steps: 171   lr:
0.0001      evaluation reward: 1.25
episode: 4   score: 1.0   memory length: 839   epsilon: 1.0   steps: 151   lr:
0.0001      evaluation reward: 1.2
episode: 5   score: 1.0   memory length: 1008   epsilon: 1.0   steps: 169
lr: 0.0001   evaluation reward: 1.1666666666666667
episode: 6   score: 0.0   memory length: 1131   epsilon: 1.0   steps: 123
lr: 0.0001   evaluation reward: 1.0
episode: 7   score: 1.0   memory length: 1303   epsilon: 1.0   steps: 172
lr: 0.0001   evaluation reward: 1.0
episode: 8   score: 0.0   memory length: 1425   epsilon: 1.0   steps: 122
lr: 0.0001   evaluation reward: 0.8888888888888888
episode: 9   score: 6.0   memory length: 1768   epsilon: 1.0   steps: 343
lr: 0.0001   evaluation reward: 1.4
episode: 10  score: 2.0   memory length: 1965   epsilon: 1.0   steps: 197
lr: 0.0001   evaluation reward: 1.4545454545454546
episode: 11  score: 2.0   memory length: 2182   epsilon: 1.0   steps: 217
lr: 0.0001   evaluation reward: 1.5
episode: 12  score: 3.0   memory length: 2407   epsilon: 1.0   steps: 225
lr: 0.0001   evaluation reward: 1.6153846153846154
episode: 13  score: 3.0   memory length: 2671   epsilon: 1.0   steps: 264
lr: 0.0001   evaluation reward: 1.7142857142857142
episode: 14  score: 3.0   memory length: 2900   epsilon: 1.0   steps: 229

```

lr: 0.0001	evaluation reward: 1.8			
episode: 15	score: 0.0	memory length: 3023	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.6875			
episode: 16	score: 1.0	memory length: 3174	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.6470588235294117			
episode: 17	score: 2.0	memory length: 3372	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6666666666666667			
episode: 18	score: 2.0	memory length: 3570	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6842105263157894			
episode: 19	score: 1.0	memory length: 3741	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.65			
episode: 20	score: 0.0	memory length: 3864	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5714285714285714			
episode: 21	score: 2.0	memory length: 4061	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.5909090909090908			
episode: 22	score: 0.0	memory length: 4184	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5217391304347827			
episode: 23	score: 3.0	memory length: 4410	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.5833333333333333			
episode: 24	score: 2.0	memory length: 4610	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.6			
episode: 25	score: 0.0	memory length: 4733	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5384615384615385			
episode: 26	score: 2.0	memory length: 4931	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5555555555555556			
episode: 27	score: 1.0	memory length: 5099	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.5357142857142858			
episode: 28	score: 3.0	memory length: 5325	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.5862068965517242			
episode: 29	score: 0.0	memory length: 5448	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5333333333333334			
episode: 30	score: 3.0	memory length: 5715	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.5806451612903225			
episode: 31	score: 2.0	memory length: 5913	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.59375			
episode: 32	score: 2.0	memory length: 6111	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.606060606060606			
episode: 33	score: 2.0	memory length: 6309	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6176470588235294			
episode: 34	score: 3.0	memory length: 6558	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.6571428571428573			
episode: 35	score: 2.0	memory length: 6773	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.6666666666666667			
episode: 36	score: 1.0	memory length: 6942	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.6486486486486487			
episode: 37	score: 1.0	memory length: 7111	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.631578947368421			
episode: 38	score: 0.0	memory length: 7234	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.5897435897435896		
episode: 39	score: 2.0	memory length: 7414	epsilon: 1.0 steps: 180
lr: 0.0001	evaluation reward: 1.6		
episode: 40	score: 2.0	memory length: 7612	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.6097560975609757		
episode: 41	score: 1.0	memory length: 7782	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.5952380952380953		
episode: 42	score: 3.0	memory length: 8010	epsilon: 1.0 steps: 228
lr: 0.0001	evaluation reward: 1.627906976744186		
episode: 43	score: 2.0	memory length: 8228	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.6363636363636365		
episode: 44	score: 3.0	memory length: 8474	epsilon: 1.0 steps: 246
lr: 0.0001	evaluation reward: 1.6666666666666667		
episode: 45	score: 0.0	memory length: 8597	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.6304347826086956		
episode: 46	score: 2.0	memory length: 8795	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.6382978723404256		
episode: 47	score: 0.0	memory length: 8918	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.6041666666666667		
episode: 48	score: 0.0	memory length: 9041	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5714285714285714		
episode: 49	score: 0.0	memory length: 9163	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.54		
episode: 50	score: 3.0	memory length: 9407	epsilon: 1.0 steps: 244
lr: 0.0001	evaluation reward: 1.5686274509803921		
episode: 51	score: 0.0	memory length: 9529	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.5384615384615385		
episode: 52	score: 0.0	memory length: 9651	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.509433962264151		
episode: 53	score: 0.0	memory length: 9774	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4814814814814814		
episode: 54	score: 5.0	memory length: 10084	epsilon: 1.0 steps: 310
lr: 0.0001	evaluation reward: 1.5454545454545454		
episode: 55	score: 1.0	memory length: 10254	epsilon: 1.0 steps: 170
lr: 0.0001	evaluation reward: 1.5357142857142858		
episode: 56	score: 2.0	memory length: 10451	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.543859649122807		
episode: 57	score: 0.0	memory length: 10573	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.5172413793103448		
episode: 58	score: 2.0	memory length: 10771	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.5254237288135593		
episode: 59	score: 2.0	memory length: 10972	epsilon: 1.0 steps: 201
lr: 0.0001	evaluation reward: 1.5333333333333334		
episode: 60	score: 1.0	memory length: 11141	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.5245901639344261		
episode: 61	score: 0.0	memory length: 11264	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 62	score: 3.0	memory length: 11510	epsilon: 1.0 steps: 246

lr: 0.0001	evaluation reward: 1.5238095238095237			
episode: 63	score: 1.0	memory length: 11678	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.515625			
episode: 64	score: 1.0	memory length: 11848	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.5076923076923077			
episode: 65	score: 0.0	memory length: 11971	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4848484848484849			
episode: 66	score: 4.0	memory length: 12247	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.5223880597014925			
episode: 67	score: 0.0	memory length: 12370	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5			
episode: 68	score: 2.0	memory length: 12567	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.5072463768115942			
episode: 69	score: 0.0	memory length: 12689	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4857142857142858			
episode: 70	score: 1.0	memory length: 12839	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4788732394366197			
episode: 71	score: 2.0	memory length: 13018	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.4861111111111112			
episode: 72	score: 0.0	memory length: 13140	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4657534246575343			
episode: 73	score: 0.0	memory length: 13263	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.445945945945946			
episode: 74	score: 3.0	memory length: 13488	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.4666666666666666			
episode: 75	score: 1.0	memory length: 13657	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.4605263157894737			
episode: 76	score: 0.0	memory length: 13779	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4415584415584415			
episode: 77	score: 0.0	memory length: 13902	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4230769230769231			
episode: 78	score: 2.0	memory length: 14102	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.4303797468354431			
episode: 79	score: 1.0	memory length: 14272	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.425			
episode: 80	score: 2.0	memory length: 14490	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.4320987654320987			
episode: 81	score: 3.0	memory length: 14736	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.451219512195122			
episode: 82	score: 1.0	memory length: 14905	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.4457831325301205			
episode: 83	score: 1.0	memory length: 15055	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4404761904761905			
episode: 84	score: 2.0	memory length: 15272	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.4470588235294117			
episode: 85	score: 5.0	memory length: 15578	epsilon: 1.0	steps: 306
lr: 0.0001	evaluation reward: 1.4883720930232558			
episode: 86	score: 2.0	memory length: 15777	epsilon: 1.0	steps: 199

lr: 0.0001	evaluation reward: 1.4942528735632183		
episode: 87	score: 2.0	memory length: 15975	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.5		
episode: 88	score: 6.0	memory length: 16304	epsilon: 1.0 steps: 329
lr: 0.0001	evaluation reward: 1.550561797752809		
episode: 89	score: 1.0	memory length: 16473	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.5444444444444445		
episode: 90	score: 0.0	memory length: 16595	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.5274725274725274		
episode: 91	score: 1.0	memory length: 16764	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.5217391304347827		
episode: 92	score: 0.0	memory length: 16887	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5053763440860215		
episode: 93	score: 0.0	memory length: 17010	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4893617021276595		
episode: 94	score: 1.0	memory length: 17181	epsilon: 1.0 steps: 171
lr: 0.0001	evaluation reward: 1.4842105263157894		
episode: 95	score: 1.0	memory length: 17332	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.4791666666666667		
episode: 96	score: 0.0	memory length: 17454	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.4639175257731958		
episode: 97	score: 2.0	memory length: 17672	epsilon: 1.0 steps: 218
lr: 0.0001	evaluation reward: 1.469387755102041		
episode: 98	score: 2.0	memory length: 17869	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.4747474747474747		
episode: 99	score: 0.0	memory length: 17991	epsilon: 1.0 steps: 122
lr: 0.0001	evaluation reward: 1.46		
episode: 100	score: 4.0	memory length: 18287	epsilon: 1.0 steps: 296
lr: 0.0001	evaluation reward: 1.48		
episode: 101	score: 0.0	memory length: 18410	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 102	score: 2.0	memory length: 18607	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.48		
episode: 103	score: 1.0	memory length: 18776	epsilon: 1.0 steps: 169
lr: 0.0001	evaluation reward: 1.48		
episode: 104	score: 2.0	memory length: 18973	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.49		
episode: 105	score: 1.0	memory length: 19123	epsilon: 1.0 steps: 150
lr: 0.0001	evaluation reward: 1.49		
episode: 106	score: 2.0	memory length: 19320	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.51		
episode: 107	score: 0.0	memory length: 19443	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 108	score: 2.0	memory length: 19641	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.52		
episode: 109	score: 2.0	memory length: 19839	epsilon: 1.0 steps: 198
lr: 0.0001	evaluation reward: 1.48		
episode: 110	score: 2.0	memory length: 20037	epsilon: 1.0 steps: 198

lr: 0.0001	evaluation reward: 1.48		
episode: 111	score: 1.0 memory length: 20188	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.47		
episode: 112	score: 1.0 memory length: 20356	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.45		
episode: 113	score: 1.0 memory length: 20526	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.43		
episode: 114	score: 2.0 memory length: 20724	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.42		
episode: 115	score: 0.0 memory length: 20846	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.42		
episode: 116	score: 0.0 memory length: 20969	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 117	score: 0.0 memory length: 21091	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39		
episode: 118	score: 0.0 memory length: 21213	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.37		
episode: 119	score: 0.0 memory length: 21336	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36		
episode: 120	score: 2.0 memory length: 21552	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.38		
episode: 121	score: 2.0 memory length: 21768	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.38		
episode: 122	score: 2.0 memory length: 21966	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4		
episode: 123	score: 3.0 memory length: 22192	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.4		
episode: 124	score: 5.0 memory length: 22519	epsilon: 1.0	steps: 327
lr: 0.0001	evaluation reward: 1.43		
episode: 125	score: 2.0 memory length: 22717	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.45		
episode: 126	score: 1.0 memory length: 22867	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.44		
episode: 127	score: 0.0 memory length: 22989	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43		
episode: 128	score: 2.0 memory length: 23189	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.42		
episode: 129	score: 1.0 memory length: 23340	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.43		
episode: 130	score: 2.0 memory length: 23557	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.42		
episode: 131	score: 1.0 memory length: 23727	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.41		
episode: 132	score: 0.0 memory length: 23850	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 133	score: 0.0 memory length: 23973	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37		
episode: 134	score: 0.0 memory length: 24095	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.34		
episode: 135	score: 2.0 memory length: 24292	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.34		
episode: 136	score: 1.0 memory length: 24443	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.34		
episode: 137	score: 1.0 memory length: 24594	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.34		
episode: 138	score: 2.0 memory length: 24791	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.36		
episode: 139	score: 0.0 memory length: 24913	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.34		
episode: 140	score: 1.0 memory length: 25064	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.33		
episode: 141	score: 0.0 memory length: 25186	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.32		
episode: 142	score: 2.0 memory length: 25384	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.31		
episode: 143	score: 2.0 memory length: 25583	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.31		
episode: 144	score: 0.0 memory length: 25706	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.28		
episode: 145	score: 2.0 memory length: 25888	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.3		
episode: 146	score: 1.0 memory length: 26057	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.29		
episode: 147	score: 1.0 memory length: 26229	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.3		
episode: 148	score: 3.0 memory length: 26455	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.33		
episode: 149	score: 0.0 memory length: 26578	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 150	score: 3.0 memory length: 26839	epsilon: 1.0	steps: 261
lr: 0.0001	evaluation reward: 1.33		
episode: 151	score: 3.0 memory length: 27085	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.36		
episode: 152	score: 1.0 memory length: 27235	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.37		
episode: 153	score: 0.0 memory length: 27358	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37		
episode: 154	score: 0.0 memory length: 27481	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 155	score: 2.0 memory length: 27682	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.33		
episode: 156	score: 2.0 memory length: 27898	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.33		
episode: 157	score: 3.0 memory length: 28145	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.36		
episode: 158	score: 3.0 memory length: 28354	epsilon: 1.0	steps: 209

lr: 0.0001	evaluation reward: 1.37		
episode: 159	score: 2.0 memory length: 28554	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.37		
episode: 160	score: 2.0 memory length: 28773	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.38		
episode: 161	score: 2.0 memory length: 28952	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.4		
episode: 162	score: 2.0 memory length: 29170	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.39		
episode: 163	score: 2.0 memory length: 29368	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4		
episode: 164	score: 2.0 memory length: 29548	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.41		
episode: 165	score: 0.0 memory length: 29671	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 166	score: 2.0 memory length: 29869	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.39		
episode: 167	score: 1.0 memory length: 30039	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.4		
episode: 168	score: 0.0 memory length: 30162	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38		
episode: 169	score: 0.0 memory length: 30285	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.38		
episode: 170	score: 2.0 memory length: 30501	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.39		
episode: 171	score: 1.0 memory length: 30669	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.38		
episode: 172	score: 2.0 memory length: 30868	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.4		
episode: 173	score: 2.0 memory length: 31086	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.42		
episode: 174	score: 0.0 memory length: 31208	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39		
episode: 175	score: 2.0 memory length: 31390	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.4		
episode: 176	score: 0.0 memory length: 31512	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4		
episode: 177	score: 0.0 memory length: 31635	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 178	score: 1.0 memory length: 31803	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.39		
episode: 179	score: 0.0 memory length: 31925	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38		
episode: 180	score: 2.0 memory length: 32140	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.38		
episode: 181	score: 2.0 memory length: 32338	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.37		
episode: 182	score: 1.0 memory length: 32507	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.37		
episode: 183	score: 0.0 memory length: 32630	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36		
episode: 184	score: 0.0 memory length: 32752	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.34		
episode: 185	score: 0.0 memory length: 32875	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.29		
episode: 186	score: 0.0 memory length: 32997	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.27		
episode: 187	score: 2.0 memory length: 33194	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.27		
episode: 188	score: 12.0 memory length: 33559	epsilon: 1.0	steps: 365
lr: 0.0001	evaluation reward: 1.33		
episode: 189	score: 2.0 memory length: 33775	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.34		
episode: 190	score: 2.0 memory length: 33973	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.36		
episode: 191	score: 2.0 memory length: 34191	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.37		
episode: 192	score: 3.0 memory length: 34453	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.4		
episode: 193	score: 0.0 memory length: 34576	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 194	score: 3.0 memory length: 34806	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.42		
episode: 195	score: 0.0 memory length: 34929	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 196	score: 0.0 memory length: 35051	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41		
episode: 197	score: 2.0 memory length: 35249	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.41		
episode: 198	score: 1.0 memory length: 35399	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4		
episode: 199	score: 9.0 memory length: 35798	epsilon: 1.0	steps: 399
lr: 0.0001	evaluation reward: 1.49		
episode: 200	score: 0.0 memory length: 35920	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.45		
episode: 201	score: 1.0 memory length: 36089	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.46		
episode: 202	score: 6.0 memory length: 36464	epsilon: 1.0	steps: 375
lr: 0.0001	evaluation reward: 1.5		
episode: 203	score: 1.0 memory length: 36633	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.5		
episode: 204	score: 1.0 memory length: 36805	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.49		
episode: 205	score: 1.0 memory length: 36956	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.49		
episode: 206	score: 2.0 memory length: 37174	epsilon: 1.0	steps: 218

lr: 0.0001	evaluation reward: 1.49		
episode: 207	score: 2.0 memory length: 37372	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.51		
episode: 208	score: 2.0 memory length: 37571	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.51		
episode: 209	score: 0.0 memory length: 37694	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 210	score: 1.0 memory length: 37845	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48		
episode: 211	score: 2.0 memory length: 38063	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.49		
episode: 212	score: 3.0 memory length: 38290	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.51		
episode: 213	score: 2.0 memory length: 38505	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.52		
episode: 214	score: 2.0 memory length: 38685	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.52		
episode: 215	score: 1.0 memory length: 38854	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.53		
episode: 216	score: 2.0 memory length: 39051	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.55		
episode: 217	score: 0.0 memory length: 39173	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.55		
episode: 218	score: 0.0 memory length: 39295	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.55		
episode: 219	score: 3.0 memory length: 39524	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.58		
episode: 220	score: 1.0 memory length: 39695	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.57		
episode: 221	score: 1.0 memory length: 39867	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.56		
episode: 222	score: 3.0 memory length: 40096	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.57		
episode: 223	score: 0.0 memory length: 40219	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.54		
episode: 224	score: 4.0 memory length: 40476	epsilon: 1.0	steps: 257
lr: 0.0001	evaluation reward: 1.53		
episode: 225	score: 0.0 memory length: 40599	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.51		
episode: 226	score: 1.0 memory length: 40768	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.51		
episode: 227	score: 0.0 memory length: 40890	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.51		
episode: 228	score: 4.0 memory length: 41206	epsilon: 1.0	steps: 316
lr: 0.0001	evaluation reward: 1.53		
episode: 229	score: 0.0 memory length: 41329	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 230	score: 0.0 memory length: 41451	epsilon: 1.0	steps: 122

lr: 0.0001	evaluation reward: 1.5		
episode: 231	score: 2.0 memory length: 41649	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.51		
episode: 232	score: 4.0 memory length: 41962	epsilon: 1.0	steps: 313
lr: 0.0001	evaluation reward: 1.55		
episode: 233	score: 5.0 memory length: 42329	epsilon: 1.0	steps: 367
lr: 0.0001	evaluation reward: 1.6		
episode: 234	score: 1.0 memory length: 42479	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.61		
episode: 235	score: 3.0 memory length: 42722	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.62		
episode: 236	score: 6.0 memory length: 43085	epsilon: 1.0	steps: 363
lr: 0.0001	evaluation reward: 1.67		
episode: 237	score: 3.0 memory length: 43296	epsilon: 1.0	steps: 211
lr: 0.0001	evaluation reward: 1.69		
episode: 238	score: 4.0 memory length: 43589	epsilon: 1.0	steps: 293
lr: 0.0001	evaluation reward: 1.71		
episode: 239	score: 1.0 memory length: 43760	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.72		
episode: 240	score: 4.0 memory length: 44057	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.75		
episode: 241	score: 0.0 memory length: 44179	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.75		
episode: 242	score: 2.0 memory length: 44397	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.75		
episode: 243	score: 3.0 memory length: 44643	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.76		
episode: 244	score: 3.0 memory length: 44891	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.79		
episode: 245	score: 2.0 memory length: 45110	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.79		
episode: 246	score: 2.0 memory length: 45308	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.8		
episode: 247	score: 3.0 memory length: 45534	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.82		
episode: 248	score: 0.0 memory length: 45656	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.79		
episode: 249	score: 0.0 memory length: 45779	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.79		
episode: 250	score: 2.0 memory length: 45995	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.78		
episode: 251	score: 3.0 memory length: 46238	epsilon: 1.0	steps: 243
lr: 0.0001	evaluation reward: 1.78		
episode: 252	score: 1.0 memory length: 46389	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.78		
episode: 253	score: 2.0 memory length: 46586	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.8		
episode: 254	score: 2.0 memory length: 46783	epsilon: 1.0	steps: 197

lr: 0.0001	evaluation reward: 1.82		
episode: 255	score: 1.0 memory length: 46934	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.81		
episode: 256	score: 2.0 memory length: 47132	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.81		
episode: 257	score: 2.0 memory length: 47330	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.8		
episode: 258	score: 3.0 memory length: 47594	epsilon: 1.0	steps: 264
lr: 0.0001	evaluation reward: 1.8		
episode: 259	score: 1.0 memory length: 47744	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.79		
episode: 260	score: 3.0 memory length: 47972	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.8		
episode: 261	score: 2.0 memory length: 48151	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.8		
episode: 262	score: 1.0 memory length: 48321	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.79		
episode: 263	score: 0.0 memory length: 48443	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.77		
episode: 264	score: 0.0 memory length: 48566	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.75		
episode: 265	score: 2.0 memory length: 48764	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.77		
episode: 266	score: 1.0 memory length: 48915	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.76		
episode: 267	score: 0.0 memory length: 49038	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.75		
episode: 268	score: 0.0 memory length: 49161	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.75		
episode: 269	score: 0.0 memory length: 49283	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.75		
episode: 270	score: 0.0 memory length: 49405	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.73		
episode: 271	score: 1.0 memory length: 49573	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.73		
episode: 272	score: 0.0 memory length: 49696	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.71		
episode: 273	score: 4.0 memory length: 49988	epsilon: 1.0	steps: 292
lr: 0.0001	evaluation reward: 1.73		
episode: 274	score: 0.0 memory length: 50110	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.73		
episode: 275	score: 3.0 memory length: 50377	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.74		
episode: 276	score: 0.0 memory length: 50500	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.74		
episode: 277	score: 0.0 memory length: 50622	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.74		
episode: 278	score: 2.0 memory length: 50820	epsilon: 1.0	steps: 198

lr: 0.0001	evaluation reward: 1.75		
episode: 279	score: 0.0 memory length: 50943	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.75		
episode: 280	score: 3.0 memory length: 51171	epsilon: 1.0	steps: 228
lr: 0.0001	evaluation reward: 1.76		
episode: 281	score: 2.0 memory length: 51386	epsilon: 1.0	steps: 215
lr: 0.0001	evaluation reward: 1.76		
episode: 282	score: 1.0 memory length: 51537	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.76		
episode: 283	score: 2.0 memory length: 51734	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.78		
episode: 284	score: 1.0 memory length: 51885	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.79		
episode: 285	score: 0.0 memory length: 52008	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.79		
episode: 286	score: 2.0 memory length: 52187	epsilon: 1.0	steps: 179
lr: 0.0001	evaluation reward: 1.81		
episode: 287	score: 1.0 memory length: 52338	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.8		
episode: 288	score: 0.0 memory length: 52460	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.68		
episode: 289	score: 1.0 memory length: 52628	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.67		
episode: 290	score: 3.0 memory length: 52894	epsilon: 1.0	steps: 266
lr: 0.0001	evaluation reward: 1.68		
episode: 291	score: 2.0 memory length: 53092	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.68		
episode: 292	score: 1.0 memory length: 53262	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.66		
episode: 293	score: 0.0 memory length: 53384	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.66		
episode: 294	score: 3.0 memory length: 53630	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.66		
episode: 295	score: 1.0 memory length: 53781	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.67		
episode: 296	score: 2.0 memory length: 53961	epsilon: 1.0	steps: 180
lr: 0.0001	evaluation reward: 1.69		
episode: 297	score: 4.0 memory length: 54279	epsilon: 1.0	steps: 318
lr: 0.0001	evaluation reward: 1.71		
episode: 298	score: 0.0 memory length: 54402	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.7		
episode: 299	score: 2.0 memory length: 54599	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.63		
episode: 300	score: 2.0 memory length: 54820	epsilon: 1.0	steps: 221
lr: 0.0001	evaluation reward: 1.65		
episode: 301	score: 1.0 memory length: 54971	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.65		
episode: 302	score: 2.0 memory length: 55168	epsilon: 1.0	steps: 197

lr: 0.0001	evaluation reward: 1.61		
episode: 303	score: 1.0 memory length: 55318	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.61		
episode: 304	score: 3.0 memory length: 55544	epsilon: 1.0	steps: 226
lr: 0.0001	evaluation reward: 1.63		
episode: 305	score: 3.0 memory length: 55795	epsilon: 1.0	steps: 251
lr: 0.0001	evaluation reward: 1.65		
episode: 306	score: 0.0 memory length: 55918	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.63		
episode: 307	score: 0.0 memory length: 56041	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61		
episode: 308	score: 0.0 memory length: 56164	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.59		
episode: 309	score: 1.0 memory length: 56333	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.6		
episode: 310	score: 1.0 memory length: 56502	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.6		
episode: 311	score: 4.0 memory length: 56799	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.62		
episode: 312	score: 1.0 memory length: 56950	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.6		
episode: 313	score: 4.0 memory length: 57264	epsilon: 1.0	steps: 314
lr: 0.0001	evaluation reward: 1.62		
episode: 314	score: 2.0 memory length: 57461	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.62		
episode: 315	score: 2.0 memory length: 57658	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.63		
episode: 316	score: 0.0 memory length: 57781	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61		
episode: 317	score: 0.0 memory length: 57903	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.61		
episode: 318	score: 0.0 memory length: 58026	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.61		
episode: 319	score: 2.0 memory length: 58208	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.6		
episode: 320	score: 4.0 memory length: 58525	epsilon: 1.0	steps: 317
lr: 0.0001	evaluation reward: 1.63		
episode: 321	score: 1.0 memory length: 58696	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.63		
episode: 322	score: 1.0 memory length: 58864	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.61		
episode: 323	score: 1.0 memory length: 59033	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.62		
episode: 324	score: 2.0 memory length: 59231	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.6		
episode: 325	score: 2.0 memory length: 59429	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.62		
episode: 326	score: 1.0 memory length: 59598	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.62		
episode: 327	score: 2.0 memory length: 59817	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.64		
episode: 328	score: 3.0 memory length: 60028	epsilon: 1.0	steps: 211
lr: 0.0001	evaluation reward: 1.63		
episode: 329	score: 2.0 memory length: 60226	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.65		
episode: 330	score: 3.0 memory length: 60476	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.68		
episode: 331	score: 0.0 memory length: 60598	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.66		
episode: 332	score: 2.0 memory length: 60816	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.64		
episode: 333	score: 2.0 memory length: 61013	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.61		
episode: 334	score: 1.0 memory length: 61181	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.61		
episode: 335	score: 1.0 memory length: 61351	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.59		
episode: 336	score: 0.0 memory length: 61474	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53		
episode: 337	score: 0.0 memory length: 61597	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 338	score: 3.0 memory length: 61844	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.49		
episode: 339	score: 3.0 memory length: 62092	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.51		
episode: 340	score: 0.0 memory length: 62215	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 341	score: 3.0 memory length: 62425	epsilon: 1.0	steps: 210
lr: 0.0001	evaluation reward: 1.5		
episode: 342	score: 0.0 memory length: 62547	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.48		
episode: 343	score: 1.0 memory length: 62697	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.46		
episode: 344	score: 1.0 memory length: 62867	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.44		
episode: 345	score: 0.0 memory length: 62989	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.42		
episode: 346	score: 0.0 memory length: 63112	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 347	score: 4.0 memory length: 63374	epsilon: 1.0	steps: 262
lr: 0.0001	evaluation reward: 1.41		
episode: 348	score: 0.0 memory length: 63496	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.41		
episode: 349	score: 0.0 memory length: 63619	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 350	score: 2.0 memory length: 63837	epsilon: 1.0	steps: 218

lr: 0.0001	evaluation reward: 1.41		
episode: 351	score: 1.0 memory length: 64006	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39		
episode: 352	score: 5.0 memory length: 64330	epsilon: 1.0	steps: 324
lr: 0.0001	evaluation reward: 1.43		
episode: 353	score: 4.0 memory length: 64643	epsilon: 1.0	steps: 313
lr: 0.0001	evaluation reward: 1.45		
episode: 354	score: 1.0 memory length: 64794	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.44		
episode: 355	score: 4.0 memory length: 65091	epsilon: 1.0	steps: 297
lr: 0.0001	evaluation reward: 1.47		
episode: 356	score: 2.0 memory length: 65289	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 357	score: 1.0 memory length: 65440	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.46		
episode: 358	score: 0.0 memory length: 65562	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.43		
episode: 359	score: 0.0 memory length: 65685	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 360	score: 0.0 memory length: 65807	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39		
episode: 361	score: 0.0 memory length: 65930	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37		
episode: 362	score: 0.0 memory length: 66053	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36		
episode: 363	score: 1.0 memory length: 66204	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.37		
episode: 364	score: 3.0 memory length: 66452	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.4		
episode: 365	score: 1.0 memory length: 66602	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.39		
episode: 366	score: 1.0 memory length: 66770	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.39		
episode: 367	score: 2.0 memory length: 66986	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.41		
episode: 368	score: 1.0 memory length: 67155	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.42		
episode: 369	score: 0.0 memory length: 67278	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 370	score: 4.0 memory length: 67554	epsilon: 1.0	steps: 276
lr: 0.0001	evaluation reward: 1.46		
episode: 371	score: 2.0 memory length: 67752	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.47		
episode: 372	score: 3.0 memory length: 67999	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.5		
episode: 373	score: 0.0 memory length: 68122	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.46		
episode: 374	score: 1.0 memory length: 68290	epsilon: 1.0	steps: 168

lr: 0.0001	evaluation reward: 1.47		
episode: 375	score: 3.0 memory length: 68536	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.47		
episode: 376	score: 1.0 memory length: 68707	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.48		
episode: 377	score: 2.0 memory length: 68904	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.5		
episode: 378	score: 1.0 memory length: 69055	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.49		
episode: 379	score: 5.0 memory length: 69371	epsilon: 1.0	steps: 316
lr: 0.0001	evaluation reward: 1.54		
episode: 380	score: 1.0 memory length: 69539	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.52		
episode: 381	score: 3.0 memory length: 69806	epsilon: 1.0	steps: 267
lr: 0.0001	evaluation reward: 1.53		
episode: 382	score: 0.0 memory length: 69929	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 383	score: 1.0 memory length: 70101	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.51		
episode: 384	score: 2.0 memory length: 70299	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.52		
episode: 385	score: 1.0 memory length: 70468	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.53		
episode: 386	score: 2.0 memory length: 70666	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.53		
episode: 387	score: 1.0 memory length: 70817	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.53		
episode: 388	score: 0.0 memory length: 70940	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.53		
episode: 389	score: 0.0 memory length: 71063	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.52		
episode: 390	score: 0.0 memory length: 71185	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.49		
episode: 391	score: 1.0 memory length: 71336	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.48		
episode: 392	score: 0.0 memory length: 71459	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 393	score: 4.0 memory length: 71732	epsilon: 1.0	steps: 273
lr: 0.0001	evaluation reward: 1.51		
episode: 394	score: 2.0 memory length: 71930	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.5		
episode: 395	score: 0.0 memory length: 72053	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 396	score: 2.0 memory length: 72251	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.49		
episode: 397	score: 1.0 memory length: 72421	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.46		
episode: 398	score: 4.0 memory length: 72697	epsilon: 1.0	steps: 276

lr: 0.0001	evaluation reward: 1.5		
episode: 399	score: 0.0 memory length: 72820	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.48		
episode: 400	score: 2.0 memory length: 73019	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.48		
episode: 401	score: 1.0 memory length: 73188	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.48		
episode: 402	score: 3.0 memory length: 73436	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.49		
episode: 403	score: 4.0 memory length: 73730	epsilon: 1.0	steps: 294
lr: 0.0001	evaluation reward: 1.52		
episode: 404	score: 2.0 memory length: 73948	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.51		
episode: 405	score: 1.0 memory length: 74118	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.49		
episode: 406	score: 2.0 memory length: 74318	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.51		
episode: 407	score: 2.0 memory length: 74517	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.53		
episode: 408	score: 3.0 memory length: 74742	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.56		
episode: 409	score: 0.0 memory length: 74864	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.55		
episode: 410	score: 0.0 memory length: 74986	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.54		
episode: 411	score: 0.0 memory length: 75109	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.5		
episode: 412	score: 0.0 memory length: 75232	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.49		
episode: 413	score: 1.0 memory length: 75401	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.46		
episode: 414	score: 1.0 memory length: 75570	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.45		
episode: 415	score: 1.0 memory length: 75721	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.44		
episode: 416	score: 1.0 memory length: 75892	epsilon: 1.0	steps: 171
lr: 0.0001	evaluation reward: 1.45		
episode: 417	score: 2.0 memory length: 76108	epsilon: 1.0	steps: 216
lr: 0.0001	evaluation reward: 1.47		
episode: 418	score: 0.0 memory length: 76231	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.47		
episode: 419	score: 1.0 memory length: 76401	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.46		
episode: 420	score: 0.0 memory length: 76524	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 421	score: 1.0 memory length: 76675	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.42		
episode: 422	score: 2.0 memory length: 76895	epsilon: 1.0	steps: 220

lr: 0.0001	evaluation reward: 1.43		
episode: 423	score: 0.0 memory length: 77018	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.42		
episode: 424	score: 1.0 memory length: 77187	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.41		
episode: 425	score: 1.0 memory length: 77337	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.4		
episode: 426	score: 2.0 memory length: 77520	epsilon: 1.0	steps: 183
lr: 0.0001	evaluation reward: 1.41		
episode: 427	score: 0.0 memory length: 77642	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.39		
episode: 428	score: 0.0 memory length: 77765	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.36		
episode: 429	score: 3.0 memory length: 78014	epsilon: 1.0	steps: 249
lr: 0.0001	evaluation reward: 1.37		
episode: 430	score: 1.0 memory length: 78186	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.35		
episode: 431	score: 0.0 memory length: 78309	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 432	score: 1.0 memory length: 78459	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.34		
episode: 433	score: 0.0 memory length: 78582	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 434	score: 2.0 memory length: 78780	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.33		
episode: 435	score: 3.0 memory length: 79007	epsilon: 1.0	steps: 227
lr: 0.0001	evaluation reward: 1.35		
episode: 436	score: 0.0 memory length: 79130	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.35		
episode: 437	score: 1.0 memory length: 79281	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.36		
episode: 438	score: 0.0 memory length: 79404	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 439	score: 3.0 memory length: 79652	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.33		
episode: 440	score: 0.0 memory length: 79774	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.33		
episode: 441	score: 1.0 memory length: 79924	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.31		
episode: 442	score: 0.0 memory length: 80047	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31		
episode: 443	score: 0.0 memory length: 80169	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3		
episode: 444	score: 1.0 memory length: 80320	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3		
episode: 445	score: 2.0 memory length: 80518	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.32		
episode: 446	score: 1.0 memory length: 80687	epsilon: 1.0	steps: 169

lr: 0.0001	evaluation reward: 1.33			
episode: 447	score: 0.0	memory length: 80810	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.29			
episode: 448	score: 7.0	memory length: 81179	epsilon: 1.0	steps: 369
lr: 0.0001	evaluation reward: 1.36			
episode: 449	score: 0.0	memory length: 81301	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.36			
episode: 450	score: 2.0	memory length: 81499	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.36			
episode: 451	score: 3.0	memory length: 81744	epsilon: 1.0	steps: 245
lr: 0.0001	evaluation reward: 1.38			
episode: 452	score: 4.0	memory length: 82021	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.37			
episode: 453	score: 1.0	memory length: 82191	epsilon: 1.0	steps: 170
lr: 0.0001	evaluation reward: 1.34			
episode: 454	score: 0.0	memory length: 82314	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33			
episode: 455	score: 2.0	memory length: 82512	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.31			
episode: 456	score: 5.0	memory length: 82846	epsilon: 1.0	steps: 334
lr: 0.0001	evaluation reward: 1.34			
episode: 457	score: 1.0	memory length: 83015	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.34			
episode: 458	score: 1.0	memory length: 83184	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.35			
episode: 459	score: 0.0	memory length: 83306	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.35			
episode: 460	score: 2.0	memory length: 83504	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.37			
episode: 461	score: 0.0	memory length: 83627	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.37			
episode: 462	score: 2.0	memory length: 83825	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.39			
episode: 463	score: 1.0	memory length: 83997	epsilon: 1.0	steps: 172
lr: 0.0001	evaluation reward: 1.39			
episode: 464	score: 2.0	memory length: 84216	epsilon: 1.0	steps: 219
lr: 0.0001	evaluation reward: 1.38			
episode: 465	score: 2.0	memory length: 84434	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.39			
episode: 466	score: 1.0	memory length: 84603	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.39			
episode: 467	score: 2.0	memory length: 84803	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.39			
episode: 468	score: 0.0	memory length: 84925	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.38			
episode: 469	score: 2.0	memory length: 85123	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.4			
episode: 470	score: 2.0	memory length: 85320	epsilon: 1.0	steps: 197

lr: 0.0001	evaluation reward: 1.38		
episode: 471	score: 2.0 memory length: 85518	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.38		
episode: 472	score: 0.0 memory length: 85640	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.35		
episode: 473	score: 2.0 memory length: 85857	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.37		
episode: 474	score: 1.0 memory length: 86025	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.37		
episode: 475	score: 1.0 memory length: 86193	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.35		
episode: 476	score: 0.0 memory length: 86315	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.34		
episode: 477	score: 0.0 memory length: 86438	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.32		
episode: 478	score: 1.0 memory length: 86590	epsilon: 1.0	steps: 152
lr: 0.0001	evaluation reward: 1.32		
episode: 479	score: 4.0 memory length: 86869	epsilon: 1.0	steps: 279
lr: 0.0001	evaluation reward: 1.31		
episode: 480	score: 2.0 memory length: 87072	epsilon: 1.0	steps: 203
lr: 0.0001	evaluation reward: 1.32		
episode: 481	score: 0.0 memory length: 87195	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.29		
episode: 482	score: 1.0 memory length: 87363	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.3		
episode: 483	score: 2.0 memory length: 87561	epsilon: 1.0	steps: 198
lr: 0.0001	evaluation reward: 1.31		
episode: 484	score: 1.0 memory length: 87712	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.3		
episode: 485	score: 0.0 memory length: 87834	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.29		
episode: 486	score: 3.0 memory length: 88099	epsilon: 1.0	steps: 265
lr: 0.0001	evaluation reward: 1.3		
episode: 487	score: 1.0 memory length: 88249	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.3		
episode: 488	score: 0.0 memory length: 88371	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.3		
episode: 489	score: 1.0 memory length: 88521	epsilon: 1.0	steps: 150
lr: 0.0001	evaluation reward: 1.31		
episode: 490	score: 0.0 memory length: 88644	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.31		
episode: 491	score: 2.0 memory length: 88845	epsilon: 1.0	steps: 201
lr: 0.0001	evaluation reward: 1.32		
episode: 492	score: 1.0 memory length: 88996	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.33		
episode: 493	score: 0.0 memory length: 89118	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.29		
episode: 494	score: 1.0 memory length: 89289	epsilon: 1.0	steps: 171

lr: 0.0001	evaluation reward: 1.28		
episode: 495	score: 3.0 memory length: 89537	epsilon: 1.0	steps: 248
lr: 0.0001	evaluation reward: 1.31		
episode: 496	score: 1.0 memory length: 89705	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.3		
episode: 497	score: 1.0 memory length: 89874	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.3		
episode: 498	score: 2.0 memory length: 90073	epsilon: 1.0	steps: 199
lr: 0.0001	evaluation reward: 1.28		
episode: 499	score: 0.0 memory length: 90196	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.28		
episode: 500	score: 0.0 memory length: 90319	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.26		
episode: 501	score: 3.0 memory length: 90587	epsilon: 1.0	steps: 268
lr: 0.0001	evaluation reward: 1.28		
episode: 502	score: 3.0 memory length: 90833	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.28		
episode: 503	score: 0.0 memory length: 90955	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.24		
episode: 504	score: 1.0 memory length: 91106	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.23		
episode: 505	score: 3.0 memory length: 91352	epsilon: 1.0	steps: 246
lr: 0.0001	evaluation reward: 1.25		
episode: 506	score: 2.0 memory length: 91552	epsilon: 1.0	steps: 200
lr: 0.0001	evaluation reward: 1.25		
episode: 507	score: 4.0 memory length: 91829	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.27		
episode: 508	score: 0.0 memory length: 91952	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.24		
episode: 509	score: 3.0 memory length: 92202	epsilon: 1.0	steps: 250
lr: 0.0001	evaluation reward: 1.27		
episode: 510	score: 1.0 memory length: 92370	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.28		
episode: 511	score: 4.0 memory length: 92664	epsilon: 1.0	steps: 294
lr: 0.0001	evaluation reward: 1.32		
episode: 512	score: 2.0 memory length: 92861	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.34		
episode: 513	score: 1.0 memory length: 93029	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.34		
episode: 514	score: 0.0 memory length: 93152	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 515	score: 0.0 memory length: 93274	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.32		
episode: 516	score: 3.0 memory length: 93499	epsilon: 1.0	steps: 225
lr: 0.0001	evaluation reward: 1.34		
episode: 517	score: 2.0 memory length: 93681	epsilon: 1.0	steps: 182
lr: 0.0001	evaluation reward: 1.34		
episode: 518	score: 0.0 memory length: 93804	epsilon: 1.0	steps: 123

lr: 0.0001	evaluation reward: 1.34		
episode: 519	score: 1.0 memory length: 93972	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.34		
episode: 520	score: 0.0 memory length: 94095	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.34		
episode: 521	score: 0.0 memory length: 94218	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 522	score: 0.0 memory length: 94340	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.31		
episode: 523	score: 2.0 memory length: 94537	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.33		
episode: 524	score: 1.0 memory length: 94706	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.33		
episode: 525	score: 2.0 memory length: 94923	epsilon: 1.0	steps: 217
lr: 0.0001	evaluation reward: 1.34		
episode: 526	score: 1.0 memory length: 95074	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.33		
episode: 527	score: 0.0 memory length: 95196	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.33		
episode: 528	score: 0.0 memory length: 95319	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.33		
episode: 529	score: 1.0 memory length: 95470	epsilon: 1.0	steps: 151
lr: 0.0001	evaluation reward: 1.31		
episode: 530	score: 3.0 memory length: 95717	epsilon: 1.0	steps: 247
lr: 0.0001	evaluation reward: 1.33		
episode: 531	score: 1.0 memory length: 95886	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.34		
episode: 532	score: 3.0 memory length: 96115	epsilon: 1.0	steps: 229
lr: 0.0001	evaluation reward: 1.36		
episode: 533	score: 2.0 memory length: 96312	epsilon: 1.0	steps: 197
lr: 0.0001	evaluation reward: 1.38		
episode: 534	score: 3.0 memory length: 96542	epsilon: 1.0	steps: 230
lr: 0.0001	evaluation reward: 1.39		
episode: 535	score: 1.0 memory length: 96710	epsilon: 1.0	steps: 168
lr: 0.0001	evaluation reward: 1.37		
episode: 536	score: 4.0 memory length: 96987	epsilon: 1.0	steps: 277
lr: 0.0001	evaluation reward: 1.41		
episode: 537	score: 1.0 memory length: 97156	epsilon: 1.0	steps: 169
lr: 0.0001	evaluation reward: 1.41		
episode: 538	score: 0.0 memory length: 97279	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.41		
episode: 539	score: 2.0 memory length: 97497	epsilon: 1.0	steps: 218
lr: 0.0001	evaluation reward: 1.4		
episode: 540	score: 0.0 memory length: 97619	epsilon: 1.0	steps: 122
lr: 0.0001	evaluation reward: 1.4		
episode: 541	score: 0.0 memory length: 97742	epsilon: 1.0	steps: 123
lr: 0.0001	evaluation reward: 1.39		
episode: 542	score: 1.0 memory length: 97892	epsilon: 1.0	steps: 150

lr: 0.0001	evaluation reward: 1.4		
episode: 543	score: 2.0	memory length: 98089	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.42		
episode: 544	score: 1.0	memory length: 98240	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.42		
episode: 545	score: 0.0	memory length: 98363	epsilon: 1.0 steps: 123
lr: 0.0001	evaluation reward: 1.4		
episode: 546	score: 2.0	memory length: 98560	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.41		
episode: 547	score: 3.0	memory length: 98806	epsilon: 1.0 steps: 246
lr: 0.0001	evaluation reward: 1.44		
episode: 548	score: 3.0	memory length: 99031	epsilon: 1.0 steps: 225
lr: 0.0001	evaluation reward: 1.4		
episode: 549	score: 2.0	memory length: 99252	epsilon: 1.0 steps: 221
lr: 0.0001	evaluation reward: 1.42		
episode: 550	score: 2.0	memory length: 99449	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.42		
episode: 551	score: 2.0	memory length: 99646	epsilon: 1.0 steps: 197
lr: 0.0001	evaluation reward: 1.41		
episode: 552	score: 1.0	memory length: 99797	epsilon: 1.0 steps: 151
lr: 0.0001	evaluation reward: 1.38		
episode: 553	score: 3.0	memory length: 100043	epsilon: 0.9999564400000001
steps: 246	lr: 0.0001	evaluation reward: 1.4	
episode: 554	score: 2.0	memory length: 100261	epsilon: 0.9997406200000056
steps: 218	lr: 0.0001	evaluation reward: 1.42	
episode: 555	score: 0.0	memory length: 100384	epsilon: 0.9996188500000083
steps: 123	lr: 0.0001	evaluation reward: 1.4	
episode: 556	score: 4.0	memory length: 100661	epsilon: 0.9993446200000142
steps: 277	lr: 0.0001	evaluation reward: 1.39	
episode: 557	score: 0.0	memory length: 100784	epsilon: 0.9992228500000169
steps: 123	lr: 0.0001	evaluation reward: 1.38	
episode: 558	score: 3.0	memory length: 101031	epsilon: 0.9989783200000222
steps: 247	lr: 0.0001	evaluation reward: 1.4	
episode: 559	score: 1.0	memory length: 101182	epsilon: 0.9988288300000254
steps: 151	lr: 0.0001	evaluation reward: 1.41	
episode: 560	score: 0.0	memory length: 101304	epsilon: 0.9987080500000028
steps: 122	lr: 0.0001	evaluation reward: 1.39	
episode: 561	score: 3.0	memory length: 101530	epsilon: 0.9984843100000329
steps: 226	lr: 0.0001	evaluation reward: 1.42	
episode: 562	score: 1.0	memory length: 101683	epsilon: 0.9983328400000362
steps: 153	lr: 0.0001	evaluation reward: 1.41	
episode: 563	score: 1.0	memory length: 101833	epsilon: 0.9981843400000394
steps: 150	lr: 0.0001	evaluation reward: 1.41	
episode: 564	score: 3.0	memory length: 102081	epsilon: 0.9979388200000447
steps: 248	lr: 0.0001	evaluation reward: 1.42	
episode: 565	score: 5.0	memory length: 102394	epsilon: 0.9976289500000515
steps: 313	lr: 0.0001	evaluation reward: 1.45	
episode: 566	score: 1.0	memory length: 102564	epsilon: 0.9974606500000551

steps: 170	lr: 0.0001	evaluation reward: 1.45
episode: 567	score: 1.0	memory length: 102716 epsilon: 0.9973101700000584
steps: 152	lr: 0.0001	evaluation reward: 1.44
episode: 568	score: 2.0	memory length: 102932 epsilon: 0.9970963300000063
steps: 216	lr: 0.0001	evaluation reward: 1.46
episode: 569	score: 3.0	memory length: 103176 epsilon: 0.9968547700000683
steps: 244	lr: 0.0001	evaluation reward: 1.47
episode: 570	score: 0.0	memory length: 103299 epsilon: 0.9967330000000709
steps: 123	lr: 0.0001	evaluation reward: 1.45
episode: 571	score: 2.0	memory length: 103497 epsilon: 0.9965369800000752
steps: 198	lr: 0.0001	evaluation reward: 1.45
episode: 572	score: 2.0	memory length: 103714 epsilon: 0.9963221500000798
steps: 217	lr: 0.0001	evaluation reward: 1.47
episode: 573	score: 0.0	memory length: 103837 epsilon: 0.9962003800000825
steps: 123	lr: 0.0001	evaluation reward: 1.45
episode: 574	score: 2.0	memory length: 104056 epsilon: 0.9959835700000872
steps: 219	lr: 0.0001	evaluation reward: 1.46
episode: 575	score: 1.0	memory length: 104224 epsilon: 0.9958172500000908
steps: 168	lr: 0.0001	evaluation reward: 1.46
episode: 576	score: 2.0	memory length: 104441 epsilon: 0.9956024200000955
steps: 217	lr: 0.0001	evaluation reward: 1.48
episode: 577	score: 2.0	memory length: 104622 epsilon: 0.9954232300000994
steps: 181	lr: 0.0001	evaluation reward: 1.5
episode: 578	score: 3.0	memory length: 104867 epsilon: 0.9951806800001046
steps: 245	lr: 0.0001	evaluation reward: 1.52
episode: 579	score: 1.0	memory length: 105039 epsilon: 0.9950104000001083
steps: 172	lr: 0.0001	evaluation reward: 1.49
episode: 580	score: 0.0	memory length: 105161 epsilon: 0.9948896200001109
steps: 122	lr: 0.0001	evaluation reward: 1.47
episode: 581	score: 0.0	memory length: 105283 epsilon: 0.9947688400001136
steps: 122	lr: 0.0001	evaluation reward: 1.47
episode: 582	score: 2.0	memory length: 105502 epsilon: 0.9945520300001183
steps: 219	lr: 0.0001	evaluation reward: 1.48
episode: 583	score: 1.0	memory length: 105674 epsilon: 0.994381750000122
steps: 172	lr: 0.0001	evaluation reward: 1.47
episode: 584	score: 0.0	memory length: 105796 epsilon: 0.9942609700001246
steps: 122	lr: 0.0001	evaluation reward: 1.46
episode: 585	score: 1.0	memory length: 105947 epsilon: 0.9941114800001278
steps: 151	lr: 0.0001	evaluation reward: 1.47
episode: 586	score: 1.0	memory length: 106115 epsilon: 0.9939451600001314
steps: 168	lr: 0.0001	evaluation reward: 1.45
episode: 587	score: 1.0	memory length: 106265 epsilon: 0.9937966600001347
steps: 150	lr: 0.0001	evaluation reward: 1.45
episode: 588	score: 3.0	memory length: 106511 epsilon: 0.99355312000014
steps: 246	lr: 0.0001	evaluation reward: 1.48
episode: 589	score: 3.0	memory length: 106737 epsilon: 0.9933293800001448
steps: 226	lr: 0.0001	evaluation reward: 1.5
episode: 590	score: 4.0	memory length: 107051 epsilon: 0.9930185200001516

steps: 314	lr: 0.0001	evaluation reward: 1.54
episode: 591	score: 2.0	memory length: 107249 epsilon: 0.9928225000001558
steps: 198	lr: 0.0001	evaluation reward: 1.54
episode: 592	score: 0.0	memory length: 107372 epsilon: 0.9927007300001585
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 593	score: 2.0	memory length: 107590 epsilon: 0.9924849100001631
steps: 218	lr: 0.0001	evaluation reward: 1.55
episode: 594	score: 2.0	memory length: 107788 epsilon: 0.9922888900001674
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 595	score: 1.0	memory length: 107958 epsilon: 0.992120590000171
steps: 170	lr: 0.0001	evaluation reward: 1.54
episode: 596	score: 0.0	memory length: 108080 epsilon: 0.9919998100001737
steps: 122	lr: 0.0001	evaluation reward: 1.53
episode: 597	score: 0.0	memory length: 108203 epsilon: 0.9918780400001763
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 598	score: 2.0	memory length: 108401 epsilon: 0.9916820200001806
steps: 198	lr: 0.0001	evaluation reward: 1.52
episode: 599	score: 2.0	memory length: 108599 epsilon: 0.9914860000001848
steps: 198	lr: 0.0001	evaluation reward: 1.54
episode: 600	score: 1.0	memory length: 108769 epsilon: 0.9913177000001885
steps: 170	lr: 0.0001	evaluation reward: 1.55
episode: 601	score: 2.0	memory length: 108971 epsilon: 0.9911177200001928
steps: 202	lr: 0.0001	evaluation reward: 1.54
episode: 602	score: 0.0	memory length: 109094 epsilon: 0.9909959500001955
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 603	score: 0.0	memory length: 109216 epsilon: 0.9908751700001981
steps: 122	lr: 0.0001	evaluation reward: 1.51
episode: 604	score: 1.0	memory length: 109385 epsilon: 0.9907078600002017
steps: 169	lr: 0.0001	evaluation reward: 1.51
episode: 605	score: 1.0	memory length: 109536 epsilon: 0.990558370000205
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 606	score: 0.0	memory length: 109658 epsilon: 0.9904375900002076
steps: 122	lr: 0.0001	evaluation reward: 1.47
episode: 607	score: 2.0	memory length: 109856 epsilon: 0.9902415700002118
steps: 198	lr: 0.0001	evaluation reward: 1.45
episode: 608	score: 2.0	memory length: 110053 epsilon: 0.9900465400002161
steps: 197	lr: 0.0001	evaluation reward: 1.47
episode: 609	score: 0.0	memory length: 110176 epsilon: 0.9899247700002187
steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 610	score: 3.0	memory length: 110424 epsilon: 0.989679250000224
steps: 248	lr: 0.0001	evaluation reward: 1.46
episode: 611	score: 2.0	memory length: 110642 epsilon: 0.9894634300002287
steps: 218	lr: 0.0001	evaluation reward: 1.44
episode: 612	score: 2.0	memory length: 110857 epsilon: 0.9892505800002334
steps: 215	lr: 0.0001	evaluation reward: 1.44
episode: 613	score: 1.0	memory length: 111008 epsilon: 0.9891010900002366
steps: 151	lr: 0.0001	evaluation reward: 1.44
episode: 614	score: 0.0	memory length: 111131 epsilon: 0.9889793200002392

steps: 123	lr: 0.0001	evaluation reward: 1.44
episode: 615	score: 1.0	memory length: 111302 epsilon: 0.9888100300002429
steps: 171	lr: 0.0001	evaluation reward: 1.45
episode: 616	score: 2.0	memory length: 111501 epsilon: 0.9886130200002472
steps: 199	lr: 0.0001	evaluation reward: 1.44
episode: 617	score: 2.0	memory length: 111699 epsilon: 0.9884170000002515
steps: 198	lr: 0.0001	evaluation reward: 1.44
episode: 618	score: 1.0	memory length: 111850 epsilon: 0.9882675100002547
steps: 151	lr: 0.0001	evaluation reward: 1.45
episode: 619	score: 2.0	memory length: 112047 epsilon: 0.9880724800002589
steps: 197	lr: 0.0001	evaluation reward: 1.46
episode: 620	score: 3.0	memory length: 112294 epsilon: 0.9878279500002642
steps: 247	lr: 0.0001	evaluation reward: 1.49
episode: 621	score: 0.0	memory length: 112417 epsilon: 0.9877061800002669
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 622	score: 1.0	memory length: 112589 epsilon: 0.9875359000002706
steps: 172	lr: 0.0001	evaluation reward: 1.5
episode: 623	score: 1.0	memory length: 112740 epsilon: 0.9873864100002738
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 624	score: 3.0	memory length: 112985 epsilon: 0.9871438600002791
steps: 245	lr: 0.0001	evaluation reward: 1.51
episode: 625	score: 0.0	memory length: 113108 epsilon: 0.9870220900002817
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 626	score: 1.0	memory length: 113278 epsilon: 0.9868537900002854
steps: 170	lr: 0.0001	evaluation reward: 1.49
episode: 627	score: 5.0	memory length: 113558 epsilon: 0.9865765900002914
steps: 280	lr: 0.0001	evaluation reward: 1.54
episode: 628	score: 1.0	memory length: 113727 epsilon: 0.986409280000295
steps: 169	lr: 0.0001	evaluation reward: 1.55
episode: 629	score: 2.0	memory length: 113945 epsilon: 0.9861934600002997
steps: 218	lr: 0.0001	evaluation reward: 1.56
episode: 630	score: 2.0	memory length: 114144 epsilon: 0.985996450000304
steps: 199	lr: 0.0001	evaluation reward: 1.55
episode: 631	score: 0.0	memory length: 114267 epsilon: 0.9858746800003066
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 632	score: 0.0	memory length: 114390 epsilon: 0.9857529100003093
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 633	score: 0.0	memory length: 114513 epsilon: 0.9856311400003119
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 634	score: 4.0	memory length: 114808 epsilon: 0.9853390900003183
steps: 295	lr: 0.0001	evaluation reward: 1.5
episode: 635	score: 3.0	memory length: 115055 epsilon: 0.9850945600003236
steps: 247	lr: 0.0001	evaluation reward: 1.52
episode: 636	score: 1.0	memory length: 115206 epsilon: 0.9849450700003268
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 637	score: 2.0	memory length: 115424 epsilon: 0.9847292500003315
steps: 218	lr: 0.0001	evaluation reward: 1.5
episode: 638	score: 1.0	memory length: 115596 epsilon: 0.9845589700003352

steps: 172	lr: 0.0001	evaluation reward: 1.51
episode: 639	score: 4.0	memory length: 115892 epsilon: 0.9842659300003416
steps: 296	lr: 0.0001	evaluation reward: 1.53
episode: 640	score: 4.0	memory length: 116167 epsilon: 0.9839936800003475
steps: 275	lr: 0.0001	evaluation reward: 1.57
episode: 641	score: 5.0	memory length: 116492 epsilon: 0.9836719300003545
steps: 325	lr: 0.0001	evaluation reward: 1.62
episode: 642	score: 4.0	memory length: 116805 epsilon: 0.9833620600003612
steps: 313	lr: 0.0001	evaluation reward: 1.65
episode: 643	score: 0.0	memory length: 116928 epsilon: 0.9832402900003638
steps: 123	lr: 0.0001	evaluation reward: 1.63
episode: 644	score: 1.0	memory length: 117079 epsilon: 0.9830908000003671
steps: 151	lr: 0.0001	evaluation reward: 1.63
episode: 645	score: 3.0	memory length: 117325 epsilon: 0.9828472600003724
steps: 246	lr: 0.0001	evaluation reward: 1.66
episode: 646	score: 2.0	memory length: 117523 epsilon: 0.9826512400003766
steps: 198	lr: 0.0001	evaluation reward: 1.66
episode: 647	score: 0.0	memory length: 117645 epsilon: 0.9825304600003792
steps: 122	lr: 0.0001	evaluation reward: 1.63
episode: 648	score: 1.0	memory length: 117817 epsilon: 0.9823601800003829
steps: 172	lr: 0.0001	evaluation reward: 1.61
episode: 649	score: 1.0	memory length: 117967 epsilon: 0.9822116800003862
steps: 150	lr: 0.0001	evaluation reward: 1.6
episode: 650	score: 1.0	memory length: 118119 epsilon: 0.9820612000003894
steps: 152	lr: 0.0001	evaluation reward: 1.59
episode: 651	score: 3.0	memory length: 118384 epsilon: 0.9817988500003951
steps: 265	lr: 0.0001	evaluation reward: 1.6
episode: 652	score: 1.0	memory length: 118553 epsilon: 0.9816315400003988
steps: 169	lr: 0.0001	evaluation reward: 1.6
episode: 653	score: 3.0	memory length: 118798 epsilon: 0.981388990000404
steps: 245	lr: 0.0001	evaluation reward: 1.6
episode: 654	score: 1.0	memory length: 118968 epsilon: 0.9812206900004077
steps: 170	lr: 0.0001	evaluation reward: 1.59
episode: 655	score: 2.0	memory length: 119166 epsilon: 0.9810246700004119
steps: 198	lr: 0.0001	evaluation reward: 1.61
episode: 656	score: 0.0	memory length: 119289 epsilon: 0.9809029000004146
steps: 123	lr: 0.0001	evaluation reward: 1.57
episode: 657	score: 1.0	memory length: 119457 epsilon: 0.9807365800004182
steps: 168	lr: 0.0001	evaluation reward: 1.58
episode: 658	score: 3.0	memory length: 119667 epsilon: 0.9805286800004227
steps: 210	lr: 0.0001	evaluation reward: 1.58
episode: 659	score: 1.0	memory length: 119836 epsilon: 0.9803613700004263
steps: 169	lr: 0.0001	evaluation reward: 1.58
episode: 660	score: 0.0	memory length: 119959 epsilon: 0.980239600000429
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 661	score: 2.0	memory length: 120177 epsilon: 0.9800237800004337
steps: 218	lr: 0.0001	evaluation reward: 1.57
episode: 662	score: 2.0	memory length: 120375 epsilon: 0.9798277600004379

steps: 198	lr: 0.0001	evaluation reward: 1.58
episode: 663	score: 0.0	memory length: 120498 epsilon: 0.9797059900004406
steps: 123	lr: 0.0001	evaluation reward: 1.57
episode: 664	score: 0.0	memory length: 120620 epsilon: 0.9795852100004432
steps: 122	lr: 0.0001	evaluation reward: 1.54
episode: 665	score: 0.0	memory length: 120743 epsilon: 0.9794634400004458
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 666	score: 1.0	memory length: 120894 epsilon: 0.9793139500004491
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 667	score: 2.0	memory length: 121091 epsilon: 0.9791189200004533
steps: 197	lr: 0.0001	evaluation reward: 1.5
episode: 668	score: 2.0	memory length: 121290 epsilon: 0.9789219100004576
steps: 199	lr: 0.0001	evaluation reward: 1.5
episode: 669	score: 4.0	memory length: 121566 epsilon: 0.9786486700004635
steps: 276	lr: 0.0001	evaluation reward: 1.51
episode: 670	score: 1.0	memory length: 121735 epsilon: 0.9784813600004671
steps: 169	lr: 0.0001	evaluation reward: 1.52
episode: 671	score: 1.0	memory length: 121904 epsilon: 0.9783140500004708
steps: 169	lr: 0.0001	evaluation reward: 1.51
episode: 672	score: 1.0	memory length: 122073 epsilon: 0.9781467400004744
steps: 169	lr: 0.0001	evaluation reward: 1.5
episode: 673	score: 2.0	memory length: 122270 epsilon: 0.9779517100004786
steps: 197	lr: 0.0001	evaluation reward: 1.52
episode: 674	score: 4.0	memory length: 122529 epsilon: 0.9776953000004842
steps: 259	lr: 0.0001	evaluation reward: 1.54
episode: 675	score: 0.0	memory length: 122652 epsilon: 0.9775735300004869
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 676	score: 0.0	memory length: 122775 epsilon: 0.9774517600004895
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 677	score: 1.0	memory length: 122943 epsilon: 0.9772854400004931
steps: 168	lr: 0.0001	evaluation reward: 1.5
episode: 678	score: 1.0	memory length: 123094 epsilon: 0.9771359500004964
steps: 151	lr: 0.0001	evaluation reward: 1.48
episode: 679	score: 1.0	memory length: 123263 epsilon: 0.9769686400005
steps: 169	lr: 0.0001	evaluation reward: 1.48
episode: 680	score: 0.0	memory length: 123386 epsilon: 0.9768468700005026
steps: 123	lr: 0.0001	evaluation reward: 1.48
episode: 681	score: 1.0	memory length: 123557 epsilon: 0.9766775800005063
steps: 171	lr: 0.0001	evaluation reward: 1.49
episode: 682	score: 2.0	memory length: 123756 epsilon: 0.9764805700005106
steps: 199	lr: 0.0001	evaluation reward: 1.49
episode: 683	score: 1.0	memory length: 123907 epsilon: 0.9763310800005138
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 684	score: 1.0	memory length: 124079 epsilon: 0.9761608000005175
steps: 172	lr: 0.0001	evaluation reward: 1.5
episode: 685	score: 1.0	memory length: 124230 epsilon: 0.9760113100005208
steps: 151	lr: 0.0001	evaluation reward: 1.5
episode: 686	score: 2.0	memory length: 124450 epsilon: 0.9757935100005255

steps: 220	lr: 0.0001	evaluation reward: 1.51
episode: 687	score: 4.0	memory length: 124745 epsilon: 0.9755014600005318
steps: 295	lr: 0.0001	evaluation reward: 1.54
episode: 688	score: 2.0	memory length: 124961 epsilon: 0.9752876200005365
steps: 216	lr: 0.0001	evaluation reward: 1.53
episode: 689	score: 3.0	memory length: 125190 epsilon: 0.9750609100005414
steps: 229	lr: 0.0001	evaluation reward: 1.53
episode: 690	score: 1.0	memory length: 125358 epsilon: 0.974894590000545
steps: 168	lr: 0.0001	evaluation reward: 1.5
episode: 691	score: 1.0	memory length: 125508 epsilon: 0.9747460900005482
steps: 150	lr: 0.0001	evaluation reward: 1.49
episode: 692	score: 1.0	memory length: 125678 epsilon: 0.9745777900005519
steps: 170	lr: 0.0001	evaluation reward: 1.5
episode: 693	score: 1.0	memory length: 125847 epsilon: 0.9744104800005555
steps: 169	lr: 0.0001	evaluation reward: 1.49
episode: 694	score: 2.0	memory length: 126067 epsilon: 0.9741926800005603
steps: 220	lr: 0.0001	evaluation reward: 1.49
episode: 695	score: 1.0	memory length: 126236 epsilon: 0.9740253700005639
steps: 169	lr: 0.0001	evaluation reward: 1.49
episode: 696	score: 0.0	memory length: 126358 epsilon: 0.9739045900005665
steps: 122	lr: 0.0001	evaluation reward: 1.49
episode: 697	score: 4.0	memory length: 126674 epsilon: 0.9735917500005733
steps: 316	lr: 0.0001	evaluation reward: 1.53
episode: 698	score: 0.0	memory length: 126797 epsilon: 0.9734699800005759
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 699	score: 1.0	memory length: 126968 epsilon: 0.9733006900005796
steps: 171	lr: 0.0001	evaluation reward: 1.5
episode: 700	score: 0.0	memory length: 127091 epsilon: 0.9731789200005823
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 701	score: 1.0	memory length: 127244 epsilon: 0.9730274500005855
steps: 153	lr: 0.0001	evaluation reward: 1.48
episode: 702	score: 1.0	memory length: 127395 epsilon: 0.9728779600005888
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 703	score: 2.0	memory length: 127614 epsilon: 0.9726611500005935
steps: 219	lr: 0.0001	evaluation reward: 1.51
episode: 704	score: 3.0	memory length: 127840 epsilon: 0.9724374100005984
steps: 226	lr: 0.0001	evaluation reward: 1.53
episode: 705	score: 0.0	memory length: 127963 epsilon: 0.972315640000601
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 706	score: 1.0	memory length: 128132 epsilon: 0.9721483300006046
steps: 169	lr: 0.0001	evaluation reward: 1.53
episode: 707	score: 4.0	memory length: 128390 epsilon: 0.9718929100006102
steps: 258	lr: 0.0001	evaluation reward: 1.55
episode: 708	score: 2.0	memory length: 128587 epsilon: 0.9716978800006144
steps: 197	lr: 0.0001	evaluation reward: 1.55
episode: 709	score: 0.0	memory length: 128710 epsilon: 0.971576110000617
steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 710	score: 1.0	memory length: 128879 epsilon: 0.9714088000006207

steps: 169	lr: 0.0001	evaluation reward: 1.53
episode: 711	score: 1.0	memory length: 129048 epsilon: 0.9712414900006243
steps: 169	lr: 0.0001	evaluation reward: 1.52
episode: 712	score: 3.0	memory length: 129292 epsilon: 0.9709999300006296
steps: 244	lr: 0.0001	evaluation reward: 1.53
episode: 713	score: 0.0	memory length: 129414 epsilon: 0.9708791500006322
steps: 122	lr: 0.0001	evaluation reward: 1.52
episode: 714	score: 2.0	memory length: 129632 epsilon: 0.9706633300006369
steps: 218	lr: 0.0001	evaluation reward: 1.54
episode: 715	score: 3.0	memory length: 129860 epsilon: 0.9704376100006418
steps: 228	lr: 0.0001	evaluation reward: 1.56
episode: 716	score: 1.0	memory length: 130029 epsilon: 0.9702703000006454
steps: 169	lr: 0.0001	evaluation reward: 1.55
episode: 717	score: 3.0	memory length: 130275 epsilon: 0.9700267600006507
steps: 246	lr: 0.0001	evaluation reward: 1.56
episode: 718	score: 2.0	memory length: 130473 epsilon: 0.969830740000655
steps: 198	lr: 0.0001	evaluation reward: 1.57
episode: 719	score: 1.0	memory length: 130643 epsilon: 0.9696624400006586
steps: 170	lr: 0.0001	evaluation reward: 1.56
episode: 720	score: 1.0	memory length: 130811 epsilon: 0.9694961200006622
steps: 168	lr: 0.0001	evaluation reward: 1.54
episode: 721	score: 2.0	memory length: 131009 epsilon: 0.9693001000006665
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 722	score: 1.0	memory length: 131178 epsilon: 0.9691327900006701
steps: 169	lr: 0.0001	evaluation reward: 1.56
episode: 723	score: 1.0	memory length: 131349 epsilon: 0.9689635000006738
steps: 171	lr: 0.0001	evaluation reward: 1.56
episode: 724	score: 0.0	memory length: 131471 epsilon: 0.9688427200006764
steps: 122	lr: 0.0001	evaluation reward: 1.53
episode: 725	score: 0.0	memory length: 131594 epsilon: 0.968720950000679
steps: 123	lr: 0.0001	evaluation reward: 1.53
episode: 726	score: 5.0	memory length: 131898 epsilon: 0.9684199900006856
steps: 304	lr: 0.0001	evaluation reward: 1.57
episode: 727	score: 1.0	memory length: 132049 epsilon: 0.9682705000006888
steps: 151	lr: 0.0001	evaluation reward: 1.53
episode: 728	score: 0.0	memory length: 132172 epsilon: 0.9681487300006915
steps: 123	lr: 0.0001	evaluation reward: 1.52
episode: 729	score: 1.0	memory length: 132341 epsilon: 0.9679814200006951
steps: 169	lr: 0.0001	evaluation reward: 1.51
episode: 730	score: 2.0	memory length: 132539 epsilon: 0.9677854000006993
steps: 198	lr: 0.0001	evaluation reward: 1.51
episode: 731	score: 2.0	memory length: 132737 epsilon: 0.9675893800007036
steps: 198	lr: 0.0001	evaluation reward: 1.53
episode: 732	score: 2.0	memory length: 132938 epsilon: 0.9673903900007079
steps: 201	lr: 0.0001	evaluation reward: 1.55
episode: 733	score: 1.0	memory length: 133107 epsilon: 0.9672230800007116
steps: 169	lr: 0.0001	evaluation reward: 1.56
episode: 734	score: 1.0	memory length: 133276 epsilon: 0.9670557700007152

steps: 169	lr: 0.0001	evaluation reward: 1.53
episode: 735	score: 0.0	memory length: 133398 epsilon: 0.9669349900007178
steps: 122	lr: 0.0001	evaluation reward: 1.5
episode: 736	score: 0.0	memory length: 133521 epsilon: 0.9668132200007205
steps: 123	lr: 0.0001	evaluation reward: 1.49
episode: 737	score: 1.0	memory length: 133693 epsilon: 0.9666429400007241
steps: 172	lr: 0.0001	evaluation reward: 1.48
episode: 738	score: 2.0	memory length: 133890 epsilon: 0.9664479100007284
steps: 197	lr: 0.0001	evaluation reward: 1.49
episode: 739	score: 0.0	memory length: 134013 epsilon: 0.966326140000731
steps: 123	lr: 0.0001	evaluation reward: 1.45
episode: 740	score: 3.0	memory length: 134263 epsilon: 0.9660786400007364
steps: 250	lr: 0.0001	evaluation reward: 1.44
episode: 741	score: 0.0	memory length: 134386 epsilon: 0.965956870000739
steps: 123	lr: 0.0001	evaluation reward: 1.39
episode: 742	score: 0.0	memory length: 134508 epsilon: 0.9658360900007417
steps: 122	lr: 0.0001	evaluation reward: 1.35
episode: 743	score: 1.0	memory length: 134679 epsilon: 0.9656668000007453
steps: 171	lr: 0.0001	evaluation reward: 1.36
episode: 744	score: 2.0	memory length: 134877 epsilon: 0.9654707800007496
steps: 198	lr: 0.0001	evaluation reward: 1.37
episode: 745	score: 3.0	memory length: 135122 epsilon: 0.9652282300007549
steps: 245	lr: 0.0001	evaluation reward: 1.37
episode: 746	score: 0.0	memory length: 135244 epsilon: 0.9651074500007575
steps: 122	lr: 0.0001	evaluation reward: 1.35
episode: 747	score: 0.0	memory length: 135367 epsilon: 0.9649856800007601
steps: 123	lr: 0.0001	evaluation reward: 1.35
episode: 748	score: 2.0	memory length: 135586 epsilon: 0.9647688700007648
steps: 219	lr: 0.0001	evaluation reward: 1.36
episode: 749	score: 1.0	memory length: 135757 epsilon: 0.9645995800007685
steps: 171	lr: 0.0001	evaluation reward: 1.36
episode: 750	score: 3.0	memory length: 135985 epsilon: 0.9643738600007734
steps: 228	lr: 0.0001	evaluation reward: 1.38
episode: 751	score: 2.0	memory length: 136167 epsilon: 0.9641936800007773
steps: 182	lr: 0.0001	evaluation reward: 1.37
episode: 752	score: 1.0	memory length: 136318 epsilon: 0.9640441900007806
steps: 151	lr: 0.0001	evaluation reward: 1.37
episode: 753	score: 0.0	memory length: 136441 epsilon: 0.9639224200007832
steps: 123	lr: 0.0001	evaluation reward: 1.34
episode: 754	score: 3.0	memory length: 136688 epsilon: 0.9636778900007885
steps: 247	lr: 0.0001	evaluation reward: 1.36
episode: 755	score: 1.0	memory length: 136856 epsilon: 0.9635115700007921
steps: 168	lr: 0.0001	evaluation reward: 1.35
episode: 756	score: 0.0	memory length: 136979 epsilon: 0.9633898000007948
steps: 123	lr: 0.0001	evaluation reward: 1.35
episode: 757	score: 2.0	memory length: 137176 epsilon: 0.963194770000799
steps: 197	lr: 0.0001	evaluation reward: 1.36
episode: 758	score: 1.0	memory length: 137345 epsilon: 0.9630274600008026

steps: 169	lr: 0.0001	evaluation reward: 1.34	
episode: 759	score: 2.0	memory length: 137543	epsilon: 0.9628314400008069
steps: 198	lr: 0.0001	evaluation reward: 1.35	
episode: 760	score: 5.0	memory length: 137867	epsilon: 0.9625106800008139
steps: 324	lr: 0.0001	evaluation reward: 1.4	
episode: 761	score: 2.0	memory length: 138086	epsilon: 0.9622938700008186
steps: 219	lr: 0.0001	evaluation reward: 1.4	
episode: 762	score: 2.0	memory length: 138304	epsilon: 0.9620780500008232
steps: 218	lr: 0.0001	evaluation reward: 1.4	
episode: 763	score: 1.0	memory length: 138457	epsilon: 0.9619265800008265
steps: 153	lr: 0.0001	evaluation reward: 1.41	
episode: 764	score: 5.0	memory length: 138821	epsilon: 0.9615662200008344
steps: 364	lr: 0.0001	evaluation reward: 1.46	
episode: 765	score: 3.0	memory length: 139049	epsilon: 0.9613405000008393
steps: 228	lr: 0.0001	evaluation reward: 1.49	
episode: 766	score: 2.0	memory length: 139251	epsilon: 0.9611405200008436
steps: 202	lr: 0.0001	evaluation reward: 1.5	
episode: 767	score: 1.0	memory length: 139402	epsilon: 0.9609910300008468
steps: 151	lr: 0.0001	evaluation reward: 1.49	
episode: 768	score: 1.0	memory length: 139570	epsilon: 0.9608247100008505
steps: 168	lr: 0.0001	evaluation reward: 1.48	
episode: 769	score: 0.0	memory length: 139693	epsilon: 0.9607029400008531
steps: 123	lr: 0.0001	evaluation reward: 1.44	
episode: 770	score: 0.0	memory length: 139816	epsilon: 0.9605811700008557
steps: 123	lr: 0.0001	evaluation reward: 1.43	
episode: 771	score: 2.0	memory length: 140034	epsilon: 0.9603653500008604
steps: 218	lr: 0.0001	evaluation reward: 1.44	
episode: 772	score: 1.0	memory length: 140203	epsilon: 0.9601980400008641
steps: 169	lr: 0.0001	evaluation reward: 1.44	
episode: 773	score: 1.0	memory length: 140374	epsilon: 0.9600287500008677
steps: 171	lr: 0.0001	evaluation reward: 1.43	
episode: 774	score: 4.0	memory length: 140650	epsilon: 0.9597555100008737
steps: 276	lr: 0.0001	evaluation reward: 1.43	
episode: 775	score: 1.0	memory length: 140801	epsilon: 0.9596060200008769
steps: 151	lr: 0.0001	evaluation reward: 1.44	
episode: 776	score: 0.0	memory length: 140924	epsilon: 0.9594842500008796
steps: 123	lr: 0.0001	evaluation reward: 1.44	
episode: 777	score: 0.0	memory length: 141047	epsilon: 0.9593624800008822
steps: 123	lr: 0.0001	evaluation reward: 1.43	
episode: 778	score: 4.0	memory length: 141363	epsilon: 0.9590496400008889
steps: 316	lr: 0.0001	evaluation reward: 1.46	
episode: 779	score: 1.0	memory length: 141513	epsilon: 0.9589011400008922
steps: 150	lr: 0.0001	evaluation reward: 1.46	
episode: 780	score: 9.0	memory length: 141899	epsilon: 0.9585190000009005
steps: 386	lr: 0.0001	evaluation reward: 1.55	
episode: 781	score: 1.0	memory length: 142069	epsilon: 0.9583507000009042
steps: 170	lr: 0.0001	evaluation reward: 1.55	
episode: 782	score: 1.0	memory length: 142241	epsilon: 0.9581804200009079

steps: 172	lr: 0.0001	evaluation reward: 1.54
episode: 783	score: 2.0	memory length: 142459 epsilon: 0.9579646000009125
steps: 218	lr: 0.0001	evaluation reward: 1.55
episode: 784	score: 3.0	memory length: 142671 epsilon: 0.9577547200009171
steps: 212	lr: 0.0001	evaluation reward: 1.57
episode: 785	score: 0.0	memory length: 142794 epsilon: 0.9576329500009197
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 786	score: 1.0	memory length: 142945 epsilon: 0.957483460000923
steps: 151	lr: 0.0001	evaluation reward: 1.55
episode: 787	score: 0.0	memory length: 143068 epsilon: 0.9573616900009256
steps: 123	lr: 0.0001	evaluation reward: 1.51
episode: 788	score: 1.0	memory length: 143219 epsilon: 0.9572122000009289
steps: 151	lr: 0.0001	evaluation reward: 1.5
episode: 789	score: 2.0	memory length: 143435 epsilon: 0.9569983600009335
steps: 216	lr: 0.0001	evaluation reward: 1.49
episode: 790	score: 1.0	memory length: 143586 epsilon: 0.9568488700009368
steps: 151	lr: 0.0001	evaluation reward: 1.49
episode: 791	score: 2.0	memory length: 143784 epsilon: 0.956652850000941
steps: 198	lr: 0.0001	evaluation reward: 1.5
episode: 792	score: 6.0	memory length: 144155 epsilon: 0.956285560000949
steps: 371	lr: 0.0001	evaluation reward: 1.55
episode: 793	score: 3.0	memory length: 144399 epsilon: 0.9560440000009542
steps: 244	lr: 0.0001	evaluation reward: 1.57
episode: 794	score: 2.0	memory length: 144596 epsilon: 0.9558489700009585
steps: 197	lr: 0.0001	evaluation reward: 1.57
episode: 795	score: 0.0	memory length: 144718 epsilon: 0.9557281900009611
steps: 122	lr: 0.0001	evaluation reward: 1.56
episode: 796	score: 2.0	memory length: 144916 epsilon: 0.9555321700009654
steps: 198	lr: 0.0001	evaluation reward: 1.58
episode: 797	score: 3.0	memory length: 145162 epsilon: 0.9552886300009706
steps: 246	lr: 0.0001	evaluation reward: 1.57
episode: 798	score: 2.0	memory length: 145359 epsilon: 0.9550936000009749
steps: 197	lr: 0.0001	evaluation reward: 1.59
episode: 799	score: 1.0	memory length: 145530 epsilon: 0.9549243100009785
steps: 171	lr: 0.0001	evaluation reward: 1.59
episode: 800	score: 0.0	memory length: 145652 epsilon: 0.9548035300009812
steps: 122	lr: 0.0001	evaluation reward: 1.59
episode: 801	score: 1.0	memory length: 145822 epsilon: 0.9546352300009848
steps: 170	lr: 0.0001	evaluation reward: 1.59
episode: 802	score: 4.0	memory length: 146138 epsilon: 0.9543223900009916
steps: 316	lr: 0.0001	evaluation reward: 1.62
episode: 803	score: 1.0	memory length: 146306 epsilon: 0.9541560700009952
steps: 168	lr: 0.0001	evaluation reward: 1.61
episode: 804	score: 0.0	memory length: 146429 epsilon: 0.9540343000009979
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 805	score: 2.0	memory length: 146648 epsilon: 0.9538174900010026
steps: 219	lr: 0.0001	evaluation reward: 1.6
episode: 806	score: 1.0	memory length: 146799 epsilon: 0.9536680000010058

steps: 151	lr: 0.0001	evaluation reward: 1.6
episode: 807	score: 1.0	memory length: 146970 epsilon: 0.9534987100010095
steps: 171	lr: 0.0001	evaluation reward: 1.57
episode: 808	score: 2.0	memory length: 147167 epsilon: 0.9533036800010137
steps: 197	lr: 0.0001	evaluation reward: 1.57
episode: 809	score: 3.0	memory length: 147392 epsilon: 0.9530809300010186
steps: 225	lr: 0.0001	evaluation reward: 1.6
episode: 810	score: 3.0	memory length: 147659 epsilon: 0.9528166000010243
steps: 267	lr: 0.0001	evaluation reward: 1.62
episode: 811	score: 4.0	memory length: 147937 epsilon: 0.9525413800010303
steps: 278	lr: 0.0001	evaluation reward: 1.65
episode: 812	score: 1.0	memory length: 148088 epsilon: 0.9523918900010335
steps: 151	lr: 0.0001	evaluation reward: 1.63
episode: 813	score: 1.0	memory length: 148238 epsilon: 0.9522433900010367
steps: 150	lr: 0.0001	evaluation reward: 1.64
episode: 814	score: 1.0	memory length: 148408 epsilon: 0.9520750900010404
steps: 170	lr: 0.0001	evaluation reward: 1.63
episode: 815	score: 2.0	memory length: 148606 epsilon: 0.9518790700010447
steps: 198	lr: 0.0001	evaluation reward: 1.62
episode: 816	score: 0.0	memory length: 148728 epsilon: 0.9517582900010473
steps: 122	lr: 0.0001	evaluation reward: 1.61
episode: 817	score: 1.0	memory length: 148896 epsilon: 0.9515919700010509
steps: 168	lr: 0.0001	evaluation reward: 1.59
episode: 818	score: 1.0	memory length: 149068 epsilon: 0.9514216900010546
steps: 172	lr: 0.0001	evaluation reward: 1.58
episode: 819	score: 0.0	memory length: 149191 epsilon: 0.9512999200010572
steps: 123	lr: 0.0001	evaluation reward: 1.57
episode: 820	score: 1.0	memory length: 149360 epsilon: 0.9511326100010609
steps: 169	lr: 0.0001	evaluation reward: 1.57
episode: 821	score: 0.0	memory length: 149482 epsilon: 0.9510118300010635
steps: 122	lr: 0.0001	evaluation reward: 1.55
episode: 822	score: 1.0	memory length: 149651 epsilon: 0.9508445200010671
steps: 169	lr: 0.0001	evaluation reward: 1.55
episode: 823	score: 2.0	memory length: 149849 epsilon: 0.9506485000010714
steps: 198	lr: 0.0001	evaluation reward: 1.56
episode: 824	score: 0.0	memory length: 149971 epsilon: 0.950527720001074
steps: 122	lr: 0.0001	evaluation reward: 1.56
episode: 825	score: 0.0	memory length: 150094 epsilon: 0.9504059500010766
steps: 123	lr: 0.0001	evaluation reward: 1.56
episode: 826	score: 3.0	memory length: 150340 epsilon: 0.9501624100010819
steps: 246	lr: 0.0001	evaluation reward: 1.54
episode: 827	score: 1.0	memory length: 150512 epsilon: 0.9499921300010856
steps: 172	lr: 0.0001	evaluation reward: 1.54
episode: 828	score: 2.0	memory length: 150712 epsilon: 0.9497941300010899
steps: 200	lr: 0.0001	evaluation reward: 1.56
episode: 829	score: 1.0	memory length: 150881 epsilon: 0.9496268200010936
steps: 169	lr: 0.0001	evaluation reward: 1.56
episode: 830	score: 3.0	memory length: 151151 epsilon: 0.9493595200010994

steps: 270	lr: 0.0001	evaluation reward: 1.57
episode: 831	score: 0.0	memory length: 151273 epsilon: 0.949238740001102
steps: 122	lr: 0.0001	evaluation reward: 1.55
episode: 832	score: 0.0	memory length: 151395 epsilon: 0.9491179600011046
steps: 122	lr: 0.0001	evaluation reward: 1.53
episode: 833	score: 3.0	memory length: 151642 epsilon: 0.9488734300011099
steps: 247	lr: 0.0001	evaluation reward: 1.55
episode: 834	score: 0.0	memory length: 151765 epsilon: 0.9487516600011126
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 835	score: 0.0	memory length: 151887 epsilon: 0.9486308800011152
steps: 122	lr: 0.0001	evaluation reward: 1.54
episode: 836	score: 3.0	memory length: 152118 epsilon: 0.9484021900011201
steps: 231	lr: 0.0001	evaluation reward: 1.57
episode: 837	score: 1.0	memory length: 152289 epsilon: 0.9482329000011238
steps: 171	lr: 0.0001	evaluation reward: 1.57
episode: 838	score: 0.0	memory length: 152412 epsilon: 0.9481111300011265
steps: 123	lr: 0.0001	evaluation reward: 1.55
episode: 839	score: 2.0	memory length: 152630 epsilon: 0.9478953100011311
steps: 218	lr: 0.0001	evaluation reward: 1.57
episode: 840	score: 0.0	memory length: 152753 epsilon: 0.9477735400011338
steps: 123	lr: 0.0001	evaluation reward: 1.54
episode: 841	score: 3.0	memory length: 153003 epsilon: 0.9475260400011392
steps: 250	lr: 0.0001	evaluation reward: 1.57
episode: 842	score: 2.0	memory length: 153183 epsilon: 0.947347840001143
steps: 180	lr: 0.0001	evaluation reward: 1.59
episode: 843	score: 0.0	memory length: 153306 epsilon: 0.9472260700011457
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 844	score: 3.0	memory length: 153553 epsilon: 0.946981540001151
steps: 247	lr: 0.0001	evaluation reward: 1.59
episode: 845	score: 2.0	memory length: 153771 epsilon: 0.9467657200011557
steps: 218	lr: 0.0001	evaluation reward: 1.58
episode: 846	score: 1.0	memory length: 153942 epsilon: 0.9465964300011593
steps: 171	lr: 0.0001	evaluation reward: 1.59
episode: 847	score: 1.0	memory length: 154093 epsilon: 0.9464469400011626
steps: 151	lr: 0.0001	evaluation reward: 1.6
episode: 848	score: 0.0	memory length: 154216 epsilon: 0.9463251700011652
steps: 123	lr: 0.0001	evaluation reward: 1.58
episode: 849	score: 3.0	memory length: 154459 epsilon: 0.9460846000011705
steps: 243	lr: 0.0001	evaluation reward: 1.6
episode: 850	score: 4.0	memory length: 154746 epsilon: 0.9458004700011766
steps: 287	lr: 0.0001	evaluation reward: 1.61
episode: 851	score: 7.0	memory length: 155065 epsilon: 0.9454846600011835
steps: 319	lr: 0.0001	evaluation reward: 1.66
episode: 852	score: 8.0	memory length: 155541 epsilon: 0.9450134200011937
steps: 476	lr: 0.0001	evaluation reward: 1.73
episode: 853	score: 2.0	memory length: 155761 epsilon: 0.9447956200011984
steps: 220	lr: 0.0001	evaluation reward: 1.75
episode: 854	score: 1.0	memory length: 155929 epsilon: 0.944629300001202

steps: 168	lr: 0.0001	evaluation reward: 1.73
episode: 855	score: 2.0	memory length: 156147 epsilon: 0.9444134800012067
steps: 218	lr: 0.0001	evaluation reward: 1.74
episode: 856	score: 2.0	memory length: 156346 epsilon: 0.944216470001211
steps: 199	lr: 0.0001	evaluation reward: 1.76
episode: 857	score: 2.0	memory length: 156544 epsilon: 0.9440204500012153
steps: 198	lr: 0.0001	evaluation reward: 1.76
episode: 858	score: 2.0	memory length: 156744 epsilon: 0.9438224500012196
steps: 200	lr: 0.0001	evaluation reward: 1.77
episode: 859	score: 0.0	memory length: 156866 epsilon: 0.9437016700012222
steps: 122	lr: 0.0001	evaluation reward: 1.75
episode: 860	score: 2.0	memory length: 157063 epsilon: 0.9435066400012264
steps: 197	lr: 0.0001	evaluation reward: 1.72
episode: 861	score: 2.0	memory length: 157260 epsilon: 0.9433116100012306
steps: 197	lr: 0.0001	evaluation reward: 1.72
episode: 862	score: 2.0	memory length: 157460 epsilon: 0.943113610001235
steps: 200	lr: 0.0001	evaluation reward: 1.72
episode: 863	score: 3.0	memory length: 157706 epsilon: 0.9428700700012402
steps: 246	lr: 0.0001	evaluation reward: 1.74
episode: 864	score: 3.0	memory length: 157951 epsilon: 0.9426275200012455
steps: 245	lr: 0.0001	evaluation reward: 1.72
episode: 865	score: 3.0	memory length: 158179 epsilon: 0.9424018000012504
steps: 228	lr: 0.0001	evaluation reward: 1.72
episode: 866	score: 1.0	memory length: 158329 epsilon: 0.9422533000012536
steps: 150	lr: 0.0001	evaluation reward: 1.71
episode: 867	score: 1.0	memory length: 158498 epsilon: 0.9420859900012573
steps: 169	lr: 0.0001	evaluation reward: 1.71
episode: 868	score: 2.0	memory length: 158696 epsilon: 0.9418899700012615
steps: 198	lr: 0.0001	evaluation reward: 1.72
episode: 869	score: 1.0	memory length: 158848 epsilon: 0.9417394900012648
steps: 152	lr: 0.0001	evaluation reward: 1.73
episode: 870	score: 4.0	memory length: 159127 epsilon: 0.9414632800012708
steps: 279	lr: 0.0001	evaluation reward: 1.77
episode: 871	score: 2.0	memory length: 159324 epsilon: 0.941268250001275
steps: 197	lr: 0.0001	evaluation reward: 1.77
episode: 872	score: 2.0	memory length: 159522 epsilon: 0.9410722300012793
steps: 198	lr: 0.0001	evaluation reward: 1.78
episode: 873	score: 3.0	memory length: 159766 epsilon: 0.9408306700012845
steps: 244	lr: 0.0001	evaluation reward: 1.8
episode: 874	score: 2.0	memory length: 159968 epsilon: 0.9406306900012888
steps: 202	lr: 0.0001	evaluation reward: 1.78
episode: 875	score: 1.0	memory length: 160136 epsilon: 0.9404643700012925
steps: 168	lr: 0.0001	evaluation reward: 1.78
episode: 876	score: 2.0	memory length: 160318 epsilon: 0.9402841900012964
steps: 182	lr: 0.0001	evaluation reward: 1.8
episode: 877	score: 2.0	memory length: 160517 epsilon: 0.9400871800013006
steps: 199	lr: 0.0001	evaluation reward: 1.82
episode: 878	score: 1.0	memory length: 160668 epsilon: 0.9399376900013039

steps: 151	lr: 0.0001	evaluation reward: 1.79
episode: 879	score: 3.0	memory length: 160894 epsilon: 0.9397139500013088
steps: 226	lr: 0.0001	evaluation reward: 1.81
episode: 880	score: 2.0	memory length: 161097 epsilon: 0.9395129800013131
steps: 203	lr: 0.0001	evaluation reward: 1.74
episode: 881	score: 4.0	memory length: 161376 epsilon: 0.9392367700013191
steps: 279	lr: 0.0001	evaluation reward: 1.77
episode: 882	score: 1.0	memory length: 161527 epsilon: 0.9390872800013224
steps: 151	lr: 0.0001	evaluation reward: 1.77
episode: 883	score: 6.0	memory length: 161903 epsilon: 0.9387150400013304
steps: 376	lr: 0.0001	evaluation reward: 1.81
episode: 884	score: 2.0	memory length: 162121 epsilon: 0.9384992200013351
steps: 218	lr: 0.0001	evaluation reward: 1.8
episode: 885	score: 2.0	memory length: 162319 epsilon: 0.9383032000013394
steps: 198	lr: 0.0001	evaluation reward: 1.82
episode: 886	score: 0.0	memory length: 162442 epsilon: 0.938181430001342
steps: 123	lr: 0.0001	evaluation reward: 1.81
episode: 887	score: 4.0	memory length: 162721 epsilon: 0.937905220001348
steps: 279	lr: 0.0001	evaluation reward: 1.85
episode: 888	score: 0.0	memory length: 162844 epsilon: 0.9377834500013507
steps: 123	lr: 0.0001	evaluation reward: 1.84
episode: 889	score: 2.0	memory length: 163025 epsilon: 0.9376042600013546
steps: 181	lr: 0.0001	evaluation reward: 1.84
episode: 890	score: 3.0	memory length: 163269 epsilon: 0.9373627000013598
steps: 244	lr: 0.0001	evaluation reward: 1.86
episode: 891	score: 2.0	memory length: 163487 epsilon: 0.9371468800013645
steps: 218	lr: 0.0001	evaluation reward: 1.86
episode: 892	score: 5.0	memory length: 163796 epsilon: 0.9368409700013711
steps: 309	lr: 0.0001	evaluation reward: 1.85
episode: 893	score: 2.0	memory length: 164013 epsilon: 0.9366261400013758
steps: 217	lr: 0.0001	evaluation reward: 1.84
episode: 894	score: 3.0	memory length: 164283 epsilon: 0.9363588400013816
steps: 270	lr: 0.0001	evaluation reward: 1.85
episode: 895	score: 2.0	memory length: 164481 epsilon: 0.9361628200013858
steps: 198	lr: 0.0001	evaluation reward: 1.87
episode: 896	score: 0.0	memory length: 164603 epsilon: 0.9360420400013885
steps: 122	lr: 0.0001	evaluation reward: 1.85
episode: 897	score: 1.0	memory length: 164774 epsilon: 0.9358727500013921
steps: 171	lr: 0.0001	evaluation reward: 1.83
episode: 898	score: 1.0	memory length: 164945 epsilon: 0.9357034600013958
steps: 171	lr: 0.0001	evaluation reward: 1.82
episode: 899	score: 2.0	memory length: 165142 epsilon: 0.9355084300014
steps: 197	lr: 0.0001	evaluation reward: 1.83
episode: 900	score: 3.0	memory length: 165367 epsilon: 0.9352856800014049
steps: 225	lr: 0.0001	evaluation reward: 1.86
episode: 901	score: 2.0	memory length: 165567 epsilon: 0.9350876800014092
steps: 200	lr: 0.0001	evaluation reward: 1.87
episode: 902	score: 1.0	memory length: 165718 epsilon: 0.9349381900014124

steps: 151	lr: 0.0001	evaluation reward: 1.84	
episode: 903	score: 0.0	memory length: 165841	epsilon: 0.9348164200014151
steps: 123	lr: 0.0001	evaluation reward: 1.83	
episode: 904	score: 3.0	memory length: 166087	epsilon: 0.9345728800014204
steps: 246	lr: 0.0001	evaluation reward: 1.86	
episode: 905	score: 7.0	memory length: 166406	epsilon: 0.9342570700014272
steps: 319	lr: 0.0001	evaluation reward: 1.91	
episode: 906	score: 5.0	memory length: 166714	epsilon: 0.9339521500014338
steps: 308	lr: 0.0001	evaluation reward: 1.95	
episode: 907	score: 1.0	memory length: 166884	epsilon: 0.9337838500014375
steps: 170	lr: 0.0001	evaluation reward: 1.95	
episode: 908	score: 1.0	memory length: 167052	epsilon: 0.9336175300014411
steps: 168	lr: 0.0001	evaluation reward: 1.94	
episode: 909	score: 2.0	memory length: 167250	epsilon: 0.9334215100014454
steps: 198	lr: 0.0001	evaluation reward: 1.93	
episode: 910	score: 0.0	memory length: 167372	epsilon: 0.9333007300014448
steps: 122	lr: 0.0001	evaluation reward: 1.9	
episode: 911	score: 2.0	memory length: 167590	epsilon: 0.9330849100014527
steps: 218	lr: 0.0001	evaluation reward: 1.88	
episode: 912	score: 5.0	memory length: 167894	epsilon: 0.9327839500014592
steps: 304	lr: 0.0001	evaluation reward: 1.92	
episode: 913	score: 4.0	memory length: 168151	epsilon: 0.9325295200014647
steps: 257	lr: 0.0001	evaluation reward: 1.95	
episode: 914	score: 2.0	memory length: 168330	epsilon: 0.9323523100014686
steps: 179	lr: 0.0001	evaluation reward: 1.96	
episode: 915	score: 0.0	memory length: 168452	epsilon: 0.9322315300014712
steps: 122	lr: 0.0001	evaluation reward: 1.94	
episode: 916	score: 2.0	memory length: 168668	epsilon: 0.9320176900014758
steps: 216	lr: 0.0001	evaluation reward: 1.96	
episode: 917	score: 1.0	memory length: 168836	epsilon: 0.9318513700014794
steps: 168	lr: 0.0001	evaluation reward: 1.96	
episode: 918	score: 0.0	memory length: 168959	epsilon: 0.9317296000014821
steps: 123	lr: 0.0001	evaluation reward: 1.95	
episode: 919	score: 4.0	memory length: 169235	epsilon: 0.931456360001488
steps: 276	lr: 0.0001	evaluation reward: 1.99	
episode: 920	score: 0.0	memory length: 169358	epsilon: 0.9313345900014907
steps: 123	lr: 0.0001	evaluation reward: 1.98	
episode: 921	score: 0.0	memory length: 169481	epsilon: 0.9312128200014933
steps: 123	lr: 0.0001	evaluation reward: 1.98	
episode: 922	score: 2.0	memory length: 169665	epsilon: 0.9310306600014973
steps: 184	lr: 0.0001	evaluation reward: 1.99	
episode: 923	score: 0.0	memory length: 169788	epsilon: 0.9309088900014999
steps: 123	lr: 0.0001	evaluation reward: 1.97	
episode: 924	score: 2.0	memory length: 169985	epsilon: 0.9307138600015041
steps: 197	lr: 0.0001	evaluation reward: 1.99	
episode: 925	score: 0.0	memory length: 170108	epsilon: 0.9305920900015068
steps: 123	lr: 0.0001	evaluation reward: 1.99	
episode: 926	score: 1.0	memory length: 170277	epsilon: 0.9304247800015104

steps: 169	lr: 0.0001	evaluation reward: 1.97
episode: 927	score: 2.0	memory length: 170476 epsilon: 0.9302277700015147
steps: 199	lr: 0.0001	evaluation reward: 1.98
episode: 928	score: 1.0	memory length: 170645 epsilon: 0.9300604600015183
steps: 169	lr: 0.0001	evaluation reward: 1.97
episode: 929	score: 2.0	memory length: 170863 epsilon: 0.929844640001523
steps: 218	lr: 0.0001	evaluation reward: 1.98
episode: 930	score: 2.0	memory length: 171042 epsilon: 0.9296674300015269
steps: 179	lr: 0.0001	evaluation reward: 1.97
episode: 931	score: 1.0	memory length: 171210 epsilon: 0.9295011100015305
steps: 168	lr: 0.0001	evaluation reward: 1.98
episode: 932	score: 3.0	memory length: 171459 epsilon: 0.9292546000015358
steps: 249	lr: 0.0001	evaluation reward: 2.01
episode: 933	score: 2.0	memory length: 171661 epsilon: 0.9290546200015402
steps: 202	lr: 0.0001	evaluation reward: 2.0
episode: 934	score: 3.0	memory length: 171911 epsilon: 0.9288071200015455
steps: 250	lr: 0.0001	evaluation reward: 2.03
episode: 935	score: 0.0	memory length: 172033 epsilon: 0.9286863400015481
steps: 122	lr: 0.0001	evaluation reward: 2.03
episode: 936	score: 0.0	memory length: 172156 epsilon: 0.9285645700015508
steps: 123	lr: 0.0001	evaluation reward: 2.0
episode: 937	score: 2.0	memory length: 172354 epsilon: 0.928368550001555
steps: 198	lr: 0.0001	evaluation reward: 2.01
episode: 938	score: 2.0	memory length: 172572 epsilon: 0.9281527300015597
steps: 218	lr: 0.0001	evaluation reward: 2.03
episode: 939	score: 3.0	memory length: 172818 epsilon: 0.927909190001565
steps: 246	lr: 0.0001	evaluation reward: 2.04
episode: 940	score: 1.0	memory length: 172969 epsilon: 0.9277597000015683
steps: 151	lr: 0.0001	evaluation reward: 2.05
episode: 941	score: 3.0	memory length: 173234 epsilon: 0.927497350001574
steps: 265	lr: 0.0001	evaluation reward: 2.05
episode: 942	score: 2.0	memory length: 173431 epsilon: 0.9273023200015782
steps: 197	lr: 0.0001	evaluation reward: 2.05
episode: 943	score: 1.0	memory length: 173599 epsilon: 0.9271360000015818
steps: 168	lr: 0.0001	evaluation reward: 2.06
episode: 944	score: 4.0	memory length: 173860 epsilon: 0.9268776100015874
steps: 261	lr: 0.0001	evaluation reward: 2.07
episode: 945	score: 4.0	memory length: 174136 epsilon: 0.9266043700015933
steps: 276	lr: 0.0001	evaluation reward: 2.09
episode: 946	score: 1.0	memory length: 174286 epsilon: 0.9264558700015966
steps: 150	lr: 0.0001	evaluation reward: 2.09
episode: 947	score: 2.0	memory length: 174504 epsilon: 0.9262400500016013
steps: 218	lr: 0.0001	evaluation reward: 2.1
episode: 948	score: 0.0	memory length: 174627 epsilon: 0.9261182800016039
steps: 123	lr: 0.0001	evaluation reward: 2.1
episode: 949	score: 0.0	memory length: 174749 epsilon: 0.9259975000016065
steps: 122	lr: 0.0001	evaluation reward: 2.07
episode: 950	score: 3.0	memory length: 174976 epsilon: 0.9257727700016114

steps: 227	lr: 0.0001	evaluation reward: 2.06
episode: 951	score: 2.0	memory length: 175192 epsilon: 0.925558930001616
steps: 216	lr: 0.0001	evaluation reward: 2.01
episode: 952	score: 3.0	memory length: 175402 epsilon: 0.9253510300016206
steps: 210	lr: 0.0001	evaluation reward: 1.96
episode: 953	score: 0.0	memory length: 175525 epsilon: 0.9252292600016232
steps: 123	lr: 0.0001	evaluation reward: 1.94
episode: 954	score: 2.0	memory length: 175722 epsilon: 0.9250342300016274
steps: 197	lr: 0.0001	evaluation reward: 1.95
episode: 955	score: 2.0	memory length: 175920 epsilon: 0.9248382100016317
steps: 198	lr: 0.0001	evaluation reward: 1.95
episode: 956	score: 2.0	memory length: 176117 epsilon: 0.9246431800016359
steps: 197	lr: 0.0001	evaluation reward: 1.95
episode: 957	score: 1.0	memory length: 176285 epsilon: 0.9244768600016395
steps: 168	lr: 0.0001	evaluation reward: 1.94
episode: 958	score: 1.0	memory length: 176436 epsilon: 0.9243273700016428
steps: 151	lr: 0.0001	evaluation reward: 1.93
episode: 959	score: 0.0	memory length: 176559 epsilon: 0.9242056000016454
steps: 123	lr: 0.0001	evaluation reward: 1.93
episode: 960	score: 2.0	memory length: 176757 epsilon: 0.9240095800016497
steps: 198	lr: 0.0001	evaluation reward: 1.93
episode: 961	score: 1.0	memory length: 176908 epsilon: 0.9238600900016529
steps: 151	lr: 0.0001	evaluation reward: 1.92
episode: 962	score: 1.0	memory length: 177077 epsilon: 0.9236927800016566
steps: 169	lr: 0.0001	evaluation reward: 1.91
episode: 963	score: 0.0	memory length: 177200 epsilon: 0.9235710100016592
steps: 123	lr: 0.0001	evaluation reward: 1.88
episode: 964	score: 1.0	memory length: 177369 epsilon: 0.9234037000016628
steps: 169	lr: 0.0001	evaluation reward: 1.86
episode: 965	score: 3.0	memory length: 177579 epsilon: 0.9231958000016673
steps: 210	lr: 0.0001	evaluation reward: 1.86
episode: 966	score: 4.0	memory length: 177840 epsilon: 0.922937410001673
steps: 261	lr: 0.0001	evaluation reward: 1.89
episode: 967	score: 2.0	memory length: 178059 epsilon: 0.9227206000016777
steps: 219	lr: 0.0001	evaluation reward: 1.9
episode: 968	score: 2.0	memory length: 178278 epsilon: 0.9225037900016824
steps: 219	lr: 0.0001	evaluation reward: 1.9
episode: 969	score: 0.0	memory length: 178400 epsilon: 0.922383010001685
steps: 122	lr: 0.0001	evaluation reward: 1.89
episode: 970	score: 0.0	memory length: 178523 epsilon: 0.9222612400016876
steps: 123	lr: 0.0001	evaluation reward: 1.85
episode: 971	score: 1.0	memory length: 178673 epsilon: 0.9221127400016909
steps: 150	lr: 0.0001	evaluation reward: 1.84
episode: 972	score: 1.0	memory length: 178841 epsilon: 0.9219464200016945
steps: 168	lr: 0.0001	evaluation reward: 1.83
episode: 973	score: 2.0	memory length: 179040 epsilon: 0.9217494100016987
steps: 199	lr: 0.0001	evaluation reward: 1.82
episode: 974	score: 3.0	memory length: 179287 epsilon: 0.921504880001704

steps: 247	lr: 0.0001	evaluation reward: 1.83
episode: 975	score: 4.0	memory length: 179566 epsilon: 0.92122867000171
steps: 279	lr: 0.0001	evaluation reward: 1.86
episode: 976	score: 2.0	memory length: 179763 epsilon: 0.9210336400017143
steps: 197	lr: 0.0001	evaluation reward: 1.86
episode: 977	score: 0.0	memory length: 179886 epsilon: 0.9209118700017169
steps: 123	lr: 0.0001	evaluation reward: 1.84
episode: 978	score: 5.0	memory length: 180202 epsilon: 0.9205990300017237
steps: 316	lr: 0.0001	evaluation reward: 1.88
episode: 979	score: 0.0	memory length: 180324 epsilon: 0.9204782500017263
steps: 122	lr: 0.0001	evaluation reward: 1.85
episode: 980	score: 2.0	memory length: 180540 epsilon: 0.920264410001731
steps: 216	lr: 0.0001	evaluation reward: 1.85
episode: 981	score: 1.0	memory length: 180710 epsilon: 0.9200961100017346
steps: 170	lr: 0.0001	evaluation reward: 1.82
episode: 982	score: 3.0	memory length: 180956 epsilon: 0.9198525700017399
steps: 246	lr: 0.0001	evaluation reward: 1.84
episode: 983	score: 2.0	memory length: 181155 epsilon: 0.9196555600017442
steps: 199	lr: 0.0001	evaluation reward: 1.8
episode: 984	score: 2.0	memory length: 181353 epsilon: 0.9194595400017485
steps: 198	lr: 0.0001	evaluation reward: 1.8
episode: 985	score: 2.0	memory length: 181551 epsilon: 0.9192635200017527
steps: 198	lr: 0.0001	evaluation reward: 1.8
episode: 986	score: 2.0	memory length: 181769 epsilon: 0.9190477000017574
steps: 218	lr: 0.0001	evaluation reward: 1.82
episode: 987	score: 2.0	memory length: 181966 epsilon: 0.9188526700017616
steps: 197	lr: 0.0001	evaluation reward: 1.8
episode: 988	score: 3.0	memory length: 182216 epsilon: 0.918605170001767
steps: 250	lr: 0.0001	evaluation reward: 1.83
episode: 989	score: 0.0	memory length: 182338 epsilon: 0.9184843900017696
steps: 122	lr: 0.0001	evaluation reward: 1.81
episode: 990	score: 2.0	memory length: 182553 epsilon: 0.9182715400017742
steps: 215	lr: 0.0001	evaluation reward: 1.8
episode: 991	score: 1.0	memory length: 182704 epsilon: 0.9181220500017775
steps: 151	lr: 0.0001	evaluation reward: 1.79
episode: 992	score: 3.0	memory length: 182952 epsilon: 0.9178765300017828
steps: 248	lr: 0.0001	evaluation reward: 1.77
episode: 993	score: 0.0	memory length: 183075 epsilon: 0.9177547600017855
steps: 123	lr: 0.0001	evaluation reward: 1.75
episode: 994	score: 1.0	memory length: 183247 epsilon: 0.9175844800017892
steps: 172	lr: 0.0001	evaluation reward: 1.73
episode: 995	score: 2.0	memory length: 183465 epsilon: 0.9173686600017938
steps: 218	lr: 0.0001	evaluation reward: 1.73
episode: 996	score: 4.0	memory length: 183741 epsilon: 0.9170954200017998
steps: 276	lr: 0.0001	evaluation reward: 1.77
episode: 997	score: 3.0	memory length: 184008 epsilon: 0.9168310900018055
steps: 267	lr: 0.0001	evaluation reward: 1.79
episode: 998	score: 3.0	memory length: 184237 epsilon: 0.9166043800018104

steps: 229	lr: 0.0001	evaluation reward: 1.81	
episode: 999	score: 2.0	memory length: 184435	epsilon: 0.9164083600018147
steps: 198	lr: 0.0001	evaluation reward: 1.81	
episode: 1000	score: 2.0	memory length: 184635	epsilon: 0.916210360001819
steps: 200	lr: 0.0001	evaluation reward: 1.8	
episode: 1001	score: 2.0	memory length: 184833	epsilon: 0.9160143400018232
steps: 198	lr: 0.0001	evaluation reward: 1.8	
episode: 1002	score: 0.0	memory length: 184956	epsilon: 0.9158925700018259
steps: 123	lr: 0.0001	evaluation reward: 1.79	
episode: 1003	score: 0.0	memory length: 185078	epsilon: 0.9157717900018285
steps: 122	lr: 0.0001	evaluation reward: 1.79	
episode: 1004	score: 2.0	memory length: 185295	epsilon: 0.9155569600018332
steps: 217	lr: 0.0001	evaluation reward: 1.78	
episode: 1005	score: 0.0	memory length: 185417	epsilon: 0.9154361800018358
steps: 122	lr: 0.0001	evaluation reward: 1.71	
episode: 1006	score: 2.0	memory length: 185616	epsilon: 0.9152391700018401
steps: 199	lr: 0.0001	evaluation reward: 1.68	
episode: 1007	score: 0.0	memory length: 185738	epsilon: 0.9151183900018427
steps: 122	lr: 0.0001	evaluation reward: 1.67	
episode: 1008	score: 0.0	memory length: 185861	epsilon: 0.9149966200018453
steps: 123	lr: 0.0001	evaluation reward: 1.66	
episode: 1009	score: 2.0	memory length: 186078	epsilon: 0.91478179000185
steps: 217	lr: 0.0001	evaluation reward: 1.66	
episode: 1010	score: 1.0	memory length: 186248	epsilon: 0.9146134900018537
steps: 170	lr: 0.0001	evaluation reward: 1.67	
episode: 1011	score: 0.0	memory length: 186371	epsilon: 0.9144917200018563
steps: 123	lr: 0.0001	evaluation reward: 1.65	
episode: 1012	score: 3.0	memory length: 186582	epsilon: 0.9142828300018608
steps: 211	lr: 0.0001	evaluation reward: 1.63	
episode: 1013	score: 0.0	memory length: 186705	epsilon: 0.9141610600018635
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 1014	score: 2.0	memory length: 186903	epsilon: 0.9139650400018677
steps: 198	lr: 0.0001	evaluation reward: 1.59	
episode: 1015	score: 5.0	memory length: 187231	epsilon: 0.9136403200018748
steps: 328	lr: 0.0001	evaluation reward: 1.64	
episode: 1016	score: 3.0	memory length: 187494	epsilon: 0.9133799500018804
steps: 263	lr: 0.0001	evaluation reward: 1.65	
episode: 1017	score: 1.0	memory length: 187664	epsilon: 0.9132116500018841
steps: 170	lr: 0.0001	evaluation reward: 1.65	
episode: 1018	score: 0.0	memory length: 187786	epsilon: 0.9130908700018867
steps: 122	lr: 0.0001	evaluation reward: 1.65	
episode: 1019	score: 1.0	memory length: 187937	epsilon: 0.91294138000189
steps: 151	lr: 0.0001	evaluation reward: 1.62	
episode: 1020	score: 0.0	memory length: 188059	epsilon: 0.9128206000018926
steps: 122	lr: 0.0001	evaluation reward: 1.62	
episode: 1021	score: 1.0	memory length: 188212	epsilon: 0.9126691300018959
steps: 153	lr: 0.0001	evaluation reward: 1.63	
episode: 1022	score: 3.0	memory length: 188457	epsilon: 0.9124265800019011

steps: 245	lr: 0.0001	evaluation reward: 1.64	
episode: 1023	score: 2.0	memory length: 188676	epsilon: 0.9122097700019058
steps: 219	lr: 0.0001	evaluation reward: 1.66	
episode: 1024	score: 1.0	memory length: 188828	epsilon: 0.9120592900019091
steps: 152	lr: 0.0001	evaluation reward: 1.65	
episode: 1025	score: 1.0	memory length: 188997	epsilon: 0.9118919800019127
steps: 169	lr: 0.0001	evaluation reward: 1.66	
episode: 1026	score: 3.0	memory length: 189267	epsilon: 0.9116246800019185
steps: 270	lr: 0.0001	evaluation reward: 1.68	
episode: 1027	score: 2.0	memory length: 189470	epsilon: 0.9114237100019229
steps: 203	lr: 0.0001	evaluation reward: 1.68	
episode: 1028	score: 3.0	memory length: 189696	epsilon: 0.9111999700019278
steps: 226	lr: 0.0001	evaluation reward: 1.7	
episode: 1029	score: 2.0	memory length: 189912	epsilon: 0.9109861300019324
steps: 216	lr: 0.0001	evaluation reward: 1.7	
episode: 1030	score: 3.0	memory length: 190181	epsilon: 0.9107198200019382
steps: 269	lr: 0.0001	evaluation reward: 1.71	
episode: 1031	score: 1.0	memory length: 190332	epsilon: 0.9105703300019414
steps: 151	lr: 0.0001	evaluation reward: 1.71	
episode: 1032	score: 1.0	memory length: 190500	epsilon: 0.910404010001945
steps: 168	lr: 0.0001	evaluation reward: 1.69	
episode: 1033	score: 1.0	memory length: 190669	epsilon: 0.9102367000019487
steps: 169	lr: 0.0001	evaluation reward: 1.68	
episode: 1034	score: 1.0	memory length: 190837	epsilon: 0.9100703800019523
steps: 168	lr: 0.0001	evaluation reward: 1.66	
episode: 1035	score: 1.0	memory length: 191006	epsilon: 0.9099030700019559
steps: 169	lr: 0.0001	evaluation reward: 1.67	
episode: 1036	score: 1.0	memory length: 191157	epsilon: 0.9097535800019592
steps: 151	lr: 0.0001	evaluation reward: 1.68	
episode: 1037	score: 2.0	memory length: 191374	epsilon: 0.9095387500019638
steps: 217	lr: 0.0001	evaluation reward: 1.68	
episode: 1038	score: 1.0	memory length: 191543	epsilon: 0.9093714400019675
steps: 169	lr: 0.0001	evaluation reward: 1.67	
episode: 1039	score: 2.0	memory length: 191722	epsilon: 0.9091942300019713
steps: 179	lr: 0.0001	evaluation reward: 1.66	
episode: 1040	score: 2.0	memory length: 191921	epsilon: 0.9089972200019756
steps: 199	lr: 0.0001	evaluation reward: 1.67	
episode: 1041	score: 2.0	memory length: 192139	epsilon: 0.9087814000019803
steps: 218	lr: 0.0001	evaluation reward: 1.66	
episode: 1042	score: 0.0	memory length: 192262	epsilon: 0.9086596300019829
steps: 123	lr: 0.0001	evaluation reward: 1.64	
episode: 1043	score: 0.0	memory length: 192385	epsilon: 0.9085378600019856
steps: 123	lr: 0.0001	evaluation reward: 1.63	
episode: 1044	score: 1.0	memory length: 192538	epsilon: 0.9083863900019888
steps: 153	lr: 0.0001	evaluation reward: 1.6	
episode: 1045	score: 0.0	memory length: 192661	epsilon: 0.9082646200019915
steps: 123	lr: 0.0001	evaluation reward: 1.56	
episode: 1046	score: 3.0	memory length: 192909	epsilon: 0.9080191000019968

steps: 248	lr: 0.0001	evaluation reward: 1.58	
episode: 1047	score: 3.0	memory length: 193180	epsilon: 0.9077508100020026
steps: 271	lr: 0.0001	evaluation reward: 1.59	
episode: 1048	score: 3.0	memory length: 193426	epsilon: 0.9075072700020079
steps: 246	lr: 0.0001	evaluation reward: 1.62	
episode: 1049	score: 1.0	memory length: 193577	epsilon: 0.9073577800020112
steps: 151	lr: 0.0001	evaluation reward: 1.63	
episode: 1050	score: 0.0	memory length: 193700	epsilon: 0.9072360100020138
steps: 123	lr: 0.0001	evaluation reward: 1.6	
episode: 1051	score: 2.0	memory length: 193915	epsilon: 0.9070231600020184
steps: 215	lr: 0.0001	evaluation reward: 1.6	
episode: 1052	score: 0.0	memory length: 194038	epsilon: 0.9069013900020211
steps: 123	lr: 0.0001	evaluation reward: 1.57	
episode: 1053	score: 3.0	memory length: 194287	epsilon: 0.9066548800020264
steps: 249	lr: 0.0001	evaluation reward: 1.6	
episode: 1054	score: 0.0	memory length: 194410	epsilon: 0.9065331100020291
steps: 123	lr: 0.0001	evaluation reward: 1.58	
episode: 1055	score: 0.0	memory length: 194532	epsilon: 0.9064123300020317
steps: 122	lr: 0.0001	evaluation reward: 1.56	
episode: 1056	score: 0.0	memory length: 194655	epsilon: 0.9062905600020343
steps: 123	lr: 0.0001	evaluation reward: 1.54	
episode: 1057	score: 3.0	memory length: 194922	epsilon: 0.9060262300020401
steps: 267	lr: 0.0001	evaluation reward: 1.56	
episode: 1058	score: 4.0	memory length: 195215	epsilon: 0.9057361600020464
steps: 293	lr: 0.0001	evaluation reward: 1.59	
episode: 1059	score: 0.0	memory length: 195338	epsilon: 0.905614390002049
steps: 123	lr: 0.0001	evaluation reward: 1.59	
episode: 1060	score: 2.0	memory length: 195555	epsilon: 0.9053995600020537
steps: 217	lr: 0.0001	evaluation reward: 1.59	
episode: 1061	score: 3.0	memory length: 195818	epsilon: 0.9051391900020593
steps: 263	lr: 0.0001	evaluation reward: 1.61	
episode: 1062	score: 6.0	memory length: 196183	epsilon: 0.9047778400020672
steps: 365	lr: 0.0001	evaluation reward: 1.66	
episode: 1063	score: 2.0	memory length: 196382	epsilon: 0.9045808300020715
steps: 199	lr: 0.0001	evaluation reward: 1.68	
episode: 1064	score: 0.0	memory length: 196505	epsilon: 0.9044590600020741
steps: 123	lr: 0.0001	evaluation reward: 1.67	
episode: 1065	score: 3.0	memory length: 196731	epsilon: 0.904235320002079
steps: 226	lr: 0.0001	evaluation reward: 1.67	
episode: 1066	score: 5.0	memory length: 197055	epsilon: 0.9039145600020859
steps: 324	lr: 0.0001	evaluation reward: 1.68	
episode: 1067	score: 4.0	memory length: 197348	epsilon: 0.9036244900020922
steps: 293	lr: 0.0001	evaluation reward: 1.7	
episode: 1068	score: 6.0	memory length: 197742	epsilon: 0.9032344300021007
steps: 394	lr: 0.0001	evaluation reward: 1.74	
episode: 1069	score: 1.0	memory length: 197911	epsilon: 0.9030671200021043
steps: 169	lr: 0.0001	evaluation reward: 1.75	
episode: 1070	score: 2.0	memory length: 198109	epsilon: 0.9028711000021086

steps: 198	lr: 0.0001	evaluation reward: 1.77	
episode: 1071	score: 4.0	memory length: 198384	epsilon: 0.9025988500021145
steps: 275	lr: 0.0001	evaluation reward: 1.8	
episode: 1072	score: 2.0	memory length: 198600	epsilon: 0.9023850100021191
steps: 216	lr: 0.0001	evaluation reward: 1.81	
episode: 1073	score: 3.0	memory length: 198849	epsilon: 0.9021385000021245
steps: 249	lr: 0.0001	evaluation reward: 1.82	
episode: 1074	score: 3.0	memory length: 199096	epsilon: 0.9018939700021298
steps: 247	lr: 0.0001	evaluation reward: 1.82	
episode: 1075	score: 2.0	memory length: 199294	epsilon: 0.901697950002134
steps: 198	lr: 0.0001	evaluation reward: 1.8	
episode: 1076	score: 3.0	memory length: 199541	epsilon: 0.9014534200021393
steps: 247	lr: 0.0001	evaluation reward: 1.81	
episode: 1077	score: 0.0	memory length: 199663	epsilon: 0.901332640002142
steps: 122	lr: 0.0001	evaluation reward: 1.81	
episode: 1078	score: 0.0	memory length: 199785	epsilon: 0.9012118600021446
steps: 122	lr: 0.0001	evaluation reward: 1.76	
episode: 1079	score: 3.0	memory length: 200013	epsilon: 0.9009861400021495
steps: 228	lr: 4e-05	evaluation reward: 1.79	
episode: 1080	score: 2.0	memory length: 200210	epsilon: 0.9007911100021537
steps: 197	lr: 4e-05	evaluation reward: 1.79	
episode: 1081	score: 1.0	memory length: 200378	epsilon: 0.9006247900021573
steps: 168	lr: 4e-05	evaluation reward: 1.79	
episode: 1082	score: 1.0	memory length: 200546	epsilon: 0.900458470002161
steps: 168	lr: 4e-05	evaluation reward: 1.77	
episode: 1083	score: 3.0	memory length: 200793	epsilon: 0.9002139400021663
steps: 247	lr: 4e-05	evaluation reward: 1.78	
episode: 1084	score: 1.0	memory length: 200944	epsilon: 0.9000644500021695
steps: 151	lr: 4e-05	evaluation reward: 1.77	
episode: 1085	score: 2.0	memory length: 201143	epsilon: 0.8998674400021738
steps: 199	lr: 4e-05	evaluation reward: 1.77	
episode: 1086	score: 3.0	memory length: 201387	epsilon: 0.899625880002179
steps: 244	lr: 4e-05	evaluation reward: 1.78	
episode: 1087	score: 2.0	memory length: 201585	epsilon: 0.8994298600021833
steps: 198	lr: 4e-05	evaluation reward: 1.78	
episode: 1088	score: 7.0	memory length: 201999	epsilon: 0.8990200000021922
steps: 414	lr: 4e-05	evaluation reward: 1.82	
episode: 1089	score: 1.0	memory length: 202170	epsilon: 0.8988507100021959
steps: 171	lr: 4e-05	evaluation reward: 1.83	
episode: 1090	score: 1.0	memory length: 202341	epsilon: 0.8986814200021995
steps: 171	lr: 4e-05	evaluation reward: 1.82	
episode: 1091	score: 4.0	memory length: 202618	epsilon: 0.8984071900022055
steps: 277	lr: 4e-05	evaluation reward: 1.85	
episode: 1092	score: 4.0	memory length: 202916	epsilon: 0.8981121700022119
steps: 298	lr: 4e-05	evaluation reward: 1.86	
episode: 1093	score: 0.0	memory length: 203039	epsilon: 0.8979904000022145
steps: 123	lr: 4e-05	evaluation reward: 1.86	
episode: 1094	score: 3.0	memory length: 203288	epsilon: 0.8977438900022199

steps: 249	lr: 4e-05	evaluation reward: 1.88	
episode: 1095	score: 3.0	memory length: 203514	epsilon: 0.8975201500022247
steps: 226	lr: 4e-05	evaluation reward: 1.89	
episode: 1096	score: 4.0	memory length: 203780	epsilon: 0.8972568100022305
steps: 266	lr: 4e-05	evaluation reward: 1.89	
episode: 1097	score: 1.0	memory length: 203931	epsilon: 0.8971073200022337
steps: 151	lr: 4e-05	evaluation reward: 1.87	
episode: 1098	score: 0.0	memory length: 204053	epsilon: 0.8969865400022363
steps: 122	lr: 4e-05	evaluation reward: 1.84	
episode: 1099	score: 1.0	memory length: 204204	epsilon: 0.8968370500022396
steps: 151	lr: 4e-05	evaluation reward: 1.83	
episode: 1100	score: 3.0	memory length: 204451	epsilon: 0.8965925200022449
steps: 247	lr: 4e-05	evaluation reward: 1.84	
episode: 1101	score: 1.0	memory length: 204621	epsilon: 0.8964242200022485
steps: 170	lr: 4e-05	evaluation reward: 1.83	
episode: 1102	score: 4.0	memory length: 204896	epsilon: 0.8961519700022544
steps: 275	lr: 4e-05	evaluation reward: 1.87	
episode: 1103	score: 2.0	memory length: 205094	epsilon: 0.8959559500022587
steps: 198	lr: 4e-05	evaluation reward: 1.89	
episode: 1104	score: 4.0	memory length: 205389	epsilon: 0.895663900002265
steps: 295	lr: 4e-05	evaluation reward: 1.91	
episode: 1105	score: 1.0	memory length: 205540	epsilon: 0.8955144100022683
steps: 151	lr: 4e-05	evaluation reward: 1.92	
episode: 1106	score: 2.0	memory length: 205759	epsilon: 0.895297600002273
steps: 219	lr: 4e-05	evaluation reward: 1.92	
episode: 1107	score: 1.0	memory length: 205910	epsilon: 0.8951481100022762
steps: 151	lr: 4e-05	evaluation reward: 1.93	
episode: 1108	score: 2.0	memory length: 206127	epsilon: 0.8949332800022809
steps: 217	lr: 4e-05	evaluation reward: 1.95	
episode: 1109	score: 5.0	memory length: 206438	epsilon: 0.8946253900022876
steps: 311	lr: 4e-05	evaluation reward: 1.98	
episode: 1110	score: 2.0	memory length: 206656	epsilon: 0.8944095700022923
steps: 218	lr: 4e-05	evaluation reward: 1.99	
episode: 1111	score: 2.0	memory length: 206856	epsilon: 0.8942115700022966
steps: 200	lr: 4e-05	evaluation reward: 2.01	
episode: 1112	score: 4.0	memory length: 207152	epsilon: 0.8939185300023029
steps: 296	lr: 4e-05	evaluation reward: 2.02	
episode: 1113	score: 1.0	memory length: 207302	epsilon: 0.8937700300023061
steps: 150	lr: 4e-05	evaluation reward: 2.03	
episode: 1114	score: 3.0	memory length: 207533	epsilon: 0.8935413400023111
steps: 231	lr: 4e-05	evaluation reward: 2.04	
episode: 1115	score: 1.0	memory length: 207703	epsilon: 0.8933730400023148
steps: 170	lr: 4e-05	evaluation reward: 2.0	
episode: 1116	score: 0.0	memory length: 207825	epsilon: 0.8932522600023174
steps: 122	lr: 4e-05	evaluation reward: 1.97	
episode: 1117	score: 2.0	memory length: 208043	epsilon: 0.8930364400023221
steps: 218	lr: 4e-05	evaluation reward: 1.98	
episode: 1118	score: 4.0	memory length: 208342	epsilon: 0.8927404300023285

steps: 299	lr: 4e-05	evaluation reward: 2.02	
episode: 1119	score: 2.0	memory length: 208559	epsilon: 0.8925256000023332
steps: 217	lr: 4e-05	evaluation reward: 2.03	
episode: 1120	score: 4.0	memory length: 208817	epsilon: 0.8922701800023387
steps: 258	lr: 4e-05	evaluation reward: 2.07	
episode: 1121	score: 2.0	memory length: 209039	epsilon: 0.8920504000023435
steps: 222	lr: 4e-05	evaluation reward: 2.08	
episode: 1122	score: 3.0	memory length: 209302	epsilon: 0.8917900300023491
steps: 263	lr: 4e-05	evaluation reward: 2.08	
episode: 1123	score: 3.0	memory length: 209551	epsilon: 0.8915435200023545
steps: 249	lr: 4e-05	evaluation reward: 2.09	
episode: 1124	score: 0.0	memory length: 209673	epsilon: 0.8914227400023571
steps: 122	lr: 4e-05	evaluation reward: 2.08	
episode: 1125	score: 0.0	memory length: 209796	epsilon: 0.8913009700023597
steps: 123	lr: 4e-05	evaluation reward: 2.07	
episode: 1126	score: 2.0	memory length: 210013	epsilon: 0.8910861400023644
steps: 217	lr: 4e-05	evaluation reward: 2.06	
episode: 1127	score: 2.0	memory length: 210229	epsilon: 0.890872300002369
steps: 216	lr: 4e-05	evaluation reward: 2.06	
episode: 1128	score: 2.0	memory length: 210427	epsilon: 0.8906762800023733
steps: 198	lr: 4e-05	evaluation reward: 2.05	
episode: 1129	score: 0.0	memory length: 210550	epsilon: 0.890554510002376
steps: 123	lr: 4e-05	evaluation reward: 2.03	
episode: 1130	score: 1.0	memory length: 210719	epsilon: 0.8903872000023796
steps: 169	lr: 4e-05	evaluation reward: 2.01	
episode: 1131	score: 3.0	memory length: 210969	epsilon: 0.890139700002385
steps: 250	lr: 4e-05	evaluation reward: 2.03	
episode: 1132	score: 3.0	memory length: 211197	epsilon: 0.8899139800023899
steps: 228	lr: 4e-05	evaluation reward: 2.05	
episode: 1133	score: 2.0	memory length: 211414	epsilon: 0.8896991500023945
steps: 217	lr: 4e-05	evaluation reward: 2.06	
episode: 1134	score: 1.0	memory length: 211565	epsilon: 0.8895496600023978
steps: 151	lr: 4e-05	evaluation reward: 2.06	
episode: 1135	score: 3.0	memory length: 211830	epsilon: 0.8892873100024035
steps: 265	lr: 4e-05	evaluation reward: 2.08	
episode: 1136	score: 0.0	memory length: 211953	epsilon: 0.8891655400024061
steps: 123	lr: 4e-05	evaluation reward: 2.07	
episode: 1137	score: 1.0	memory length: 212122	epsilon: 0.8889982300024097
steps: 169	lr: 4e-05	evaluation reward: 2.06	
episode: 1138	score: 2.0	memory length: 212323	epsilon: 0.8887992400024141
steps: 201	lr: 4e-05	evaluation reward: 2.07	
episode: 1139	score: 0.0	memory length: 212446	epsilon: 0.8886774700024167
steps: 123	lr: 4e-05	evaluation reward: 2.05	
episode: 1140	score: 2.0	memory length: 212649	epsilon: 0.8884765000024211
steps: 203	lr: 4e-05	evaluation reward: 2.05	
episode: 1141	score: 2.0	memory length: 212831	epsilon: 0.888296320002425
steps: 182	lr: 4e-05	evaluation reward: 2.05	
episode: 1142	score: 4.0	memory length: 213108	epsilon: 0.8880220900024309

steps: 277	lr: 4e-05	evaluation reward: 2.09	
episode: 1143	score: 0.0	memory length: 213230	epsilon: 0.8879013100024336
steps: 122	lr: 4e-05	evaluation reward: 2.09	
episode: 1144	score: 1.0	memory length: 213381	epsilon: 0.8877518200024368
steps: 151	lr: 4e-05	evaluation reward: 2.09	
episode: 1145	score: 3.0	memory length: 213630	epsilon: 0.8875053100024421
steps: 249	lr: 4e-05	evaluation reward: 2.12	
episode: 1146	score: 4.0	memory length: 213924	epsilon: 0.8872142500024485
steps: 294	lr: 4e-05	evaluation reward: 2.13	
episode: 1147	score: 3.0	memory length: 214153	epsilon: 0.8869875400024534
steps: 229	lr: 4e-05	evaluation reward: 2.13	
episode: 1148	score: 0.0	memory length: 214276	epsilon: 0.886865770002456
steps: 123	lr: 4e-05	evaluation reward: 2.1	
episode: 1149	score: 3.0	memory length: 214509	epsilon: 0.886635100002461
steps: 233	lr: 4e-05	evaluation reward: 2.12	
episode: 1150	score: 1.0	memory length: 214680	epsilon: 0.8864658100024647
steps: 171	lr: 4e-05	evaluation reward: 2.13	
episode: 1151	score: 0.0	memory length: 214802	epsilon: 0.8863450300024673
steps: 122	lr: 4e-05	evaluation reward: 2.11	
episode: 1152	score: 1.0	memory length: 214953	epsilon: 0.8861955400024706
steps: 151	lr: 4e-05	evaluation reward: 2.12	
episode: 1153	score: 3.0	memory length: 215202	epsilon: 0.8859490300024759
steps: 249	lr: 4e-05	evaluation reward: 2.12	
episode: 1154	score: 6.0	memory length: 215539	epsilon: 0.8856154000024832
steps: 337	lr: 4e-05	evaluation reward: 2.18	
episode: 1155	score: 2.0	memory length: 215759	epsilon: 0.8853976000024879
steps: 220	lr: 4e-05	evaluation reward: 2.2	
episode: 1156	score: 0.0	memory length: 215881	epsilon: 0.8852768200024905
steps: 122	lr: 4e-05	evaluation reward: 2.2	
episode: 1157	score: 5.0	memory length: 216224	epsilon: 0.8849372500024979
steps: 343	lr: 4e-05	evaluation reward: 2.22	
episode: 1158	score: 0.0	memory length: 216347	epsilon: 0.8848154800025005
steps: 123	lr: 4e-05	evaluation reward: 2.18	
episode: 1159	score: 3.0	memory length: 216614	epsilon: 0.8845511500025063
steps: 267	lr: 4e-05	evaluation reward: 2.21	
episode: 1160	score: 4.0	memory length: 216888	epsilon: 0.8842798900025122
steps: 274	lr: 4e-05	evaluation reward: 2.23	
episode: 1161	score: 2.0	memory length: 217088	epsilon: 0.8840818900025165
steps: 200	lr: 4e-05	evaluation reward: 2.22	
episode: 1162	score: 1.0	memory length: 217257	epsilon: 0.8839145800025201
steps: 169	lr: 4e-05	evaluation reward: 2.17	
episode: 1163	score: 4.0	memory length: 217571	epsilon: 0.8836037200025268
steps: 314	lr: 4e-05	evaluation reward: 2.19	
episode: 1164	score: 1.0	memory length: 217741	epsilon: 0.8834354200025305
steps: 170	lr: 4e-05	evaluation reward: 2.2	
episode: 1165	score: 2.0	memory length: 217960	epsilon: 0.8832186100025352
steps: 219	lr: 4e-05	evaluation reward: 2.19	
episode: 1166	score: 3.0	memory length: 218186	epsilon: 0.8829948700025401

steps: 226	lr: 4e-05	evaluation reward: 2.17	
episode: 1167	score: 5.0	memory length: 218528	epsilon: 0.8826562900025474
steps: 342	lr: 4e-05	evaluation reward: 2.18	
episode: 1168	score: 2.0	memory length: 218731	epsilon: 0.8824553200025518
steps: 203	lr: 4e-05	evaluation reward: 2.14	
episode: 1169	score: 0.0	memory length: 218854	epsilon: 0.8823335500025544
steps: 123	lr: 4e-05	evaluation reward: 2.13	
episode: 1170	score: 2.0	memory length: 219051	epsilon: 0.8821385200025587
steps: 197	lr: 4e-05	evaluation reward: 2.13	
episode: 1171	score: 2.0	memory length: 219267	epsilon: 0.8819246800025633
steps: 216	lr: 4e-05	evaluation reward: 2.11	
episode: 1172	score: 7.0	memory length: 219663	epsilon: 0.8815326400025718
steps: 396	lr: 4e-05	evaluation reward: 2.16	
episode: 1173	score: 3.0	memory length: 219907	epsilon: 0.881291080002577
steps: 244	lr: 4e-05	evaluation reward: 2.16	
episode: 1174	score: 5.0	memory length: 220246	epsilon: 0.8809554700025843
steps: 339	lr: 4e-05	evaluation reward: 2.18	
episode: 1175	score: 4.0	memory length: 220538	epsilon: 0.8806663900025906
steps: 292	lr: 4e-05	evaluation reward: 2.2	
episode: 1176	score: 3.0	memory length: 220782	epsilon: 0.8804248300025959
steps: 244	lr: 4e-05	evaluation reward: 2.2	
episode: 1177	score: 0.0	memory length: 220905	epsilon: 0.8803030600025985
steps: 123	lr: 4e-05	evaluation reward: 2.2	
episode: 1178	score: 4.0	memory length: 221200	epsilon: 0.8800110100026048
steps: 295	lr: 4e-05	evaluation reward: 2.24	
episode: 1179	score: 2.0	memory length: 221398	epsilon: 0.8798149900026091
steps: 198	lr: 4e-05	evaluation reward: 2.23	
episode: 1180	score: 2.0	memory length: 221579	epsilon: 0.879635800002613
steps: 181	lr: 4e-05	evaluation reward: 2.23	
episode: 1181	score: 2.0	memory length: 221776	epsilon: 0.8794407700026172
steps: 197	lr: 4e-05	evaluation reward: 2.24	
episode: 1182	score: 2.0	memory length: 221974	epsilon: 0.8792447500026215
steps: 198	lr: 4e-05	evaluation reward: 2.25	
episode: 1183	score: 1.0	memory length: 222126	epsilon: 0.8790942700026247
steps: 152	lr: 4e-05	evaluation reward: 2.23	
episode: 1184	score: 1.0	memory length: 222295	epsilon: 0.8789269600026284
steps: 169	lr: 4e-05	evaluation reward: 2.23	
episode: 1185	score: 3.0	memory length: 222522	epsilon: 0.8787022300026333
steps: 227	lr: 4e-05	evaluation reward: 2.24	
episode: 1186	score: 2.0	memory length: 222719	epsilon: 0.8785072000026375
steps: 197	lr: 4e-05	evaluation reward: 2.23	
episode: 1187	score: 2.0	memory length: 222901	epsilon: 0.8783270200026414
steps: 182	lr: 4e-05	evaluation reward: 2.23	
episode: 1188	score: 2.0	memory length: 223099	epsilon: 0.8781310000026457
steps: 198	lr: 4e-05	evaluation reward: 2.18	
episode: 1189	score: 3.0	memory length: 223348	epsilon: 0.877884490002651
steps: 249	lr: 4e-05	evaluation reward: 2.2	
episode: 1190	score: 1.0	memory length: 223499	epsilon: 0.8777350000026543

```

steps: 151    lr: 4e-05    evaluation reward: 2.2
episode: 1191  score: 10.0  memory length: 223899  epsilon:
0.8773390000026628  steps: 400    lr: 4e-05    evaluation reward: 2.26
episode: 1192  score: 0.0  memory length: 224022  epsilon: 0.8772172300026655
steps: 123    lr: 4e-05    evaluation reward: 2.22
episode: 1193  score: 1.0  memory length: 224173  epsilon: 0.8770677400026687
steps: 151    lr: 4e-05    evaluation reward: 2.23
episode: 1194  score: 4.0  memory length: 224482  epsilon: 0.8767618300026754
steps: 309    lr: 4e-05    evaluation reward: 2.24
episode: 1195  score: 0.0  memory length: 224605  epsilon: 0.876640060002678
steps: 123    lr: 4e-05    evaluation reward: 2.21
episode: 1196  score: 2.0  memory length: 224826  epsilon: 0.8764212700026828
steps: 221    lr: 4e-05    evaluation reward: 2.19
episode: 1197  score: 2.0  memory length: 225028  epsilon: 0.8762212900026871
steps: 202    lr: 4e-05    evaluation reward: 2.2
episode: 1198  score: 1.0  memory length: 225179  epsilon: 0.8760718000026904
steps: 151    lr: 4e-05    evaluation reward: 2.21
episode: 1199  score: 0.0  memory length: 225301  epsilon: 0.875951020002693
steps: 122    lr: 4e-05    evaluation reward: 2.2
episode: 1200  score: 1.0  memory length: 225452  epsilon: 0.8758015300026962
steps: 151    lr: 4e-05    evaluation reward: 2.18
episode: 1201  score: 0.0  memory length: 225574  epsilon: 0.8756807500026988
steps: 122    lr: 4e-05    evaluation reward: 2.17
episode: 1202  score: 3.0  memory length: 225820  epsilon: 0.8754372100027041
steps: 246    lr: 4e-05    evaluation reward: 2.16
episode: 1203  score: 3.0  memory length: 226030  epsilon: 0.8752293100027086
steps: 210    lr: 4e-05    evaluation reward: 2.17
episode: 1204  score: 4.0  memory length: 226305  epsilon: 0.8749570600027146
steps: 275    lr: 4e-05    evaluation reward: 2.17
episode: 1205  score: 0.0  memory length: 226428  epsilon: 0.8748352900027172
steps: 123    lr: 4e-05    evaluation reward: 2.16
episode: 1206  score: 2.0  memory length: 226647  epsilon: 0.8746184800027219
steps: 219    lr: 4e-05    evaluation reward: 2.16
episode: 1207  score: 1.0  memory length: 226816  epsilon: 0.8744511700027255
steps: 169    lr: 4e-05    evaluation reward: 2.16
episode: 1208  score: 3.0  memory length: 227062  epsilon: 0.8742076300027308
steps: 246    lr: 4e-05    evaluation reward: 2.17
episode: 1209  score: 3.0  memory length: 227271  epsilon: 0.8740007200027353
steps: 209    lr: 4e-05    evaluation reward: 2.15
episode: 1210  score: 7.0  memory length: 227693  epsilon: 0.8735829400027444
steps: 422    lr: 4e-05    evaluation reward: 2.2
episode: 1211  score: 3.0  memory length: 227940  epsilon: 0.8733384100027497
steps: 247    lr: 4e-05    evaluation reward: 2.21
episode: 1212  score: 11.0  memory length: 228436  epsilon:
0.8728473700027604  steps: 496    lr: 4e-05    evaluation reward: 2.28
episode: 1213  score: 1.0  memory length: 228605  epsilon: 0.872680060002764
steps: 169    lr: 4e-05    evaluation reward: 2.28
episode: 1214  score: 2.0  memory length: 228805  epsilon: 0.8724820600027683

```

steps: 200	lr: 4e-05	evaluation reward: 2.27	
episode: 1215	score: 2.0	memory length: 229003	epsilon: 0.8722860400027725
steps: 198	lr: 4e-05	evaluation reward: 2.28	
episode: 1216	score: 0.0	memory length: 229126	epsilon: 0.8721642700027752
steps: 123	lr: 4e-05	evaluation reward: 2.28	
episode: 1217	score: 0.0	memory length: 229248	epsilon: 0.8720434900027778
steps: 122	lr: 4e-05	evaluation reward: 2.26	
episode: 1218	score: 0.0	memory length: 229371	epsilon: 0.8719217200027805
steps: 123	lr: 4e-05	evaluation reward: 2.22	
episode: 1219	score: 3.0	memory length: 229620	epsilon: 0.8716752100027858
steps: 249	lr: 4e-05	evaluation reward: 2.23	
episode: 1220	score: 3.0	memory length: 229871	epsilon: 0.8714267200027912
steps: 251	lr: 4e-05	evaluation reward: 2.22	
episode: 1221	score: 4.0	memory length: 230148	epsilon: 0.8711524900027972
steps: 277	lr: 4e-05	evaluation reward: 2.24	
episode: 1222	score: 2.0	memory length: 230327	epsilon: 0.870975280002801
steps: 179	lr: 4e-05	evaluation reward: 2.23	
episode: 1223	score: 5.0	memory length: 230637	epsilon: 0.8706683800028077
steps: 310	lr: 4e-05	evaluation reward: 2.25	
episode: 1224	score: 9.0	memory length: 230997	epsilon: 0.8703119800028154
steps: 360	lr: 4e-05	evaluation reward: 2.34	
episode: 1225	score: 5.0	memory length: 231310	epsilon: 0.8700021100028221
steps: 313	lr: 4e-05	evaluation reward: 2.39	
episode: 1226	score: 4.0	memory length: 231585	epsilon: 0.869729860002828
steps: 275	lr: 4e-05	evaluation reward: 2.41	
episode: 1227	score: 1.0	memory length: 231755	epsilon: 0.8695615600028317
steps: 170	lr: 4e-05	evaluation reward: 2.4	
episode: 1228	score: 2.0	memory length: 231937	epsilon: 0.8693813800028356
steps: 182	lr: 4e-05	evaluation reward: 2.4	
episode: 1229	score: 5.0	memory length: 232241	epsilon: 0.8690804200028421
steps: 304	lr: 4e-05	evaluation reward: 2.45	
episode: 1230	score: 3.0	memory length: 232487	epsilon: 0.8688368800028474
steps: 246	lr: 4e-05	evaluation reward: 2.47	
episode: 1231	score: 1.0	memory length: 232656	epsilon: 0.868669570002851
steps: 169	lr: 4e-05	evaluation reward: 2.45	
episode: 1232	score: 7.0	memory length: 233073	epsilon: 0.86825674000286
steps: 417	lr: 4e-05	evaluation reward: 2.49	
episode: 1233	score: 0.0	memory length: 233196	epsilon: 0.8681349700028627
steps: 123	lr: 4e-05	evaluation reward: 2.47	
episode: 1234	score: 3.0	memory length: 233422	epsilon: 0.8679112300028675
steps: 226	lr: 4e-05	evaluation reward: 2.49	
episode: 1235	score: 4.0	memory length: 233697	epsilon: 0.8676389800028734
steps: 275	lr: 4e-05	evaluation reward: 2.5	
episode: 1236	score: 5.0	memory length: 233986	epsilon: 0.8673528700028796
steps: 289	lr: 4e-05	evaluation reward: 2.55	
episode: 1237	score: 2.0	memory length: 234184	epsilon: 0.8671568500028839
steps: 198	lr: 4e-05	evaluation reward: 2.56	
episode: 1238	score: 2.0	memory length: 234403	epsilon: 0.8669400400028886

steps: 219	lr: 4e-05	evaluation reward: 2.56	
episode: 1239	score: 2.0	memory length: 234583	epsilon: 0.8667618400028925
steps: 180	lr: 4e-05	evaluation reward: 2.58	
episode: 1240	score: 2.0	memory length: 234781	epsilon: 0.8665658200028967
steps: 198	lr: 4e-05	evaluation reward: 2.58	
episode: 1241	score: 1.0	memory length: 234931	epsilon: 0.8664173200029
steps: 150	lr: 4e-05	evaluation reward: 2.57	
episode: 1242	score: 1.0	memory length: 235100	epsilon: 0.8662500100029036
steps: 169	lr: 4e-05	evaluation reward: 2.54	
episode: 1243	score: 6.0	memory length: 235455	epsilon: 0.8658985600029112
steps: 355	lr: 4e-05	evaluation reward: 2.6	
episode: 1244	score: 0.0	memory length: 235577	epsilon: 0.8657777800029138
steps: 122	lr: 4e-05	evaluation reward: 2.59	
episode: 1245	score: 5.0	memory length: 235944	epsilon: 0.8654144500029217
steps: 367	lr: 4e-05	evaluation reward: 2.61	
episode: 1246	score: 3.0	memory length: 236215	epsilon: 0.8651461600029275
steps: 271	lr: 4e-05	evaluation reward: 2.6	
episode: 1247	score: 1.0	memory length: 236386	epsilon: 0.8649768700029312
steps: 171	lr: 4e-05	evaluation reward: 2.58	
episode: 1248	score: 2.0	memory length: 236583	epsilon: 0.8647818400029355
steps: 197	lr: 4e-05	evaluation reward: 2.6	
episode: 1249	score: 5.0	memory length: 236947	epsilon: 0.8644214800029433
steps: 364	lr: 4e-05	evaluation reward: 2.62	
episode: 1250	score: 4.0	memory length: 237244	epsilon: 0.8641274500029497
steps: 297	lr: 4e-05	evaluation reward: 2.65	
episode: 1251	score: 5.0	memory length: 237568	epsilon: 0.8638066900029566
steps: 324	lr: 4e-05	evaluation reward: 2.7	
episode: 1252	score: 3.0	memory length: 237797	epsilon: 0.8635799800029615
steps: 229	lr: 4e-05	evaluation reward: 2.72	
episode: 1253	score: 4.0	memory length: 238087	epsilon: 0.8632928800029678
steps: 290	lr: 4e-05	evaluation reward: 2.73	
episode: 1254	score: 0.0	memory length: 238209	epsilon: 0.8631721000029704
steps: 122	lr: 4e-05	evaluation reward: 2.67	
episode: 1255	score: 3.0	memory length: 238440	epsilon: 0.8629434100029754
steps: 231	lr: 4e-05	evaluation reward: 2.68	
episode: 1256	score: 2.0	memory length: 238638	epsilon: 0.8627473900029796
steps: 198	lr: 4e-05	evaluation reward: 2.7	
episode: 1257	score: 3.0	memory length: 238886	epsilon: 0.862501870002985
steps: 248	lr: 4e-05	evaluation reward: 2.68	
episode: 1258	score: 1.0	memory length: 239056	epsilon: 0.8623335700029886
steps: 170	lr: 4e-05	evaluation reward: 2.69	
episode: 1259	score: 5.0	memory length: 239382	epsilon: 0.8620108300029956
steps: 326	lr: 4e-05	evaluation reward: 2.71	
episode: 1260	score: 4.0	memory length: 239659	epsilon: 0.8617366000030016
steps: 277	lr: 4e-05	evaluation reward: 2.71	
episode: 1261	score: 5.0	memory length: 239986	epsilon: 0.8614128700030086
steps: 327	lr: 4e-05	evaluation reward: 2.74	
episode: 1262	score: 3.0	memory length: 240213	epsilon: 0.8611881400030135

steps: 227	lr: 4e-05	evaluation reward: 2.76	
episode: 1263	score: 1.0	memory length: 240363	epsilon: 0.8610396400030167
steps: 150	lr: 4e-05	evaluation reward: 2.73	
episode: 1264	score: 2.0	memory length: 240560	epsilon: 0.8608446100030209
steps: 197	lr: 4e-05	evaluation reward: 2.74	
episode: 1265	score: 5.0	memory length: 240885	epsilon: 0.8605228600030279
steps: 325	lr: 4e-05	evaluation reward: 2.77	
episode: 1266	score: 0.0	memory length: 241008	epsilon: 0.8604010900030306
steps: 123	lr: 4e-05	evaluation reward: 2.74	
episode: 1267	score: 0.0	memory length: 241131	epsilon: 0.8602793200030332
steps: 123	lr: 4e-05	evaluation reward: 2.69	
episode: 1268	score: 2.0	memory length: 241349	epsilon: 0.8600635000030379
steps: 218	lr: 4e-05	evaluation reward: 2.69	
episode: 1269	score: 0.0	memory length: 241471	epsilon: 0.8599427200030405
steps: 122	lr: 4e-05	evaluation reward: 2.69	
episode: 1270	score: 4.0	memory length: 241770	epsilon: 0.8596467100030469
steps: 299	lr: 4e-05	evaluation reward: 2.71	
episode: 1271	score: 3.0	memory length: 241997	epsilon: 0.8594219800030518
steps: 227	lr: 4e-05	evaluation reward: 2.72	
episode: 1272	score: 1.0	memory length: 242165	epsilon: 0.8592556600030554
steps: 168	lr: 4e-05	evaluation reward: 2.66	
episode: 1273	score: 3.0	memory length: 242410	epsilon: 0.8590131100030607
steps: 245	lr: 4e-05	evaluation reward: 2.66	
episode: 1274	score: 2.0	memory length: 242611	epsilon: 0.858814120003065
steps: 201	lr: 4e-05	evaluation reward: 2.63	
episode: 1275	score: 1.0	memory length: 242783	epsilon: 0.8586438400030687
steps: 172	lr: 4e-05	evaluation reward: 2.6	
episode: 1276	score: 2.0	memory length: 242985	epsilon: 0.858443860003073
steps: 202	lr: 4e-05	evaluation reward: 2.59	
episode: 1277	score: 5.0	memory length: 243279	epsilon: 0.8581528000030794
steps: 294	lr: 4e-05	evaluation reward: 2.64	
episode: 1278	score: 4.0	memory length: 243600	epsilon: 0.8578350100030863
steps: 321	lr: 4e-05	evaluation reward: 2.64	
episode: 1279	score: 1.0	memory length: 243751	epsilon: 0.8576855200030895
steps: 151	lr: 4e-05	evaluation reward: 2.63	
episode: 1280	score: 4.0	memory length: 244071	epsilon: 0.8573687200030964
steps: 320	lr: 4e-05	evaluation reward: 2.65	
episode: 1281	score: 0.0	memory length: 244194	epsilon: 0.857246950003099
steps: 123	lr: 4e-05	evaluation reward: 2.63	
episode: 1282	score: 0.0	memory length: 244317	epsilon: 0.8571251800031017
steps: 123	lr: 4e-05	evaluation reward: 2.61	
episode: 1283	score: 4.0	memory length: 244576	epsilon: 0.8568687700031072
steps: 259	lr: 4e-05	evaluation reward: 2.64	
episode: 1284	score: 3.0	memory length: 244825	epsilon: 0.8566222600031126
steps: 249	lr: 4e-05	evaluation reward: 2.66	
episode: 1285	score: 3.0	memory length: 245069	epsilon: 0.8563807000031178
steps: 244	lr: 4e-05	evaluation reward: 2.66	
episode: 1286	score: 2.0	memory length: 245267	epsilon: 0.8561846800031221

steps: 198	lr: 4e-05	evaluation reward: 2.66	
episode: 1287	score: 5.0	memory length: 245588	epsilon: 0.855866890003129
steps: 321	lr: 4e-05	evaluation reward: 2.69	
episode: 1288	score: 4.0	memory length: 245879	epsilon: 0.8555788000031352
steps: 291	lr: 4e-05	evaluation reward: 2.71	
episode: 1289	score: 5.0	memory length: 246202	epsilon: 0.8552590300031422
steps: 323	lr: 4e-05	evaluation reward: 2.73	
episode: 1290	score: 2.0	memory length: 246399	epsilon: 0.8550640000031464
steps: 197	lr: 4e-05	evaluation reward: 2.74	
episode: 1291	score: 3.0	memory length: 246667	epsilon: 0.8547986800031522
steps: 268	lr: 4e-05	evaluation reward: 2.67	
episode: 1292	score: 2.0	memory length: 246865	epsilon: 0.8546026600031564
steps: 198	lr: 4e-05	evaluation reward: 2.69	
episode: 1293	score: 3.0	memory length: 247113	epsilon: 0.8543571400031618
steps: 248	lr: 4e-05	evaluation reward: 2.71	
episode: 1294	score: 3.0	memory length: 247339	epsilon: 0.8541334000031666
steps: 226	lr: 4e-05	evaluation reward: 2.7	
episode: 1295	score: 2.0	memory length: 247536	epsilon: 0.8539383700031709
steps: 197	lr: 4e-05	evaluation reward: 2.72	
episode: 1296	score: 5.0	memory length: 247886	epsilon: 0.8535918700031784
steps: 350	lr: 4e-05	evaluation reward: 2.75	
episode: 1297	score: 5.0	memory length: 248208	epsilon: 0.8532730900031853
steps: 322	lr: 4e-05	evaluation reward: 2.78	
episode: 1298	score: 0.0	memory length: 248330	epsilon: 0.8531523100031879
steps: 122	lr: 4e-05	evaluation reward: 2.77	
episode: 1299	score: 5.0	memory length: 248654	epsilon: 0.8528315500031949
steps: 324	lr: 4e-05	evaluation reward: 2.82	
episode: 1300	score: 4.0	memory length: 248969	epsilon: 0.8525197000032017
steps: 315	lr: 4e-05	evaluation reward: 2.85	
episode: 1301	score: 1.0	memory length: 249120	epsilon: 0.8523702100032049
steps: 151	lr: 4e-05	evaluation reward: 2.86	
episode: 1302	score: 3.0	memory length: 249367	epsilon: 0.8521256800032102
steps: 247	lr: 4e-05	evaluation reward: 2.86	
episode: 1303	score: 3.0	memory length: 249613	epsilon: 0.8518821400032155
steps: 246	lr: 4e-05	evaluation reward: 2.86	
episode: 1304	score: 6.0	memory length: 250004	epsilon: 0.8514950500032239
steps: 391	lr: 4e-05	evaluation reward: 2.88	
episode: 1305	score: 4.0	memory length: 250280	epsilon: 0.8512218100032298
steps: 276	lr: 4e-05	evaluation reward: 2.92	
episode: 1306	score: 1.0	memory length: 250430	epsilon: 0.851073310003233
steps: 150	lr: 4e-05	evaluation reward: 2.91	
episode: 1307	score: 3.0	memory length: 250661	epsilon: 0.850844620003238
steps: 231	lr: 4e-05	evaluation reward: 2.93	
episode: 1308	score: 3.0	memory length: 250888	epsilon: 0.8506198900032429
steps: 227	lr: 4e-05	evaluation reward: 2.93	
episode: 1309	score: 1.0	memory length: 251039	epsilon: 0.8504704000032461
steps: 151	lr: 4e-05	evaluation reward: 2.91	
episode: 1310	score: 2.0	memory length: 251239	epsilon: 0.8502724000032504

steps: 200	lr: 4e-05	evaluation reward: 2.86	
episode: 1311	score: 2.0	memory length: 251437	epsilon: 0.8500763800032547
steps: 198	lr: 4e-05	evaluation reward: 2.85	
episode: 1312	score: 3.0	memory length: 251663	epsilon: 0.8498526400032596
steps: 226	lr: 4e-05	evaluation reward: 2.77	
episode: 1313	score: 1.0	memory length: 251834	epsilon: 0.8496833500032632
steps: 171	lr: 4e-05	evaluation reward: 2.77	
episode: 1314	score: 2.0	memory length: 252034	epsilon: 0.8494853500032675
steps: 200	lr: 4e-05	evaluation reward: 2.77	
episode: 1315	score: 7.0	memory length: 252446	epsilon: 0.8490774700032764
steps: 412	lr: 4e-05	evaluation reward: 2.82	
episode: 1316	score: 1.0	memory length: 252597	epsilon: 0.8489279800032796
steps: 151	lr: 4e-05	evaluation reward: 2.83	
episode: 1317	score: 1.0	memory length: 252767	epsilon: 0.8487596800032833
steps: 170	lr: 4e-05	evaluation reward: 2.84	
episode: 1318	score: 3.0	memory length: 252998	epsilon: 0.8485309900032882
steps: 231	lr: 4e-05	evaluation reward: 2.87	
episode: 1319	score: 2.0	memory length: 253196	epsilon: 0.8483349700032925
steps: 198	lr: 4e-05	evaluation reward: 2.86	
episode: 1320	score: 3.0	memory length: 253441	epsilon: 0.8480924200032978
steps: 245	lr: 4e-05	evaluation reward: 2.86	
episode: 1321	score: 2.0	memory length: 253623	epsilon: 0.8479122400033017
steps: 182	lr: 4e-05	evaluation reward: 2.84	
episode: 1322	score: 12.0	memory length: 254119	epsilon:
0.8474212000033123	steps: 496	lr: 4e-05	evaluation reward: 2.94
episode: 1323	score: 3.0	memory length: 254386	epsilon: 0.8471568700033181
steps: 267	lr: 4e-05	evaluation reward: 2.92	
episode: 1324	score: 8.0	memory length: 254809	epsilon: 0.8467381000033272
steps: 423	lr: 4e-05	evaluation reward: 2.91	
episode: 1325	score: 3.0	memory length: 255053	epsilon: 0.8464965400033324
steps: 244	lr: 4e-05	evaluation reward: 2.89	
episode: 1326	score: 2.0	memory length: 255237	epsilon: 0.8463143800033364
steps: 184	lr: 4e-05	evaluation reward: 2.87	
episode: 1327	score: 4.0	memory length: 255511	epsilon: 0.8460431200033423
steps: 274	lr: 4e-05	evaluation reward: 2.9	
episode: 1328	score: 1.0	memory length: 255664	epsilon: 0.8458916500033455
steps: 153	lr: 4e-05	evaluation reward: 2.89	
episode: 1329	score: 1.0	memory length: 255833	epsilon: 0.8457243400033492
steps: 169	lr: 4e-05	evaluation reward: 2.85	
episode: 1330	score: 3.0	memory length: 256060	epsilon: 0.845499610003354
steps: 227	lr: 4e-05	evaluation reward: 2.85	
episode: 1331	score: 6.0	memory length: 256418	epsilon: 0.8451451900033617
steps: 358	lr: 4e-05	evaluation reward: 2.9	
episode: 1332	score: 4.0	memory length: 256676	epsilon: 0.8448897700033673
steps: 258	lr: 4e-05	evaluation reward: 2.87	
episode: 1333	score: 5.0	memory length: 257001	epsilon: 0.8445680200033743
steps: 325	lr: 4e-05	evaluation reward: 2.92	
episode: 1334	score: 4.0	memory length: 257316	epsilon: 0.844256170003381

steps: 315	lr: 4e-05	evaluation reward: 2.93	
episode: 1335	score: 2.0	memory length: 257535	epsilon: 0.8440393600033858
steps: 219	lr: 4e-05	evaluation reward: 2.91	
episode: 1336	score: 0.0	memory length: 257657	epsilon: 0.8439185800033884
steps: 122	lr: 4e-05	evaluation reward: 2.86	
episode: 1337	score: 1.0	memory length: 257807	epsilon: 0.8437700800033916
steps: 150	lr: 4e-05	evaluation reward: 2.85	
episode: 1338	score: 1.0	memory length: 257958	epsilon: 0.8436205900033948
steps: 151	lr: 4e-05	evaluation reward: 2.84	
episode: 1339	score: 5.0	memory length: 258283	epsilon: 0.8432988400034018
steps: 325	lr: 4e-05	evaluation reward: 2.87	
episode: 1340	score: 6.0	memory length: 258617	epsilon: 0.842968180003409
steps: 334	lr: 4e-05	evaluation reward: 2.91	
episode: 1341	score: 2.0	memory length: 258797	epsilon: 0.8427899800034129
steps: 180	lr: 4e-05	evaluation reward: 2.92	
episode: 1342	score: 0.0	memory length: 258920	epsilon: 0.8426682100034155
steps: 123	lr: 4e-05	evaluation reward: 2.91	
episode: 1343	score: 0.0	memory length: 259043	epsilon: 0.8425464400034182
steps: 123	lr: 4e-05	evaluation reward: 2.85	
episode: 1344	score: 0.0	memory length: 259165	epsilon: 0.8424256600034208
steps: 122	lr: 4e-05	evaluation reward: 2.85	
episode: 1345	score: 4.0	memory length: 259458	epsilon: 0.8421355900034271
steps: 293	lr: 4e-05	evaluation reward: 2.84	
episode: 1346	score: 4.0	memory length: 259733	epsilon: 0.841863340003433
steps: 275	lr: 4e-05	evaluation reward: 2.85	
episode: 1347	score: 4.0	memory length: 259991	epsilon: 0.8416079200034385
steps: 258	lr: 4e-05	evaluation reward: 2.88	
episode: 1348	score: 0.0	memory length: 260113	epsilon: 0.8414871400034412
steps: 122	lr: 4e-05	evaluation reward: 2.86	
episode: 1349	score: 4.0	memory length: 260389	epsilon: 0.8412139000034471
steps: 276	lr: 4e-05	evaluation reward: 2.85	
episode: 1350	score: 6.0	memory length: 260747	epsilon: 0.8408594800034548
steps: 358	lr: 4e-05	evaluation reward: 2.87	
episode: 1351	score: 3.0	memory length: 260972	epsilon: 0.8406367300034596
steps: 225	lr: 4e-05	evaluation reward: 2.85	
episode: 1352	score: 5.0	memory length: 261303	epsilon: 0.8403090400034667
steps: 331	lr: 4e-05	evaluation reward: 2.87	
episode: 1353	score: 0.0	memory length: 261425	epsilon: 0.8401882600034694
steps: 122	lr: 4e-05	evaluation reward: 2.83	
episode: 1354	score: 3.0	memory length: 261689	epsilon: 0.839926900003475
steps: 264	lr: 4e-05	evaluation reward: 2.86	
episode: 1355	score: 6.0	memory length: 262033	epsilon: 0.8395863400034824
steps: 344	lr: 4e-05	evaluation reward: 2.89	
episode: 1356	score: 3.0	memory length: 262280	epsilon: 0.8393418100034877
steps: 247	lr: 4e-05	evaluation reward: 2.9	
episode: 1357	score: 2.0	memory length: 262497	epsilon: 0.8391269800034924
steps: 217	lr: 4e-05	evaluation reward: 2.89	
episode: 1358	score: 0.0	memory length: 262619	epsilon: 0.839006200003495

steps: 122	lr: 4e-05	evaluation reward: 2.88	
episode: 1359	score: 2.0	memory length: 262816	epsilon: 0.8388111700034993
steps: 197	lr: 4e-05	evaluation reward: 2.85	
episode: 1360	score: 2.0	memory length: 263033	epsilon: 0.8385963400035039
steps: 217	lr: 4e-05	evaluation reward: 2.83	
episode: 1361	score: 0.0	memory length: 263156	epsilon: 0.8384745700035066
steps: 123	lr: 4e-05	evaluation reward: 2.78	
episode: 1362	score: 5.0	memory length: 263480	epsilon: 0.8381538100035135
steps: 324	lr: 4e-05	evaluation reward: 2.8	
episode: 1363	score: 4.0	memory length: 263735	epsilon: 0.837901360003519
steps: 255	lr: 4e-05	evaluation reward: 2.83	
episode: 1364	score: 2.0	memory length: 263932	epsilon: 0.8377063300035232
steps: 197	lr: 4e-05	evaluation reward: 2.83	
episode: 1365	score: 3.0	memory length: 264158	epsilon: 0.8374825900035281
steps: 226	lr: 4e-05	evaluation reward: 2.81	
episode: 1366	score: 2.0	memory length: 264357	epsilon: 0.8372855800035324
steps: 199	lr: 4e-05	evaluation reward: 2.83	
episode: 1367	score: 4.0	memory length: 264654	epsilon: 0.8369915500035388
steps: 297	lr: 4e-05	evaluation reward: 2.87	
episode: 1368	score: 4.0	memory length: 264933	epsilon: 0.8367153400035447
steps: 279	lr: 4e-05	evaluation reward: 2.89	
episode: 1369	score: 3.0	memory length: 265164	epsilon: 0.8364866500035497
steps: 231	lr: 4e-05	evaluation reward: 2.92	
episode: 1370	score: 3.0	memory length: 265412	epsilon: 0.836241130003555
steps: 248	lr: 4e-05	evaluation reward: 2.91	
episode: 1371	score: 2.0	memory length: 265628	epsilon: 0.8360272900035597
steps: 216	lr: 4e-05	evaluation reward: 2.9	
episode: 1372	score: 3.0	memory length: 265857	epsilon: 0.8358005800035646
steps: 229	lr: 4e-05	evaluation reward: 2.92	
episode: 1373	score: 5.0	memory length: 266183	epsilon: 0.8354778400035716
steps: 326	lr: 4e-05	evaluation reward: 2.94	
episode: 1374	score: 5.0	memory length: 266490	epsilon: 0.8351739100035782
steps: 307	lr: 4e-05	evaluation reward: 2.97	
episode: 1375	score: 1.0	memory length: 266641	epsilon: 0.8350244200035815
steps: 151	lr: 4e-05	evaluation reward: 2.97	
episode: 1376	score: 1.0	memory length: 266792	epsilon: 0.8348749300035847
steps: 151	lr: 4e-05	evaluation reward: 2.96	
episode: 1377	score: 0.0	memory length: 266915	epsilon: 0.8347531600035873
steps: 123	lr: 4e-05	evaluation reward: 2.91	
episode: 1378	score: 3.0	memory length: 267146	epsilon: 0.8345244700035923
steps: 231	lr: 4e-05	evaluation reward: 2.9	
episode: 1379	score: 3.0	memory length: 267395	epsilon: 0.8342779600035977
steps: 249	lr: 4e-05	evaluation reward: 2.92	
episode: 1380	score: 3.0	memory length: 267643	epsilon: 0.834032440003603
steps: 248	lr: 4e-05	evaluation reward: 2.91	
episode: 1381	score: 4.0	memory length: 267918	epsilon: 0.8337601900036089
steps: 275	lr: 4e-05	evaluation reward: 2.95	
episode: 1382	score: 3.0	memory length: 268149	epsilon: 0.8335315000036139

steps: 231	lr: 4e-05	evaluation reward: 2.98	
episode: 1383	score: 2.0	memory length: 268331	epsilon: 0.8333513200036178
steps: 182	lr: 4e-05	evaluation reward: 2.96	
episode: 1384	score: 1.0	memory length: 268482	epsilon: 0.833201830003621
steps: 151	lr: 4e-05	evaluation reward: 2.94	
episode: 1385	score: 4.0	memory length: 268779	epsilon: 0.8329078000036274
steps: 297	lr: 4e-05	evaluation reward: 2.95	
episode: 1386	score: 1.0	memory length: 268951	epsilon: 0.8327375200036311
steps: 172	lr: 4e-05	evaluation reward: 2.94	
episode: 1387	score: 3.0	memory length: 269198	epsilon: 0.8324929900036364
steps: 247	lr: 4e-05	evaluation reward: 2.92	
episode: 1388	score: 2.0	memory length: 269418	epsilon: 0.8322751900036411
steps: 220	lr: 4e-05	evaluation reward: 2.9	
episode: 1389	score: 4.0	memory length: 269691	epsilon: 0.832004920003647
steps: 273	lr: 4e-05	evaluation reward: 2.89	
episode: 1390	score: 2.0	memory length: 269910	epsilon: 0.8317881100036517
steps: 219	lr: 4e-05	evaluation reward: 2.89	
episode: 1391	score: 3.0	memory length: 270177	epsilon: 0.8315237800036575
steps: 267	lr: 4e-05	evaluation reward: 2.89	
episode: 1392	score: 3.0	memory length: 270426	epsilon: 0.8312772700036628
steps: 249	lr: 4e-05	evaluation reward: 2.9	
episode: 1393	score: 3.0	memory length: 270672	epsilon: 0.8310337300036681
steps: 246	lr: 4e-05	evaluation reward: 2.9	
episode: 1394	score: 3.0	memory length: 270899	epsilon: 0.830809000003673
steps: 227	lr: 4e-05	evaluation reward: 2.9	
episode: 1395	score: 2.0	memory length: 271115	epsilon: 0.8305951600036776
steps: 216	lr: 4e-05	evaluation reward: 2.9	
episode: 1396	score: 2.0	memory length: 271295	epsilon: 0.8304169600036815
steps: 180	lr: 4e-05	evaluation reward: 2.87	
episode: 1397	score: 5.0	memory length: 271601	epsilon: 0.8301140200036881
steps: 306	lr: 4e-05	evaluation reward: 2.87	
episode: 1398	score: 2.0	memory length: 271798	epsilon: 0.8299189900036923
steps: 197	lr: 4e-05	evaluation reward: 2.89	
episode: 1399	score: 9.0	memory length: 272202	epsilon: 0.829519030003701
steps: 404	lr: 4e-05	evaluation reward: 2.93	
episode: 1400	score: 0.0	memory length: 272324	epsilon: 0.8293982500037036
steps: 122	lr: 4e-05	evaluation reward: 2.89	
episode: 1401	score: 3.0	memory length: 272571	epsilon: 0.8291537200037089
steps: 247	lr: 4e-05	evaluation reward: 2.91	
episode: 1402	score: 4.0	memory length: 272866	epsilon: 0.8288616700037152
steps: 295	lr: 4e-05	evaluation reward: 2.92	
episode: 1403	score: 3.0	memory length: 273113	epsilon: 0.8286171400037206
steps: 247	lr: 4e-05	evaluation reward: 2.92	
episode: 1404	score: 3.0	memory length: 273327	epsilon: 0.8284052800037252
steps: 214	lr: 4e-05	evaluation reward: 2.89	
episode: 1405	score: 5.0	memory length: 273665	epsilon: 0.8280706600037324
steps: 338	lr: 4e-05	evaluation reward: 2.9	
episode: 1406	score: 0.0	memory length: 273788	epsilon: 0.8279488900037351

steps: 123	lr: 4e-05	evaluation reward: 2.89	
episode: 1407	score: 7.0	memory length: 274186	epsilon: 0.8275548700037436
steps: 398	lr: 4e-05	evaluation reward: 2.93	
episode: 1408	score: 1.0	memory length: 274337	epsilon: 0.8274053800037469
steps: 151	lr: 4e-05	evaluation reward: 2.91	
episode: 1409	score: 4.0	memory length: 274632	epsilon: 0.8271133300037532
steps: 295	lr: 4e-05	evaluation reward: 2.94	
episode: 1410	score: 2.0	memory length: 274849	epsilon: 0.8268985000037579
steps: 217	lr: 4e-05	evaluation reward: 2.94	
episode: 1411	score: 3.0	memory length: 275116	epsilon: 0.8266341700037636
steps: 267	lr: 4e-05	evaluation reward: 2.95	
episode: 1412	score: 6.0	memory length: 275509	epsilon: 0.826245100003772
steps: 393	lr: 4e-05	evaluation reward: 2.98	
episode: 1413	score: 0.0	memory length: 275631	epsilon: 0.8261243200037747
steps: 122	lr: 4e-05	evaluation reward: 2.97	
episode: 1414	score: 4.0	memory length: 275875	epsilon: 0.8258827600037799
steps: 244	lr: 4e-05	evaluation reward: 2.99	
episode: 1415	score: 2.0	memory length: 276054	epsilon: 0.8257055500037838
steps: 179	lr: 4e-05	evaluation reward: 2.94	
episode: 1416	score: 1.0	memory length: 276206	epsilon: 0.825555070003787
steps: 152	lr: 4e-05	evaluation reward: 2.94	
episode: 1417	score: 0.0	memory length: 276328	epsilon: 0.8254342900037897
steps: 122	lr: 4e-05	evaluation reward: 2.93	
episode: 1418	score: 3.0	memory length: 276558	epsilon: 0.8252065900037946
steps: 230	lr: 4e-05	evaluation reward: 2.93	
episode: 1419	score: 1.0	memory length: 276727	epsilon: 0.8250392800037982
steps: 169	lr: 4e-05	evaluation reward: 2.92	
episode: 1420	score: 3.0	memory length: 276996	epsilon: 0.824772970003804
steps: 269	lr: 4e-05	evaluation reward: 2.92	
episode: 1421	score: 2.0	memory length: 277194	epsilon: 0.8245769500038083
steps: 198	lr: 4e-05	evaluation reward: 2.92	
episode: 1422	score: 5.0	memory length: 277563	epsilon: 0.8242116400038162
steps: 369	lr: 4e-05	evaluation reward: 2.85	
episode: 1423	score: 1.0	memory length: 277732	epsilon: 0.8240443300038198
steps: 169	lr: 4e-05	evaluation reward: 2.83	
episode: 1424	score: 5.0	memory length: 278042	epsilon: 0.8237374300038265
steps: 310	lr: 4e-05	evaluation reward: 2.8	
episode: 1425	score: 4.0	memory length: 278316	epsilon: 0.8234661700038324
steps: 274	lr: 4e-05	evaluation reward: 2.81	
episode: 1426	score: 3.0	memory length: 278563	epsilon: 0.8232216400038377
steps: 247	lr: 4e-05	evaluation reward: 2.82	
episode: 1427	score: 3.0	memory length: 278811	epsilon: 0.822976120003843
steps: 248	lr: 4e-05	evaluation reward: 2.81	
episode: 1428	score: 4.0	memory length: 279086	epsilon: 0.8227038700038489
steps: 275	lr: 4e-05	evaluation reward: 2.84	
episode: 1429	score: 4.0	memory length: 279346	epsilon: 0.8224464700038545
steps: 260	lr: 4e-05	evaluation reward: 2.87	
episode: 1430	score: 3.0	memory length: 279594	epsilon: 0.8222009500038598

steps: 248	lr: 4e-05	evaluation reward: 2.87	
episode: 1431	score: 3.0	memory length: 279858	epsilon: 0.8219395900038655
steps: 264	lr: 4e-05	evaluation reward: 2.84	
episode: 1432	score: 5.0	memory length: 280206	epsilon: 0.821595070003873
steps: 348	lr: 4e-05	evaluation reward: 2.85	
episode: 1433	score: 1.0	memory length: 280356	epsilon: 0.8214465700038762
steps: 150	lr: 4e-05	evaluation reward: 2.81	
episode: 1434	score: 4.0	memory length: 280616	epsilon: 0.8211891700038818
steps: 260	lr: 4e-05	evaluation reward: 2.81	
episode: 1435	score: 2.0	memory length: 280797	epsilon: 0.8210099800038857
steps: 181	lr: 4e-05	evaluation reward: 2.81	
episode: 1436	score: 6.0	memory length: 281189	epsilon: 0.8206219000038941
steps: 392	lr: 4e-05	evaluation reward: 2.87	
episode: 1437	score: 4.0	memory length: 281487	epsilon: 0.8203268800039005
steps: 298	lr: 4e-05	evaluation reward: 2.9	
episode: 1438	score: 1.0	memory length: 281637	epsilon: 0.8201783800039038
steps: 150	lr: 4e-05	evaluation reward: 2.9	
episode: 1439	score: 3.0	memory length: 281886	epsilon: 0.8199318700039091
steps: 249	lr: 4e-05	evaluation reward: 2.88	
episode: 1440	score: 6.0	memory length: 282244	epsilon: 0.8195774500039168
steps: 358	lr: 4e-05	evaluation reward: 2.88	
episode: 1441	score: 2.0	memory length: 282425	epsilon: 0.8193982600039207
steps: 181	lr: 4e-05	evaluation reward: 2.88	
episode: 1442	score: 2.0	memory length: 282623	epsilon: 0.8192022400039249
steps: 198	lr: 4e-05	evaluation reward: 2.9	
episode: 1443	score: 2.0	memory length: 282822	epsilon: 0.8190052300039292
steps: 199	lr: 4e-05	evaluation reward: 2.92	
episode: 1444	score: 0.0	memory length: 282945	epsilon: 0.8188834600039319
steps: 123	lr: 4e-05	evaluation reward: 2.92	
episode: 1445	score: 3.0	memory length: 283192	epsilon: 0.8186389300039372
steps: 247	lr: 4e-05	evaluation reward: 2.91	
episode: 1446	score: 6.0	memory length: 283575	epsilon: 0.8182597600039454
steps: 383	lr: 4e-05	evaluation reward: 2.93	
episode: 1447	score: 4.0	memory length: 283872	epsilon: 0.8179657300039518
steps: 297	lr: 4e-05	evaluation reward: 2.93	
episode: 1448	score: 3.0	memory length: 284144	epsilon: 0.8176964500039576
steps: 272	lr: 4e-05	evaluation reward: 2.96	
episode: 1449	score: 3.0	memory length: 284372	epsilon: 0.8174707300039625
steps: 228	lr: 4e-05	evaluation reward: 2.95	
episode: 1450	score: 2.0	memory length: 284551	epsilon: 0.8172935200039664
steps: 179	lr: 4e-05	evaluation reward: 2.91	
episode: 1451	score: 1.0	memory length: 284719	epsilon: 0.81712720000397
steps: 168	lr: 4e-05	evaluation reward: 2.89	
episode: 1452	score: 3.0	memory length: 284933	epsilon: 0.8169153400039746
steps: 214	lr: 4e-05	evaluation reward: 2.87	
episode: 1453	score: 2.0	memory length: 285131	epsilon: 0.8167193200039788
steps: 198	lr: 4e-05	evaluation reward: 2.89	
episode: 1454	score: 4.0	memory length: 285410	epsilon: 0.8164431100039848

steps: 279	lr: 4e-05	evaluation reward: 2.9	
episode: 1455	score: 3.0	memory length: 285656	epsilon: 0.8161995700039901
steps: 246	lr: 4e-05	evaluation reward: 2.87	
episode: 1456	score: 5.0	memory length: 285982	epsilon: 0.8158768300039971
steps: 326	lr: 4e-05	evaluation reward: 2.89	
episode: 1457	score: 2.0	memory length: 286204	epsilon: 0.8156570500040019
steps: 222	lr: 4e-05	evaluation reward: 2.89	
episode: 1458	score: 3.0	memory length: 286453	epsilon: 0.8154105400040073
steps: 249	lr: 4e-05	evaluation reward: 2.92	
episode: 1459	score: 7.0	memory length: 286865	epsilon: 0.8150026600040161
steps: 412	lr: 4e-05	evaluation reward: 2.97	
episode: 1460	score: 5.0	memory length: 287191	epsilon: 0.8146799200040231
steps: 326	lr: 4e-05	evaluation reward: 3.0	
episode: 1461	score: 2.0	memory length: 287409	epsilon: 0.8144641000040278
steps: 218	lr: 4e-05	evaluation reward: 3.02	
episode: 1462	score: 4.0	memory length: 287684	epsilon: 0.8141918500040337
steps: 275	lr: 4e-05	evaluation reward: 3.01	
episode: 1463	score: 2.0	memory length: 287906	epsilon: 0.8139720700040385
steps: 222	lr: 4e-05	evaluation reward: 2.99	
episode: 1464	score: 3.0	memory length: 288152	epsilon: 0.8137285300040438
steps: 246	lr: 4e-05	evaluation reward: 3.0	
episode: 1465	score: 2.0	memory length: 288349	epsilon: 0.813533500004048
steps: 197	lr: 4e-05	evaluation reward: 2.99	
episode: 1466	score: 0.0	memory length: 288472	epsilon: 0.8134117300040506
steps: 123	lr: 4e-05	evaluation reward: 2.97	
episode: 1467	score: 4.0	memory length: 288744	epsilon: 0.8131424500040565
steps: 272	lr: 4e-05	evaluation reward: 2.97	
episode: 1468	score: 3.0	memory length: 288970	epsilon: 0.8129187100040614
steps: 226	lr: 4e-05	evaluation reward: 2.96	
episode: 1469	score: 3.0	memory length: 289195	epsilon: 0.8126959600040662
steps: 225	lr: 4e-05	evaluation reward: 2.96	
episode: 1470	score: 5.0	memory length: 289523	epsilon: 0.8123712400040732
steps: 328	lr: 4e-05	evaluation reward: 2.98	
episode: 1471	score: 3.0	memory length: 289769	epsilon: 0.8121277000040785
steps: 246	lr: 4e-05	evaluation reward: 2.99	
episode: 1472	score: 3.0	memory length: 289999	epsilon: 0.8119000000040835
steps: 230	lr: 4e-05	evaluation reward: 2.99	
episode: 1473	score: 2.0	memory length: 290197	epsilon: 0.8117039800040877
steps: 198	lr: 4e-05	evaluation reward: 2.96	
episode: 1474	score: 6.0	memory length: 290534	epsilon: 0.811370350004095
steps: 337	lr: 4e-05	evaluation reward: 2.97	
episode: 1475	score: 1.0	memory length: 290684	epsilon: 0.8112218500040982
steps: 150	lr: 4e-05	evaluation reward: 2.97	
episode: 1476	score: 3.0	memory length: 290928	epsilon: 0.8109802900041034
steps: 244	lr: 4e-05	evaluation reward: 2.99	
episode: 1477	score: 1.0	memory length: 291096	epsilon: 0.810813970004107
steps: 168	lr: 4e-05	evaluation reward: 3.0	
episode: 1478	score: 1.0	memory length: 291246	epsilon: 0.8106654700041103

steps: 150	lr: 4e-05	evaluation reward: 2.98	
episode: 1479	score: 3.0	memory length: 291490	epsilon: 0.8104239100041155
steps: 244	lr: 4e-05	evaluation reward: 2.98	
episode: 1480	score: 1.0	memory length: 291641	epsilon: 0.8102744200041188
steps: 151	lr: 4e-05	evaluation reward: 2.96	
episode: 1481	score: 4.0	memory length: 291900	epsilon: 0.8100180100041243
steps: 259	lr: 4e-05	evaluation reward: 2.96	
episode: 1482	score: 3.0	memory length: 292168	epsilon: 0.8097526900041301
steps: 268	lr: 4e-05	evaluation reward: 2.96	
episode: 1483	score: 2.0	memory length: 292366	epsilon: 0.8095566700041343
steps: 198	lr: 4e-05	evaluation reward: 2.96	
episode: 1484	score: 5.0	memory length: 292664	epsilon: 0.8092616500041407
steps: 298	lr: 4e-05	evaluation reward: 3.0	
episode: 1485	score: 1.0	memory length: 292832	epsilon: 0.8090953300041444
steps: 168	lr: 4e-05	evaluation reward: 2.97	
episode: 1486	score: 2.0	memory length: 293052	epsilon: 0.8088775300041491
steps: 220	lr: 4e-05	evaluation reward: 2.98	
episode: 1487	score: 1.0	memory length: 293220	epsilon: 0.8087112100041527
steps: 168	lr: 4e-05	evaluation reward: 2.96	
episode: 1488	score: 2.0	memory length: 293422	epsilon: 0.808511230004157
steps: 202	lr: 4e-05	evaluation reward: 2.96	
episode: 1489	score: 7.0	memory length: 293834	epsilon: 0.8081033500041659
steps: 412	lr: 4e-05	evaluation reward: 2.99	
episode: 1490	score: 4.0	memory length: 294112	epsilon: 0.8078281300041719
steps: 278	lr: 4e-05	evaluation reward: 3.01	
episode: 1491	score: 0.0	memory length: 294235	epsilon: 0.8077063600041745
steps: 123	lr: 4e-05	evaluation reward: 2.98	
episode: 1492	score: 2.0	memory length: 294433	epsilon: 0.8075103400041788
steps: 198	lr: 4e-05	evaluation reward: 2.97	
episode: 1493	score: 1.0	memory length: 294583	epsilon: 0.807361840004182
steps: 150	lr: 4e-05	evaluation reward: 2.95	
episode: 1494	score: 2.0	memory length: 294763	epsilon: 0.8071836400041859
steps: 180	lr: 4e-05	evaluation reward: 2.94	
episode: 1495	score: 5.0	memory length: 295072	epsilon: 0.8068777300041925
steps: 309	lr: 4e-05	evaluation reward: 2.97	
episode: 1496	score: 0.0	memory length: 295195	epsilon: 0.8067559600041951
steps: 123	lr: 4e-05	evaluation reward: 2.95	
episode: 1497	score: 4.0	memory length: 295470	epsilon: 0.806483710004201
steps: 275	lr: 4e-05	evaluation reward: 2.94	
episode: 1498	score: 4.0	memory length: 295764	epsilon: 0.8061926500042074
steps: 294	lr: 4e-05	evaluation reward: 2.96	
episode: 1499	score: 3.0	memory length: 295992	epsilon: 0.8059669300042123
steps: 228	lr: 4e-05	evaluation reward: 2.9	
episode: 1500	score: 3.0	memory length: 296222	epsilon: 0.8057392300042172
steps: 230	lr: 4e-05	evaluation reward: 2.93	
episode: 1501	score: 3.0	memory length: 296448	epsilon: 0.8055154900042221
steps: 226	lr: 4e-05	evaluation reward: 2.93	
episode: 1502	score: 8.0	memory length: 296833	epsilon: 0.8051343400042303

steps: 385	lr: 4e-05	evaluation reward: 2.97	
episode: 1503	score: 2.0	memory length: 297031	epsilon: 0.8049383200042346
steps: 198	lr: 4e-05	evaluation reward: 2.96	
episode: 1504	score: 0.0	memory length: 297154	epsilon: 0.8048165500042372
steps: 123	lr: 4e-05	evaluation reward: 2.93	
episode: 1505	score: 4.0	memory length: 297470	epsilon: 0.804503710004244
steps: 316	lr: 4e-05	evaluation reward: 2.92	
episode: 1506	score: 2.0	memory length: 297670	epsilon: 0.8043057100042483
steps: 200	lr: 4e-05	evaluation reward: 2.94	
episode: 1507	score: 1.0	memory length: 297841	epsilon: 0.804136420004252
steps: 171	lr: 4e-05	evaluation reward: 2.88	
episode: 1508	score: 3.0	memory length: 298071	epsilon: 0.803908720004257
steps: 230	lr: 4e-05	evaluation reward: 2.9	
episode: 1509	score: 4.0	memory length: 298369	epsilon: 0.8036137000042634
steps: 298	lr: 4e-05	evaluation reward: 2.9	
episode: 1510	score: 3.0	memory length: 298600	epsilon: 0.8033850100042683
steps: 231	lr: 4e-05	evaluation reward: 2.91	
episode: 1511	score: 4.0	memory length: 298874	epsilon: 0.8031137500042742
steps: 274	lr: 4e-05	evaluation reward: 2.92	
episode: 1512	score: 4.0	memory length: 299170	epsilon: 0.8028207100042806
steps: 296	lr: 4e-05	evaluation reward: 2.9	
episode: 1513	score: 5.0	memory length: 299499	epsilon: 0.8024950000042876
steps: 329	lr: 4e-05	evaluation reward: 2.95	
episode: 1514	score: 2.0	memory length: 299715	epsilon: 0.8022811600042923
steps: 216	lr: 4e-05	evaluation reward: 2.93	
episode: 1515	score: 3.0	memory length: 299942	epsilon: 0.8020564300042972
steps: 227	lr: 4e-05	evaluation reward: 2.94	
episode: 1516	score: 2.0	memory length: 300140	epsilon: 0.8018604100043014
steps: 198	lr: 1.6000000000000003e-05	evaluation reward: 2.95	
episode: 1517	score: 3.0	memory length: 300389	epsilon: 0.8016139000043068
steps: 249	lr: 1.6000000000000003e-05	evaluation reward: 2.98	
episode: 1518	score: 6.0	memory length: 300786	epsilon: 0.8012208700043153
steps: 397	lr: 1.6000000000000003e-05	evaluation reward: 3.01	
episode: 1519	score: 2.0	memory length: 300985	epsilon: 0.8010238600043196
steps: 199	lr: 1.6000000000000003e-05	evaluation reward: 3.02	
episode: 1520	score: 1.0	memory length: 301135	epsilon: 0.8008753600043228
steps: 150	lr: 1.6000000000000003e-05	evaluation reward: 3.0	
episode: 1521	score: 3.0	memory length: 301384	epsilon: 0.8006288500043282
steps: 249	lr: 1.6000000000000003e-05	evaluation reward: 3.01	
episode: 1522	score: 0.0	memory length: 301507	epsilon: 0.8005070800043308
steps: 123	lr: 1.6000000000000003e-05	evaluation reward: 2.96	
episode: 1523	score: 4.0	memory length: 301800	epsilon: 0.8002170100043371
steps: 293	lr: 1.6000000000000003e-05	evaluation reward: 2.99	
episode: 1524	score: 6.0	memory length: 302175	epsilon: 0.7998457600043452
steps: 375	lr: 1.6000000000000003e-05	evaluation reward: 3.0	
episode: 1525	score: 3.0	memory length: 302401	epsilon: 0.79962202000435
steps: 226	lr: 1.6000000000000003e-05	evaluation reward: 2.99	
episode: 1526	score: 0.0	memory length: 302524	epsilon: 0.7995002500043527

```

steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1527  score: 3.0    memory length: 302768    epsilon: 0.7992586900043579
steps: 244    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1528  score: 2.0    memory length: 302949    epsilon: 0.7990795000043618
steps: 181    lr: 1.6000000000000003e-05    evaluation reward: 2.94
episode: 1529  score: 6.0    memory length: 303276    epsilon: 0.7987557700043688
steps: 327    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1530  score: 2.0    memory length: 303477    epsilon: 0.7985567800043731
steps: 201    lr: 1.6000000000000003e-05    evaluation reward: 2.95
episode: 1531  score: 4.0    memory length: 303739    epsilon: 0.7982974000043788
steps: 262    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1532  score: 9.0    memory length: 304198    epsilon: 0.7978429900043886
steps: 459    lr: 1.6000000000000003e-05    evaluation reward: 3.0
episode: 1533  score: 2.0    memory length: 304416    epsilon: 0.7976271700043933
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1534  score: 0.0    memory length: 304539    epsilon: 0.797505400004396
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 2.97
episode: 1535  score: 3.0    memory length: 304770    epsilon: 0.7972767100044009
steps: 231    lr: 1.6000000000000003e-05    evaluation reward: 2.98
episode: 1536  score: 4.0    memory length: 305046    epsilon: 0.7970034700044069
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 2.96
episode: 1537  score: 8.0    memory length: 305399    epsilon: 0.7966540000044144
steps: 353    lr: 1.6000000000000003e-05    evaluation reward: 3.0
episode: 1538  score: 2.0    memory length: 305597    epsilon: 0.7964579800044187
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1539  score: 3.0    memory length: 305823    epsilon: 0.7962342400044236
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.01
episode: 1540  score: 4.0    memory length: 306077    epsilon: 0.795982780004429
steps: 254    lr: 1.6000000000000003e-05    evaluation reward: 2.99
episode: 1541  score: 3.0    memory length: 306305    epsilon: 0.7957570600044339
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 3.0
episode: 1542  score: 7.0    memory length: 306683    epsilon: 0.795382840004442
steps: 378    lr: 1.6000000000000003e-05    evaluation reward: 3.05
episode: 1543  score: 3.0    memory length: 306931    epsilon: 0.7951373200044474
steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 3.06
episode: 1544  score: 4.0    memory length: 307205    epsilon: 0.7948660600044533
steps: 274    lr: 1.6000000000000003e-05    evaluation reward: 3.1
episode: 1545  score: 3.0    memory length: 307472    epsilon: 0.794601730004459
steps: 267    lr: 1.6000000000000003e-05    evaluation reward: 3.1
episode: 1546  score: 2.0    memory length: 307670    epsilon: 0.7944057100044633
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.06
episode: 1547  score: 3.0    memory length: 307919    epsilon: 0.7941592000044686
steps: 249    lr: 1.6000000000000003e-05    evaluation reward: 3.05
episode: 1548  score: 4.0    memory length: 308195    epsilon: 0.7938859600044745
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 3.06
episode: 1549  score: 6.0    memory length: 308580    epsilon: 0.7935048100044828
steps: 385    lr: 1.6000000000000003e-05    evaluation reward: 3.09
episode: 1550  score: 6.0    memory length: 308938    epsilon: 0.7931503900044905

```

```

steps: 358    lr: 1.6000000000000003e-05    evaluation reward: 3.13
episode: 1551  score: 5.0    memory length: 309245    epsilon: 0.7928464600044971
steps: 307    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1552  score: 2.0    memory length: 309463    epsilon: 0.7926306400045018
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.16
episode: 1553  score: 6.0    memory length: 309838    epsilon: 0.7922593900045098
steps: 375    lr: 1.6000000000000003e-05    evaluation reward: 3.2
episode: 1554  score: 0.0    memory length: 309961    epsilon: 0.7921376200045125
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.16
episode: 1555  score: 3.0    memory length: 310190    epsilon: 0.7919109100045174
steps: 229    lr: 1.6000000000000003e-05    evaluation reward: 3.16
episode: 1556  score: 6.0    memory length: 310537    epsilon: 0.7915673800045249
steps: 347    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1557  score: 2.0    memory length: 310735    epsilon: 0.7913713600045291
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1558  score: 6.0    memory length: 311068    epsilon: 0.7910416900045363
steps: 333    lr: 1.6000000000000003e-05    evaluation reward: 3.2
episode: 1559  score: 2.0    memory length: 311248    epsilon: 0.7908634900045401
steps: 180    lr: 1.6000000000000003e-05    evaluation reward: 3.15
episode: 1560  score: 4.0    memory length: 311523    epsilon: 0.7905912400045461
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.14
episode: 1561  score: 5.0    memory length: 311829    epsilon: 0.7902883000045526
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1562  score: 2.0    memory length: 312050    epsilon: 0.7900695100045574
steps: 221    lr: 1.6000000000000003e-05    evaluation reward: 3.15
episode: 1563  score: 2.0    memory length: 312248    epsilon: 0.7898734900045616
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.15
episode: 1564  score: 4.0    memory length: 312530    epsilon: 0.7895943100045677
steps: 282    lr: 1.6000000000000003e-05    evaluation reward: 3.16
episode: 1565  score: 3.0    memory length: 312741    epsilon: 0.7893854200045722
steps: 211    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1566  score: 0.0    memory length: 312864    epsilon: 0.7892636500045749
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1567  score: 5.0    memory length: 313188    epsilon: 0.7889428900045818
steps: 324    lr: 1.6000000000000003e-05    evaluation reward: 3.18
episode: 1568  score: 3.0    memory length: 313415    epsilon: 0.7887181600045867
steps: 227    lr: 1.6000000000000003e-05    evaluation reward: 3.18
episode: 1569  score: 4.0    memory length: 313675    epsilon: 0.7884607600045923
steps: 260    lr: 1.6000000000000003e-05    evaluation reward: 3.19
episode: 1570  score: 3.0    memory length: 313908    epsilon: 0.7882300900045973
steps: 233    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1571  score: 3.0    memory length: 314151    epsilon: 0.7879895200046025
steps: 243    lr: 1.6000000000000003e-05    evaluation reward: 3.17
episode: 1572  score: 4.0    memory length: 314392    epsilon: 0.7877509300046077
steps: 241    lr: 1.6000000000000003e-05    evaluation reward: 3.18
episode: 1573  score: 8.0    memory length: 314789    epsilon: 0.7873579000046163
steps: 397    lr: 1.6000000000000003e-05    evaluation reward: 3.24
episode: 1574  score: 6.0    memory length: 315149    epsilon: 0.787001500004624

```

```

steps: 360    lr: 1.6000000000000003e-05    evaluation reward: 3.24
episode: 1575  score: 3.0    memory length: 315375    epsilon: 0.7867777600046288
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.26
episode: 1576  score: 5.0    memory length: 315698    epsilon: 0.7864579900046358
steps: 323    lr: 1.6000000000000003e-05    evaluation reward: 3.28
episode: 1577  score: 6.0    memory length: 316046    epsilon: 0.7861134700046433
steps: 348    lr: 1.6000000000000003e-05    evaluation reward: 3.33
episode: 1578  score: 6.0    memory length: 316425    epsilon: 0.7857382600046514
steps: 379    lr: 1.6000000000000003e-05    evaluation reward: 3.38
episode: 1579  score: 2.0    memory length: 316643    epsilon: 0.7855224400046561
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.37
episode: 1580  score: 4.0    memory length: 316921    epsilon: 0.7852472200046621
steps: 278    lr: 1.6000000000000003e-05    evaluation reward: 3.4
episode: 1581  score: 4.0    memory length: 317216    epsilon: 0.7849551700046684
steps: 295    lr: 1.6000000000000003e-05    evaluation reward: 3.4
episode: 1582  score: 2.0    memory length: 317433    epsilon: 0.7847403400046731
steps: 217    lr: 1.6000000000000003e-05    evaluation reward: 3.39
episode: 1583  score: 1.0    memory length: 317584    epsilon: 0.7845908500046763
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.38
episode: 1584  score: 1.0    memory length: 317754    epsilon: 0.78442255000468
steps: 170    lr: 1.6000000000000003e-05    evaluation reward: 3.34
episode: 1585  score: 3.0    memory length: 317967    epsilon: 0.7842116800046846
steps: 213    lr: 1.6000000000000003e-05    evaluation reward: 3.36
episode: 1586  score: 1.0    memory length: 318136    epsilon: 0.7840443700046882
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 3.35
episode: 1587  score: 1.0    memory length: 318304    epsilon: 0.7838780500046918
steps: 168    lr: 1.6000000000000003e-05    evaluation reward: 3.35
episode: 1588  score: 3.0    memory length: 318574    epsilon: 0.7836107500046976
steps: 270    lr: 1.6000000000000003e-05    evaluation reward: 3.36
episode: 1589  score: 5.0    memory length: 318896    epsilon: 0.7832919700047045
steps: 322    lr: 1.6000000000000003e-05    evaluation reward: 3.34
episode: 1590  score: 3.0    memory length: 319108    epsilon: 0.7830820900047091
steps: 212    lr: 1.6000000000000003e-05    evaluation reward: 3.33
episode: 1591  score: 4.0    memory length: 319369    epsilon: 0.7828237000047147
steps: 261    lr: 1.6000000000000003e-05    evaluation reward: 3.37
episode: 1592  score: 1.0    memory length: 319520    epsilon: 0.7826742100047179
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.36
episode: 1593  score: 8.0    memory length: 319943    epsilon: 0.782255440004727
steps: 423    lr: 1.6000000000000003e-05    evaluation reward: 3.43
episode: 1594  score: 0.0    memory length: 320066    epsilon: 0.7821336700047297
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.41
episode: 1595  score: 4.0    memory length: 320358    epsilon: 0.7818445900047359
steps: 292    lr: 1.6000000000000003e-05    evaluation reward: 3.4
episode: 1596  score: 9.0    memory length: 320660    epsilon: 0.7815456100047424
steps: 302    lr: 1.6000000000000003e-05    evaluation reward: 3.49
episode: 1597  score: 3.0    memory length: 320886    epsilon: 0.7813218700047473
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.48
episode: 1598  score: 5.0    memory length: 321191    epsilon: 0.7810199200047538

```

```

steps: 305    lr: 1.6000000000000003e-05    evaluation reward: 3.49
episode: 1599  score: 3.0    memory length: 321400    epsilon: 0.7808130100047583
steps: 209    lr: 1.6000000000000003e-05    evaluation reward: 3.49
episode: 1600  score: 6.0    memory length: 321744    epsilon: 0.7804724500047657
steps: 344    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1601  score: 4.0    memory length: 322037    epsilon: 0.780182380004772
steps: 293    lr: 1.6000000000000003e-05    evaluation reward: 3.53
episode: 1602  score: 4.0    memory length: 322313    epsilon: 0.779909140004778
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 3.49
episode: 1603  score: 6.0    memory length: 322681    epsilon: 0.7795448200047859
steps: 368    lr: 1.6000000000000003e-05    evaluation reward: 3.53
episode: 1604  score: 4.0    memory length: 322978    epsilon: 0.7792507900047922
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 3.57
episode: 1605  score: 3.0    memory length: 323226    epsilon: 0.7790052700047976
steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 3.56
episode: 1606  score: 2.0    memory length: 323444    epsilon: 0.7787894500048023
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.56
episode: 1607  score: 2.0    memory length: 323644    epsilon: 0.7785914500048066
steps: 200    lr: 1.6000000000000003e-05    evaluation reward: 3.57
episode: 1608  score: 2.0    memory length: 323825    epsilon: 0.7784122600048105
steps: 181    lr: 1.6000000000000003e-05    evaluation reward: 3.56
episode: 1609  score: 6.0    memory length: 324182    epsilon: 0.7780588300048181
steps: 357    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1610  score: 4.0    memory length: 324457    epsilon: 0.777786580004824
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.59
episode: 1611  score: 7.0    memory length: 324824    epsilon: 0.7774232500048319
steps: 367    lr: 1.6000000000000003e-05    evaluation reward: 3.62
episode: 1612  score: 4.0    memory length: 325100    epsilon: 0.7771500100048379
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 3.62
episode: 1613  score: 7.0    memory length: 325547    epsilon: 0.7767074800048475
steps: 447    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1614  score: 7.0    memory length: 325955    epsilon: 0.7763035600048562
steps: 408    lr: 1.6000000000000003e-05    evaluation reward: 3.69
episode: 1615  score: 8.0    memory length: 326407    epsilon: 0.7758560800048659
steps: 452    lr: 1.6000000000000003e-05    evaluation reward: 3.74
episode: 1616  score: 3.0    memory length: 326635    epsilon: 0.7756303600048708
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1617  score: 0.0    memory length: 326758    epsilon: 0.7755085900048735
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1618  score: 6.0    memory length: 327134    epsilon: 0.7751363500048816
steps: 376    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1619  score: 4.0    memory length: 327410    epsilon: 0.7748631100048875
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 3.74
episode: 1620  score: 2.0    memory length: 327628    epsilon: 0.7746472900048922
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1621  score: 3.0    memory length: 327854    epsilon: 0.774423550004897
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1622  score: 7.0    memory length: 328271    epsilon: 0.774010720004906

```

```

steps: 417    lr: 1.6000000000000003e-05    evaluation reward: 3.82
episode: 1623  score: 5.0    memory length: 328592    epsilon: 0.7736929300049129
steps: 321    lr: 1.6000000000000003e-05    evaluation reward: 3.83
episode: 1624  score: 0.0    memory length: 328715    epsilon: 0.7735711600049155
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.77
episode: 1625  score: 3.0    memory length: 328925    epsilon: 0.7733632600049201
steps: 210    lr: 1.6000000000000003e-05    evaluation reward: 3.77
episode: 1626  score: 7.0    memory length: 329349    epsilon: 0.7729435000049292
steps: 424    lr: 1.6000000000000003e-05    evaluation reward: 3.84
episode: 1627  score: 2.0    memory length: 329568    epsilon: 0.7727266900049339
steps: 219    lr: 1.6000000000000003e-05    evaluation reward: 3.83
episode: 1628  score: 1.0    memory length: 329719    epsilon: 0.7725772000049371
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.82
episode: 1629  score: 2.0    memory length: 329937    epsilon: 0.7723613800049418
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.78
episode: 1630  score: 2.0    memory length: 330135    epsilon: 0.7721653600049461
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.78
episode: 1631  score: 4.0    memory length: 330431    epsilon: 0.7718723200049524
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 3.78
episode: 1632  score: 3.0    memory length: 330659    epsilon: 0.7716466000049573
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1633  score: 6.0    memory length: 331038    epsilon: 0.7712713900049655
steps: 379    lr: 1.6000000000000003e-05    evaluation reward: 3.76
episode: 1634  score: 3.0    memory length: 331266    epsilon: 0.7710456700049704
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 3.79
episode: 1635  score: 3.0    memory length: 331511    epsilon: 0.7708031200049756
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 3.79
episode: 1636  score: 2.0    memory length: 331709    epsilon: 0.7706071000049799
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.77
episode: 1637  score: 3.0    memory length: 331920    epsilon: 0.7703982100049844
steps: 211    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1638  score: 1.0    memory length: 332091    epsilon: 0.7702289200049881
steps: 171    lr: 1.6000000000000003e-05    evaluation reward: 3.71
episode: 1639  score: 3.0    memory length: 332358    epsilon: 0.7699645900049938
steps: 267    lr: 1.6000000000000003e-05    evaluation reward: 3.71
episode: 1640  score: 6.0    memory length: 332706    epsilon: 0.7696200700050013
steps: 348    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1641  score: 3.0    memory length: 332952    epsilon: 0.7693765300050066
steps: 246    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1642  score: 0.0    memory length: 333074    epsilon: 0.7692557500050092
steps: 122    lr: 1.6000000000000003e-05    evaluation reward: 3.66
episode: 1643  score: 2.0    memory length: 333271    epsilon: 0.7690607200050135
steps: 197    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1644  score: 1.0    memory length: 333422    epsilon: 0.7689112300050167
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.62
episode: 1645  score: 4.0    memory length: 333715    epsilon: 0.768621160005023
steps: 293    lr: 1.6000000000000003e-05    evaluation reward: 3.63
episode: 1646  score: 4.0    memory length: 333992    epsilon: 0.768346930005029

```

```

steps: 277    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1647    score: 11.0    memory length: 334585    epsilon:
0.7677598600050417    steps: 593    lr: 1.6000000000000003e-05    evaluation
reward: 3.73
episode: 1648    score: 3.0    memory length: 334834    epsilon: 0.7675133500050471
steps: 249    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1649    score: 2.0    memory length: 335016    epsilon: 0.767333170005051
steps: 182    lr: 1.6000000000000003e-05    evaluation reward: 3.68
episode: 1650    score: 3.0    memory length: 335283    epsilon: 0.7670688400050567
steps: 267    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1651    score: 3.0    memory length: 335492    epsilon: 0.7668619300050612
steps: 209    lr: 1.6000000000000003e-05    evaluation reward: 3.63
episode: 1652    score: 5.0    memory length: 335816    epsilon: 0.7665411700050682
steps: 324    lr: 1.6000000000000003e-05    evaluation reward: 3.66
episode: 1653    score: 4.0    memory length: 336095    epsilon: 0.7662649600050742
steps: 279    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1654    score: 0.0    memory length: 336217    epsilon: 0.7661441800050768
steps: 122    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1655    score: 3.0    memory length: 336464    epsilon: 0.7658996500050821
steps: 247    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1656    score: 4.0    memory length: 336758    epsilon: 0.7656085900050884
steps: 294    lr: 1.6000000000000003e-05    evaluation reward: 3.62
episode: 1657    score: 1.0    memory length: 336909    epsilon: 0.7654591000050917
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.61
episode: 1658    score: 4.0    memory length: 337184    epsilon: 0.7651868500050976
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.59
episode: 1659    score: 6.0    memory length: 337521    epsilon: 0.7648532200051048
steps: 337    lr: 1.6000000000000003e-05    evaluation reward: 3.63
episode: 1660    score: 4.0    memory length: 337796    epsilon: 0.7645809700051107
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.63
episode: 1661    score: 6.0    memory length: 338151    epsilon: 0.7642295200051183
steps: 355    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1662    score: 3.0    memory length: 338380    epsilon: 0.7640028100051233
steps: 229    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1663    score: 2.0    memory length: 338578    epsilon: 0.7638067900051275
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1664    score: 1.0    memory length: 338728    epsilon: 0.7636582900051307
steps: 150    lr: 1.6000000000000003e-05    evaluation reward: 3.62
episode: 1665    score: 6.0    memory length: 339080    epsilon: 0.7633098100051383
steps: 352    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1666    score: 5.0    memory length: 339400    epsilon: 0.7629930100051452
steps: 320    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1667    score: 5.0    memory length: 339725    epsilon: 0.7626712600051522
steps: 325    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1668    score: 11.0    memory length: 340294    epsilon:
0.7621079500051644    steps: 569    lr: 1.6000000000000003e-05    evaluation
reward: 3.78
episode: 1669    score: 6.0    memory length: 340639    epsilon: 0.7617664000051718

```

```

steps: 345    lr: 1.6000000000000003e-05    evaluation reward: 3.8
episode: 1670  score: 1.0    memory length: 340789    epsilon: 0.761617900005175
steps: 150    lr: 1.6000000000000003e-05    evaluation reward: 3.78
episode: 1671  score: 5.0    memory length: 341064    epsilon: 0.761345650005181
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.8
episode: 1672  score: 6.0    memory length: 341390    epsilon: 0.761022910005188
steps: 326    lr: 1.6000000000000003e-05    evaluation reward: 3.82
episode: 1673  score: 7.0    memory length: 341750    epsilon: 0.7606665100051957
steps: 360    lr: 1.6000000000000003e-05    evaluation reward: 3.81
episode: 1674  score: 4.0    memory length: 342029    epsilon: 0.7603903000052017
steps: 279    lr: 1.6000000000000003e-05    evaluation reward: 3.79
episode: 1675  score: 0.0    memory length: 342151    epsilon: 0.7602695200052043
steps: 122    lr: 1.6000000000000003e-05    evaluation reward: 3.76
episode: 1676  score: 4.0    memory length: 342447    epsilon: 0.7599764800052107
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1677  score: 4.0    memory length: 342741    epsilon: 0.759685420005217
steps: 294    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1678  score: 2.0    memory length: 342939    epsilon: 0.7594894000052212
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.69
episode: 1679  score: 1.0    memory length: 343090    epsilon: 0.7593399100052245
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.68
episode: 1680  score: 5.0    memory length: 343435    epsilon: 0.7589983600052319
steps: 345    lr: 1.6000000000000003e-05    evaluation reward: 3.69
episode: 1681  score: 5.0    memory length: 343739    epsilon: 0.7586974000052384
steps: 304    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1682  score: 3.0    memory length: 344007    epsilon: 0.7584320800052442
steps: 268    lr: 1.6000000000000003e-05    evaluation reward: 3.71
episode: 1683  score: 3.0    memory length: 344217    epsilon: 0.7582241800052487
steps: 210    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1684  score: 2.0    memory length: 344415    epsilon: 0.758028160005253
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.74
episode: 1685  score: 3.0    memory length: 344661    epsilon: 0.7577846200052583
steps: 246    lr: 1.6000000000000003e-05    evaluation reward: 3.74
episode: 1686  score: 2.0    memory length: 344861    epsilon: 0.7575866200052626
steps: 200    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1687  score: 4.0    memory length: 345122    epsilon: 0.7573282300052682
steps: 261    lr: 1.6000000000000003e-05    evaluation reward: 3.78
episode: 1688  score: 3.0    memory length: 345350    epsilon: 0.7571025100052731
steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 3.78
episode: 1689  score: 0.0    memory length: 345472    epsilon: 0.7569817300052757
steps: 122    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1690  score: 2.0    memory length: 345670    epsilon: 0.7567857100052799
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1691  score: 2.0    memory length: 345868    epsilon: 0.7565896900052842
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1692  score: 6.0    memory length: 346227    epsilon: 0.7562342800052919
steps: 359    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1693  score: 5.0    memory length: 346498    epsilon: 0.7559659900052977

```



```

steps: 271    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1694  score: 3.0    memory length: 346746    epsilon: 0.7557204700053031
steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1695  score: 1.0    memory length: 346897    epsilon: 0.7555709800053063
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1696  score: 1.0    memory length: 347066    epsilon: 0.75540367000531
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1697  score: 4.0    memory length: 347363    epsilon: 0.7551096400053163
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1698  score: 4.0    memory length: 347620    epsilon: 0.7548552100053219
steps: 257    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1699  score: 2.0    memory length: 347838    epsilon: 0.7546393900053265
steps: 218    lr: 1.6000000000000003e-05    evaluation reward: 3.63
episode: 1700  score: 5.0    memory length: 348145    epsilon: 0.7543354600053331
steps: 307    lr: 1.6000000000000003e-05    evaluation reward: 3.62
episode: 1701  score: 1.0    memory length: 348317    epsilon: 0.7541651800053368
steps: 172    lr: 1.6000000000000003e-05    evaluation reward: 3.59
episode: 1702  score: 0.0    memory length: 348440    epsilon: 0.7540434100053395
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.55
episode: 1703  score: 3.0    memory length: 348688    epsilon: 0.7537978900053448
steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1704  score: 4.0    memory length: 348964    epsilon: 0.7535246500053507
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1705  score: 3.0    memory length: 349209    epsilon: 0.753282100005356
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1706  score: 4.0    memory length: 349483    epsilon: 0.7530108400053619
steps: 274    lr: 1.6000000000000003e-05    evaluation reward: 3.54
episode: 1707  score: 0.0    memory length: 349606    epsilon: 0.7528890700053645
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1708  score: 8.0    memory length: 349929    epsilon: 0.7525693000053715
steps: 323    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1709  score: 5.0    memory length: 350264    epsilon: 0.7522376500053787
steps: 335    lr: 1.6000000000000003e-05    evaluation reward: 3.57
episode: 1710  score: 1.0    memory length: 350414    epsilon: 0.7520891500053819
steps: 150    lr: 1.6000000000000003e-05    evaluation reward: 3.54
episode: 1711  score: 1.0    memory length: 350586    epsilon: 0.7519188700053856
steps: 172    lr: 1.6000000000000003e-05    evaluation reward: 3.48
episode: 1712  score: 6.0    memory length: 350945    epsilon: 0.7515634600053933
steps: 359    lr: 1.6000000000000003e-05    evaluation reward: 3.5
episode: 1713  score: 9.0    memory length: 351416    epsilon: 0.7510971700054034
steps: 471    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1714  score: 1.0    memory length: 351567    epsilon: 0.7509476800054067
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 3.46
episode: 1715  score: 2.0    memory length: 351765    epsilon: 0.7507516600054109
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.4
episode: 1716  score: 5.0    memory length: 352080    epsilon: 0.7504398100054177
steps: 315    lr: 1.6000000000000003e-05    evaluation reward: 3.42
episode: 1717  score: 4.0    memory length: 352358    epsilon: 0.7501645900054237

```

```

steps: 278    lr: 1.6000000000000003e-05    evaluation reward: 3.46
episode: 1718  score: 3.0    memory length: 352606    epsilon: 0.749919070005429
steps: 248    lr: 1.6000000000000003e-05    evaluation reward: 3.43
episode: 1719  score: 5.0    memory length: 352879    epsilon: 0.7496488000054349
steps: 273    lr: 1.6000000000000003e-05    evaluation reward: 3.44
episode: 1720  score: 2.0    memory length: 353059    epsilon: 0.7494706000054387
steps: 180    lr: 1.6000000000000003e-05    evaluation reward: 3.44
episode: 1721  score: 5.0    memory length: 353388    epsilon: 0.7491448900054458
steps: 329    lr: 1.6000000000000003e-05    evaluation reward: 3.46
episode: 1722  score: 2.0    memory length: 353586    epsilon: 0.7489488700054501
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.41
episode: 1723  score: 4.0    memory length: 353884    epsilon: 0.7486538500054565
steps: 298    lr: 1.6000000000000003e-05    evaluation reward: 3.4
episode: 1724  score: 6.0    memory length: 354215    epsilon: 0.7483261600054636
steps: 331    lr: 1.6000000000000003e-05    evaluation reward: 3.46
episode: 1725  score: 2.0    memory length: 354414    epsilon: 0.7481291500054679
steps: 199    lr: 1.6000000000000003e-05    evaluation reward: 3.45
episode: 1726  score: 2.0    memory length: 354612    epsilon: 0.7479331300054721
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 3.4
episode: 1727  score: 4.0    memory length: 354873    epsilon: 0.7476747400054777
steps: 261    lr: 1.6000000000000003e-05    evaluation reward: 3.42
episode: 1728  score: 4.0    memory length: 355131    epsilon: 0.7474193200054833
steps: 258    lr: 1.6000000000000003e-05    evaluation reward: 3.45
episode: 1729  score: 6.0    memory length: 355505    epsilon: 0.7470490600054913
steps: 374    lr: 1.6000000000000003e-05    evaluation reward: 3.49
episode: 1730  score: 5.0    memory length: 355827    epsilon: 0.7467302800054982
steps: 322    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1731  score: 4.0    memory length: 356070    epsilon: 0.7464897100055035
steps: 243    lr: 1.6000000000000003e-05    evaluation reward: 3.52
episode: 1732  score: 4.0    memory length: 356365    epsilon: 0.7461976600055098
steps: 295    lr: 1.6000000000000003e-05    evaluation reward: 3.53
episode: 1733  score: 4.0    memory length: 356625    epsilon: 0.7459402600055154
steps: 260    lr: 1.6000000000000003e-05    evaluation reward: 3.51
episode: 1734  score: 5.0    memory length: 356972    epsilon: 0.7455967300055228
steps: 347    lr: 1.6000000000000003e-05    evaluation reward: 3.53
episode: 1735  score: 8.0    memory length: 357403    epsilon: 0.7451700400055321
steps: 431    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1736  score: 2.0    memory length: 357603    epsilon: 0.7449720400055364
steps: 200    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1737  score: 2.0    memory length: 357805    epsilon: 0.7447720600055407
steps: 202    lr: 1.6000000000000003e-05    evaluation reward: 3.57
episode: 1738  score: 4.0    memory length: 358061    epsilon: 0.7445186200055462
steps: 256    lr: 1.6000000000000003e-05    evaluation reward: 3.6
episode: 1739  score: 3.0    memory length: 358306    epsilon: 0.7442760700055515
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 3.6
episode: 1740  score: 2.0    memory length: 358489    epsilon: 0.7440949000055554
steps: 183    lr: 1.6000000000000003e-05    evaluation reward: 3.56
episode: 1741  score: 1.0    memory length: 358657    epsilon: 0.7439285800055591

```

```

steps: 168    lr: 1.6000000000000003e-05    evaluation reward: 3.54
episode: 1742  score: 2.0    memory length: 358874    epsilon: 0.7437137500055637
steps: 217    lr: 1.6000000000000003e-05    evaluation reward: 3.56
episode: 1743  score: 4.0    memory length: 359172    epsilon: 0.7434187300055701
steps: 298    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1744  score: 4.0    memory length: 359429    epsilon: 0.7431643000055757
steps: 257    lr: 1.6000000000000003e-05    evaluation reward: 3.61
episode: 1745  score: 2.0    memory length: 359645    epsilon: 0.7429504600055803
steps: 216    lr: 1.6000000000000003e-05    evaluation reward: 3.59
episode: 1746  score: 6.0    memory length: 360056    epsilon: 0.7425435700055891
steps: 411    lr: 1.6000000000000003e-05    evaluation reward: 3.61
episode: 1747  score: 4.0    memory length: 360331    epsilon: 0.742271320005595
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.54
episode: 1748  score: 3.0    memory length: 360557    epsilon: 0.7420475800055999
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 3.54
episode: 1749  score: 4.0    memory length: 360850    epsilon: 0.7417575100056062
steps: 293    lr: 1.6000000000000003e-05    evaluation reward: 3.56
episode: 1750  score: 5.0    memory length: 361156    epsilon: 0.7414545700056128
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1751  score: 5.0    memory length: 361445    epsilon: 0.741168460005619
steps: 289    lr: 1.6000000000000003e-05    evaluation reward: 3.6
episode: 1752  score: 4.0    memory length: 361743    epsilon: 0.7408734400056254
steps: 298    lr: 1.6000000000000003e-05    evaluation reward: 3.59
episode: 1753  score: 2.0    memory length: 361923    epsilon: 0.7406952400056293
steps: 180    lr: 1.6000000000000003e-05    evaluation reward: 3.57
episode: 1754  score: 6.0    memory length: 362261    epsilon: 0.7403606200056365
steps: 338    lr: 1.6000000000000003e-05    evaluation reward: 3.63
episode: 1755  score: 7.0    memory length: 362669    epsilon: 0.7399567000056453
steps: 408    lr: 1.6000000000000003e-05    evaluation reward: 3.67
episode: 1756  score: 7.0    memory length: 363116    epsilon: 0.7395141700056549
steps: 447    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1757  score: 6.0    memory length: 363489    epsilon: 0.7391449000056629
steps: 373    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1758  score: 4.0    memory length: 363764    epsilon: 0.7388726500056688
steps: 275    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1759  score: 6.0    memory length: 364160    epsilon: 0.7384806100056773
steps: 396    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1760  score: 4.0    memory length: 364418    epsilon: 0.7382251900056829
steps: 258    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1761  score: 5.0    memory length: 364761    epsilon: 0.7378856200056902
steps: 343    lr: 1.6000000000000003e-05    evaluation reward: 3.74
episode: 1762  score: 2.0    memory length: 364943    epsilon: 0.7377054400056942
steps: 182    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1763  score: 4.0    memory length: 365219    epsilon: 0.7374322000057001
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1764  score: 3.0    memory length: 365448    epsilon: 0.737205490005705
steps: 229    lr: 1.6000000000000003e-05    evaluation reward: 3.77
episode: 1765  score: 1.0    memory length: 365620    epsilon: 0.7370352100057087

```

```

steps: 172    lr: 1.6000000000000003e-05    evaluation reward: 3.72
episode: 1766  score: 6.0    memory length: 365956    epsilon: 0.7367025700057159
steps: 336    lr: 1.6000000000000003e-05    evaluation reward: 3.73
episode: 1767  score: 2.0    memory length: 366175    epsilon: 0.7364857600057206
steps: 219    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1768  score: 5.0    memory length: 366481    epsilon: 0.7361828200057272
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1769  score: 2.0    memory length: 366683    epsilon: 0.7359828400057316
steps: 202    lr: 1.6000000000000003e-05    evaluation reward: 3.6
episode: 1770  score: 4.0    memory length: 366962    epsilon: 0.7357066300057375
steps: 279    lr: 1.6000000000000003e-05    evaluation reward: 3.63
episode: 1771  score: 3.0    memory length: 367206    epsilon: 0.7354650700057428
steps: 244    lr: 1.6000000000000003e-05    evaluation reward: 3.61
episode: 1772  score: 3.0    memory length: 367451    epsilon: 0.7352225200057481
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1773  score: 7.0    memory length: 367836    epsilon: 0.7348413700057563
steps: 385    lr: 1.6000000000000003e-05    evaluation reward: 3.58
episode: 1774  score: 6.0    memory length: 368194    epsilon: 0.734486950005764
steps: 358    lr: 1.6000000000000003e-05    evaluation reward: 3.6
episode: 1775  score: 0.0    memory length: 368317    epsilon: 0.7343651800057667
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.6
episode: 1776  score: 6.0    memory length: 368711    epsilon: 0.7339751200057751
steps: 394    lr: 1.6000000000000003e-05    evaluation reward: 3.62
episode: 1777  score: 6.0    memory length: 369064    epsilon: 0.7336256500057827
steps: 353    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1778  score: 3.0    memory length: 369328    epsilon: 0.7333642900057884
steps: 264    lr: 1.6000000000000003e-05    evaluation reward: 3.65
episode: 1779  score: 3.0    memory length: 369553    epsilon: 0.7331415400057932
steps: 225    lr: 1.6000000000000003e-05    evaluation reward: 3.67
episode: 1780  score: 2.0    memory length: 369768    epsilon: 0.7329286900057979
steps: 215    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1781  score: 5.0    memory length: 370093    epsilon: 0.7326069400058048
steps: 325    lr: 1.6000000000000003e-05    evaluation reward: 3.64
episode: 1782  score: 5.0    memory length: 370419    epsilon: 0.7322842000058118
steps: 326    lr: 1.6000000000000003e-05    evaluation reward: 3.66
episode: 1783  score: 3.0    memory length: 370649    epsilon: 0.7320565000058168
steps: 230    lr: 1.6000000000000003e-05    evaluation reward: 3.66
episode: 1784  score: 4.0    memory length: 370946    epsilon: 0.7317624700058232
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 3.68
episode: 1785  score: 4.0    memory length: 371203    epsilon: 0.7315080400058287
steps: 257    lr: 1.6000000000000003e-05    evaluation reward: 3.69
episode: 1786  score: 3.0    memory length: 371473    epsilon: 0.7312407400058345
steps: 270    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1787  score: 2.0    memory length: 371654    epsilon: 0.7310615500058384
steps: 181    lr: 1.6000000000000003e-05    evaluation reward: 3.68
episode: 1788  score: 3.0    memory length: 371864    epsilon: 0.7308536500058429
steps: 210    lr: 1.6000000000000003e-05    evaluation reward: 3.68
episode: 1789  score: 2.0    memory length: 372061    epsilon: 0.7306586200058471

```

```

steps: 197    lr: 1.6000000000000003e-05    evaluation reward: 3.7
episode: 1790  score: 7.0    memory length: 372472    epsilon: 0.730251730005856
steps: 411    lr: 1.6000000000000003e-05    evaluation reward: 3.75
episode: 1791  score: 4.0    memory length: 372788    epsilon: 0.7299388900058628
steps: 316    lr: 1.6000000000000003e-05    evaluation reward: 3.77
episode: 1792  score: 7.0    memory length: 373218    epsilon: 0.729513190005872
steps: 430    lr: 1.6000000000000003e-05    evaluation reward: 3.78
episode: 1793  score: 6.0    memory length: 373556    epsilon: 0.7291785700058793
steps: 338    lr: 1.6000000000000003e-05    evaluation reward: 3.79
episode: 1794  score: 6.0    memory length: 373895    epsilon: 0.7288429600058866
steps: 339    lr: 1.6000000000000003e-05    evaluation reward: 3.82
episode: 1795  score: 6.0    memory length: 374287    epsilon: 0.728454880005895
steps: 392    lr: 1.6000000000000003e-05    evaluation reward: 3.87
episode: 1796  score: 7.0    memory length: 374692    epsilon: 0.7280539300059037
steps: 405    lr: 1.6000000000000003e-05    evaluation reward: 3.93
episode: 1797  score: 2.0    memory length: 374895    epsilon: 0.727852960005908
steps: 203    lr: 1.6000000000000003e-05    evaluation reward: 3.91
episode: 1798  score: 4.0    memory length: 375193    epsilon: 0.7275579400059144
steps: 298    lr: 1.6000000000000003e-05    evaluation reward: 3.91
episode: 1799  score: 5.0    memory length: 375554    epsilon: 0.7272005500059222
steps: 361    lr: 1.6000000000000003e-05    evaluation reward: 3.94
episode: 1800  score: 0.0    memory length: 375677    epsilon: 0.7270787800059249
steps: 123    lr: 1.6000000000000003e-05    evaluation reward: 3.89
episode: 1801  score: 7.0    memory length: 376081    epsilon: 0.7266788200059335
steps: 404    lr: 1.6000000000000003e-05    evaluation reward: 3.95
episode: 1802  score: 5.0    memory length: 376411    epsilon: 0.7263521200059406
steps: 330    lr: 1.6000000000000003e-05    evaluation reward: 4.0
episode: 1803  score: 3.0    memory length: 376640    epsilon: 0.7261254100059455
steps: 229    lr: 1.6000000000000003e-05    evaluation reward: 4.0
episode: 1804  score: 7.0    memory length: 377063    epsilon: 0.7257066400059546
steps: 423    lr: 1.6000000000000003e-05    evaluation reward: 4.03
episode: 1805  score: 2.0    memory length: 377279    epsilon: 0.7254928000059593
steps: 216    lr: 1.6000000000000003e-05    evaluation reward: 4.02
episode: 1806  score: 2.0    memory length: 377477    epsilon: 0.7252967800059635
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 4.0
episode: 1807  score: 5.0    memory length: 377806    epsilon: 0.7249710700059706
steps: 329    lr: 1.6000000000000003e-05    evaluation reward: 4.05
episode: 1808  score: 7.0    memory length: 378193    epsilon: 0.7245879400059789
steps: 387    lr: 1.6000000000000003e-05    evaluation reward: 4.04
episode: 1809  score: 8.0    memory length: 378589    epsilon: 0.7241959000059874
steps: 396    lr: 1.6000000000000003e-05    evaluation reward: 4.07
episode: 1810  score: 1.0    memory length: 378758    epsilon: 0.7240285900059911
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 4.07
episode: 1811  score: 4.0    memory length: 379054    epsilon: 0.7237355500059974
steps: 296    lr: 1.6000000000000003e-05    evaluation reward: 4.1
episode: 1812  score: 6.0    memory length: 379410    epsilon: 0.7233831100060051
steps: 356    lr: 1.6000000000000003e-05    evaluation reward: 4.1
episode: 1813  score: 2.0    memory length: 379609    epsilon: 0.7231861000060094

```

```

steps: 199    lr: 1.6000000000000003e-05    evaluation reward: 4.03
episode: 1814  score: 5.0    memory length: 379915    epsilon: 0.7228831600060159
steps: 306    lr: 1.6000000000000003e-05    evaluation reward: 4.07
episode: 1815  score: 6.0    memory length: 380270    epsilon: 0.7225317100060236
steps: 355    lr: 1.6000000000000003e-05    evaluation reward: 4.11
episode: 1816  score: 6.0    memory length: 380594    epsilon: 0.7222109500060305
steps: 324    lr: 1.6000000000000003e-05    evaluation reward: 4.12
episode: 1817  score: 1.0    memory length: 380745    epsilon: 0.7220614600060338
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 4.09
episode: 1818  score: 1.0    memory length: 380914    epsilon: 0.7218941500060374
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 4.07
episode: 1819  score: 11.0   memory length: 381365    epsilon:
0.7214476600060471    steps: 451    lr: 1.6000000000000003e-05    evaluation
reward: 4.13
episode: 1820  score: 1.0    memory length: 381516    epsilon: 0.7212981700060503
steps: 151    lr: 1.6000000000000003e-05    evaluation reward: 4.12
episode: 1821  score: 18.0   memory length: 382081    epsilon:
0.7207388200060625    steps: 565    lr: 1.6000000000000003e-05    evaluation
reward: 4.25
episode: 1822  score: 2.0    memory length: 382279    epsilon: 0.7205428000060667
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1823  score: 4.0    memory length: 382539    epsilon: 0.7202854000060723
steps: 260    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1824  score: 3.0    memory length: 382785    epsilon: 0.7200418600060776
steps: 246    lr: 1.6000000000000003e-05    evaluation reward: 4.22
episode: 1825  score: 4.0    memory length: 383044    epsilon: 0.7197854500060832
steps: 259    lr: 1.6000000000000003e-05    evaluation reward: 4.24
episode: 1826  score: 11.0   memory length: 383522    epsilon:
0.7193122300060935    steps: 478    lr: 1.6000000000000003e-05    evaluation
reward: 4.33
episode: 1827  score: 2.0    memory length: 383742    epsilon: 0.7190944300060982
steps: 220    lr: 1.6000000000000003e-05    evaluation reward: 4.31
episode: 1828  score: 1.0    memory length: 383911    epsilon: 0.7189271200061018
steps: 169    lr: 1.6000000000000003e-05    evaluation reward: 4.28
episode: 1829  score: 3.0    memory length: 384156    epsilon: 0.7186845700061071
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1830  score: 7.0    memory length: 384531    epsilon: 0.7183133200061151
steps: 375    lr: 1.6000000000000003e-05    evaluation reward: 4.27
episode: 1831  score: 4.0    memory length: 384792    epsilon: 0.7180549300061208
steps: 261    lr: 1.6000000000000003e-05    evaluation reward: 4.27
episode: 1832  score: 9.0    memory length: 385247    epsilon: 0.7176044800061305
steps: 455    lr: 1.6000000000000003e-05    evaluation reward: 4.32
episode: 1833  score: 4.0    memory length: 385523    epsilon: 0.7173312400061365
steps: 276    lr: 1.6000000000000003e-05    evaluation reward: 4.32
episode: 1834  score: 4.0    memory length: 385785    epsilon: 0.7170718600061421
steps: 262    lr: 1.6000000000000003e-05    evaluation reward: 4.31
episode: 1835  score: 2.0    memory length: 385983    epsilon: 0.7168758400061463
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 4.25

```

episode: 1836 score: 4.0 memory length: 386238 epsilon: 0.7166233900061518
 steps: 255 lr: 1.6000000000000003e-05 evaluation reward: 4.27
 episode: 1837 score: 6.0 memory length: 386592 epsilon: 0.7162729300061594
 steps: 354 lr: 1.6000000000000003e-05 evaluation reward: 4.31
 episode: 1838 score: 2.0 memory length: 386807 epsilon: 0.7160600800061641
 steps: 215 lr: 1.6000000000000003e-05 evaluation reward: 4.29
 episode: 1839 score: 3.0 memory length: 387053 epsilon: 0.7158165400061693
 steps: 246 lr: 1.6000000000000003e-05 evaluation reward: 4.29
 episode: 1840 score: 6.0 memory length: 387407 epsilon: 0.715466080006177
 steps: 354 lr: 1.6000000000000003e-05 evaluation reward: 4.33
 episode: 1841 score: 7.0 memory length: 387790 epsilon: 0.7150869100061852
 steps: 383 lr: 1.6000000000000003e-05 evaluation reward: 4.39
 episode: 1842 score: 1.0 memory length: 387942 epsilon: 0.7149364300061885
 steps: 152 lr: 1.6000000000000003e-05 evaluation reward: 4.38
 episode: 1843 score: 2.0 memory length: 388123 epsilon: 0.7147572400061923
 steps: 181 lr: 1.6000000000000003e-05 evaluation reward: 4.36
 episode: 1844 score: 5.0 memory length: 388411 epsilon: 0.7144721200061985
 steps: 288 lr: 1.6000000000000003e-05 evaluation reward: 4.37
 episode: 1845 score: 5.0 memory length: 388721 epsilon: 0.7141652200062052
 steps: 310 lr: 1.6000000000000003e-05 evaluation reward: 4.4
 episode: 1846 score: 2.0 memory length: 388920 epsilon: 0.7139682100062095
 steps: 199 lr: 1.6000000000000003e-05 evaluation reward: 4.36
 episode: 1847 score: 3.0 memory length: 389130 epsilon: 0.713760310006214
 steps: 210 lr: 1.6000000000000003e-05 evaluation reward: 4.35
 episode: 1848 score: 4.0 memory length: 389405 epsilon: 0.7134880600062199
 steps: 275 lr: 1.6000000000000003e-05 evaluation reward: 4.36
 episode: 1849 score: 5.0 memory length: 389713 epsilon: 0.7131831400062265
 steps: 308 lr: 1.6000000000000003e-05 evaluation reward: 4.37
 episode: 1850 score: 1.0 memory length: 389882 epsilon: 0.7130158300062301
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 4.33
 episode: 1851 score: 1.0 memory length: 390034 epsilon: 0.7128653500062334
 steps: 152 lr: 1.6000000000000003e-05 evaluation reward: 4.29
 episode: 1852 score: 3.0 memory length: 390262 epsilon: 0.7126396300062383
 steps: 228 lr: 1.6000000000000003e-05 evaluation reward: 4.28
 episode: 1853 score: 12.0 memory length: 390657 epsilon:
 0.7122485800062468 steps: 395 lr: 1.6000000000000003e-05 evaluation
 reward: 4.38
 episode: 1854 score: 7.0 memory length: 391039 epsilon: 0.711870400006255
 steps: 382 lr: 1.6000000000000003e-05 evaluation reward: 4.39
 episode: 1855 score: 5.0 memory length: 391367 epsilon: 0.7115456800062621
 steps: 328 lr: 1.6000000000000003e-05 evaluation reward: 4.37
 episode: 1856 score: 4.0 memory length: 391646 epsilon: 0.7112694700062681
 steps: 279 lr: 1.6000000000000003e-05 evaluation reward: 4.34
 episode: 1857 score: 1.0 memory length: 391797 epsilon: 0.7111199800062713
 steps: 151 lr: 1.6000000000000003e-05 evaluation reward: 4.29
 episode: 1858 score: 4.0 memory length: 392075 epsilon: 0.7108447600062773
 steps: 278 lr: 1.6000000000000003e-05 evaluation reward: 4.29
 episode: 1859 score: 3.0 memory length: 392303 epsilon: 0.7106190400062822

```

steps: 228    lr: 1.6000000000000003e-05    evaluation reward: 4.26
episode: 1860  score: 6.0    memory length: 392640    epsilon: 0.7102854100062894
steps: 337    lr: 1.6000000000000003e-05    evaluation reward: 4.28
episode: 1861  score: 2.0    memory length: 392860    epsilon: 0.7100676100062941
steps: 220    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1862  score: 3.0    memory length: 393073    epsilon: 0.7098567400062987
steps: 213    lr: 1.6000000000000003e-05    evaluation reward: 4.26
episode: 1863  score: 2.0    memory length: 393271    epsilon: 0.709660720006303
steps: 198    lr: 1.6000000000000003e-05    evaluation reward: 4.24
episode: 1864  score: 5.0    memory length: 393580    epsilon: 0.7093548100063096
steps: 309    lr: 1.6000000000000003e-05    evaluation reward: 4.26
episode: 1865  score: 3.0    memory length: 393792    epsilon: 0.7091449300063142
steps: 212    lr: 1.6000000000000003e-05    evaluation reward: 4.28
episode: 1866  score: 4.0    memory length: 394052    epsilon: 0.7088875300063198
steps: 260    lr: 1.6000000000000003e-05    evaluation reward: 4.26
episode: 1867  score: 4.0    memory length: 394325    epsilon: 0.7086172600063256
steps: 273    lr: 1.6000000000000003e-05    evaluation reward: 4.28
episode: 1868  score: 1.0    memory length: 394477    epsilon: 0.7084667800063289
steps: 152    lr: 1.6000000000000003e-05    evaluation reward: 4.24
episode: 1869  score: 3.0    memory length: 394704    epsilon: 0.7082420500063338
steps: 227    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1870  score: 4.0    memory length: 394982    epsilon: 0.7079668300063398
steps: 278    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1871  score: 7.0    memory length: 395283    epsilon: 0.7076688400063462
steps: 301    lr: 1.6000000000000003e-05    evaluation reward: 4.29
episode: 1872  score: 3.0    memory length: 395528    epsilon: 0.7074262900063515
steps: 245    lr: 1.6000000000000003e-05    evaluation reward: 4.29
episode: 1873  score: 2.0    memory length: 395748    epsilon: 0.7072084900063562
steps: 220    lr: 1.6000000000000003e-05    evaluation reward: 4.24
episode: 1874  score: 3.0    memory length: 395994    epsilon: 0.7069649500063615
steps: 246    lr: 1.6000000000000003e-05    evaluation reward: 4.21
episode: 1875  score: 4.0    memory length: 396250    epsilon: 0.706711510006367
steps: 256    lr: 1.6000000000000003e-05    evaluation reward: 4.25
episode: 1876  score: 5.0    memory length: 396547    epsilon: 0.7064174800063734
steps: 297    lr: 1.6000000000000003e-05    evaluation reward: 4.24
episode: 1877  score: 3.0    memory length: 396773    epsilon: 0.7061937400063782
steps: 226    lr: 1.6000000000000003e-05    evaluation reward: 4.21
episode: 1878  score: 4.0    memory length: 397090    epsilon: 0.7058799100063851
steps: 317    lr: 1.6000000000000003e-05    evaluation reward: 4.22
episode: 1879  score: 4.0    memory length: 397353    epsilon: 0.7056195400063907
steps: 263    lr: 1.6000000000000003e-05    evaluation reward: 4.23
episode: 1880  score: 3.0    memory length: 397563    epsilon: 0.7054116400063952
steps: 210    lr: 1.6000000000000003e-05    evaluation reward: 4.24
episode: 1881  score: 8.0    memory length: 398007    epsilon: 0.7049720800064048
steps: 444    lr: 1.6000000000000003e-05    evaluation reward: 4.27
episode: 1882  score: 12.0   memory length: 398619    epsilon:
0.7043662000064179    steps: 612    lr: 1.6000000000000003e-05    evaluation
reward: 4.34

```


episode: 1883 score: 3.0 memory length: 398869 epsilon: 0.7041187000064233
 steps: 250 lr: 1.6000000000000003e-05 evaluation reward: 4.34
 episode: 1884 score: 4.0 memory length: 399146 epsilon: 0.7038444700064292
 steps: 277 lr: 1.6000000000000003e-05 evaluation reward: 4.34
 episode: 1885 score: 1.0 memory length: 399315 epsilon: 0.7036771600064329
 steps: 169 lr: 1.6000000000000003e-05 evaluation reward: 4.31
 episode: 1886 score: 5.0 memory length: 399642 epsilon: 0.7033534300064399
 steps: 327 lr: 1.6000000000000003e-05 evaluation reward: 4.33
 episode: 1887 score: 0.0 memory length: 399764 epsilon: 0.7032326500064425
 steps: 122 lr: 1.6000000000000003e-05 evaluation reward: 4.31
 episode: 1888 score: 7.0 memory length: 400152 epsilon: 0.7028485300064509
 steps: 388 lr: 6.400000000000001e-06 evaluation reward: 4.35
 episode: 1889 score: 4.0 memory length: 400430 epsilon: 0.7025733100064568
 steps: 278 lr: 6.400000000000001e-06 evaluation reward: 4.37
 episode: 1890 score: 6.0 memory length: 400801 epsilon: 0.7022060200064648
 steps: 371 lr: 6.400000000000001e-06 evaluation reward: 4.36
 episode: 1891 score: 3.0 memory length: 401069 epsilon: 0.7019407000064706
 steps: 268 lr: 6.400000000000001e-06 evaluation reward: 4.35
 episode: 1892 score: 3.0 memory length: 401320 epsilon: 0.701692210006476
 steps: 251 lr: 6.400000000000001e-06 evaluation reward: 4.31
 episode: 1893 score: 6.0 memory length: 401659 epsilon: 0.7013566000064833
 steps: 339 lr: 6.400000000000001e-06 evaluation reward: 4.31
 episode: 1894 score: 4.0 memory length: 401933 epsilon: 0.7010853400064891
 steps: 274 lr: 6.400000000000001e-06 evaluation reward: 4.29
 episode: 1895 score: 12.0 memory length: 402532 epsilon: 0.700492330006502
 steps: 599 lr: 6.400000000000001e-06 evaluation reward: 4.35
 episode: 1896 score: 3.0 memory length: 402797 epsilon: 0.7002299800065077
 steps: 265 lr: 6.400000000000001e-06 evaluation reward: 4.31
 episode: 1897 score: 3.0 memory length: 403026 epsilon: 0.7000032700065126
 steps: 229 lr: 6.400000000000001e-06 evaluation reward: 4.32
 episode: 1898 score: 5.0 memory length: 403350 epsilon: 0.6996825100065196
 steps: 324 lr: 6.400000000000001e-06 evaluation reward: 4.33
 episode: 1899 score: 3.0 memory length: 403578 epsilon: 0.6994567900065245
 steps: 228 lr: 6.400000000000001e-06 evaluation reward: 4.31
 episode: 1900 score: 4.0 memory length: 403850 epsilon: 0.6991875100065303
 steps: 272 lr: 6.400000000000001e-06 evaluation reward: 4.35
 episode: 1901 score: 1.0 memory length: 404002 epsilon: 0.6990370300065336
 steps: 152 lr: 6.400000000000001e-06 evaluation reward: 4.29
 episode: 1902 score: 5.0 memory length: 404301 epsilon: 0.69874102000654
 steps: 299 lr: 6.400000000000001e-06 evaluation reward: 4.29
 episode: 1903 score: 7.0 memory length: 404748 epsilon: 0.6982984900065496
 steps: 447 lr: 6.400000000000001e-06 evaluation reward: 4.33
 episode: 1904 score: 12.0 memory length: 405209 epsilon:
 0.6978421000065596 steps: 461 lr: 6.400000000000001e-06 evaluation
 reward: 4.38
 episode: 1905 score: 3.0 memory length: 405458 epsilon: 0.6975955900065649
 steps: 249 lr: 6.400000000000001e-06 evaluation reward: 4.39
 episode: 1906 score: 8.0 memory length: 405930 epsilon: 0.697128310006575

```

steps: 472    lr: 6.4000000000000001e-06    evaluation reward: 4.45
episode: 1907  score: 3.0    memory length: 406196    epsilon: 0.6968649700065808
steps: 266    lr: 6.4000000000000001e-06    evaluation reward: 4.43
episode: 1908  score: 3.0    memory length: 406443    epsilon: 0.6966204400065861
steps: 247    lr: 6.4000000000000001e-06    evaluation reward: 4.39
episode: 1909  score: 6.0    memory length: 406806    epsilon: 0.6962610700065939
steps: 363    lr: 6.4000000000000001e-06    evaluation reward: 4.37
episode: 1910  score: 7.0    memory length: 407211    epsilon: 0.6958601200066026
steps: 405    lr: 6.4000000000000001e-06    evaluation reward: 4.43
episode: 1911  score: 6.0    memory length: 407555    epsilon: 0.69551956000661
steps: 344    lr: 6.4000000000000001e-06    evaluation reward: 4.45
episode: 1912  score: 4.0    memory length: 407833    epsilon: 0.695244340006616
steps: 278    lr: 6.4000000000000001e-06    evaluation reward: 4.43
episode: 1913  score: 9.0    memory length: 408286    epsilon: 0.6947958700066257
steps: 453    lr: 6.4000000000000001e-06    evaluation reward: 4.5
episode: 1914  score: 12.0   memory length: 408777    epsilon:
0.6943097800066362    steps: 491    lr: 6.4000000000000001e-06    evaluation
reward: 4.57
episode: 1915  score: 5.0    memory length: 409103    epsilon: 0.6939870400066432
steps: 326    lr: 6.4000000000000001e-06    evaluation reward: 4.56
episode: 1916  score: 8.0    memory length: 409555    epsilon: 0.693539560006653
steps: 452    lr: 6.4000000000000001e-06    evaluation reward: 4.58
episode: 1917  score: 4.0    memory length: 409797    epsilon: 0.6932999800066582
steps: 242    lr: 6.4000000000000001e-06    evaluation reward: 4.61
episode: 1918  score: 4.0    memory length: 410073    epsilon: 0.6930267400066641
steps: 276    lr: 6.4000000000000001e-06    evaluation reward: 4.64
episode: 1919  score: 4.0    memory length: 410371    epsilon: 0.6927317200066705
steps: 298    lr: 6.4000000000000001e-06    evaluation reward: 4.57
episode: 1920  score: 8.0    memory length: 410796    epsilon: 0.6923109700066796
steps: 425    lr: 6.4000000000000001e-06    evaluation reward: 4.64
episode: 1921  score: 3.0    memory length: 411041    epsilon: 0.6920684200066849
steps: 245    lr: 6.4000000000000001e-06    evaluation reward: 4.49
episode: 1922  score: 13.0   memory length: 411474    epsilon:
0.6916397500066942    steps: 433    lr: 6.4000000000000001e-06    evaluation
reward: 4.6
episode: 1923  score: 2.0    memory length: 411672    epsilon: 0.6914437300066985
steps: 198    lr: 6.4000000000000001e-06    evaluation reward: 4.58
episode: 1924  score: 6.0    memory length: 412024    epsilon: 0.691095250006706
steps: 352    lr: 6.4000000000000001e-06    evaluation reward: 4.61
episode: 1925  score: 3.0    memory length: 412255    epsilon: 0.690866560006711
steps: 231    lr: 6.4000000000000001e-06    evaluation reward: 4.6
episode: 1926  score: 8.0    memory length: 412728    epsilon: 0.6903982900067211
steps: 473    lr: 6.4000000000000001e-06    evaluation reward: 4.57
episode: 1927  score: 2.0    memory length: 412946    epsilon: 0.6901824700067258
steps: 218    lr: 6.4000000000000001e-06    evaluation reward: 4.57
episode: 1928  score: 4.0    memory length: 413188    epsilon: 0.689942890006731
steps: 242    lr: 6.4000000000000001e-06    evaluation reward: 4.6
episode: 1929  score: 4.0    memory length: 413447    epsilon: 0.6896864800067366

```

steps: 259 lr: 6.4000000000000001e-06 evaluation reward: 4.61
 episode: 1930 score: 7.0 memory length: 413869 epsilon: 0.6892687000067457
 steps: 422 lr: 6.4000000000000001e-06 evaluation reward: 4.61
 episode: 1931 score: 7.0 memory length: 414263 epsilon: 0.6888786400067541
 steps: 394 lr: 6.4000000000000001e-06 evaluation reward: 4.64
 episode: 1932 score: 4.0 memory length: 414535 epsilon: 0.68860936000676
 steps: 272 lr: 6.4000000000000001e-06 evaluation reward: 4.59
 episode: 1933 score: 2.0 memory length: 414752 epsilon: 0.6883945300067646
 steps: 217 lr: 6.4000000000000001e-06 evaluation reward: 4.57
 episode: 1934 score: 3.0 memory length: 415001 epsilon: 0.68814802000677
 steps: 249 lr: 6.4000000000000001e-06 evaluation reward: 4.56
 episode: 1935 score: 5.0 memory length: 415308 epsilon: 0.6878440900067766
 steps: 307 lr: 6.4000000000000001e-06 evaluation reward: 4.59
 episode: 1936 score: 2.0 memory length: 415506 epsilon: 0.6876480700067809
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 4.57
 episode: 1937 score: 6.0 memory length: 415870 epsilon: 0.6872877100067887
 steps: 364 lr: 6.4000000000000001e-06 evaluation reward: 4.57
 episode: 1938 score: 6.0 memory length: 416217 epsilon: 0.6869441800067961
 steps: 347 lr: 6.4000000000000001e-06 evaluation reward: 4.61
 episode: 1939 score: 2.0 memory length: 416435 epsilon: 0.6867283600068008
 steps: 218 lr: 6.4000000000000001e-06 evaluation reward: 4.6
 episode: 1940 score: 2.0 memory length: 416633 epsilon: 0.6865323400068051
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 4.56
 episode: 1941 score: 4.0 memory length: 416931 epsilon: 0.6862373200068115
 steps: 298 lr: 6.4000000000000001e-06 evaluation reward: 4.53
 episode: 1942 score: 6.0 memory length: 417262 epsilon: 0.6859096300068186
 steps: 331 lr: 6.4000000000000001e-06 evaluation reward: 4.58
 episode: 1943 score: 6.0 memory length: 417617 epsilon: 0.6855581800068262
 steps: 355 lr: 6.4000000000000001e-06 evaluation reward: 4.62
 episode: 1944 score: 5.0 memory length: 417978 epsilon: 0.685200790006834
 steps: 361 lr: 6.4000000000000001e-06 evaluation reward: 4.62
 episode: 1945 score: 2.0 memory length: 418177 epsilon: 0.6850037800068383
 steps: 199 lr: 6.4000000000000001e-06 evaluation reward: 4.59
 episode: 1946 score: 5.0 memory length: 418519 epsilon: 0.6846652000068456
 steps: 342 lr: 6.4000000000000001e-06 evaluation reward: 4.62
 episode: 1947 score: 6.0 memory length: 418885 epsilon: 0.6843028600068535
 steps: 366 lr: 6.4000000000000001e-06 evaluation reward: 4.65
 episode: 1948 score: 9.0 memory length: 419332 epsilon: 0.6838603300068631
 steps: 447 lr: 6.4000000000000001e-06 evaluation reward: 4.7
 episode: 1949 score: 13.0 memory length: 419933 epsilon: 0.683265340006876
 steps: 601 lr: 6.4000000000000001e-06 evaluation reward: 4.78
 episode: 1950 score: 2.0 memory length: 420131 epsilon: 0.6830693200068803
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 4.79
 episode: 1951 score: 7.0 memory length: 420501 epsilon: 0.6827030200068882
 steps: 370 lr: 6.4000000000000001e-06 evaluation reward: 4.85
 episode: 1952 score: 7.0 memory length: 420890 epsilon: 0.6823179100068966
 steps: 389 lr: 6.4000000000000001e-06 evaluation reward: 4.89
 episode: 1953 score: 7.0 memory length: 421253 epsilon: 0.6819585400069044

```

steps: 363    lr: 6.4000000000000001e-06    evaluation reward: 4.84
episode: 1954  score: 5.0    memory length: 421561    epsilon: 0.681653620006911
steps: 308    lr: 6.4000000000000001e-06    evaluation reward: 4.82
episode: 1955  score: 3.0    memory length: 421826    epsilon: 0.6813912700069167
steps: 265    lr: 6.4000000000000001e-06    evaluation reward: 4.8
episode: 1956  score: 5.0    memory length: 422135    epsilon: 0.6810853600069233
steps: 309    lr: 6.4000000000000001e-06    evaluation reward: 4.81
episode: 1957  score: 4.0    memory length: 422388    epsilon: 0.6808348900069288
steps: 253    lr: 6.4000000000000001e-06    evaluation reward: 4.84
episode: 1958  score: 6.0    memory length: 422743    epsilon: 0.6804834400069364
steps: 355    lr: 6.4000000000000001e-06    evaluation reward: 4.86
episode: 1959  score: 7.0    memory length: 423119    epsilon: 0.6801112000069445
steps: 376    lr: 6.4000000000000001e-06    evaluation reward: 4.9
episode: 1960  score: 4.0    memory length: 423396    epsilon: 0.6798369700069504
steps: 277    lr: 6.4000000000000001e-06    evaluation reward: 4.88
episode: 1961  score: 6.0    memory length: 423770    epsilon: 0.6794667100069585
steps: 374    lr: 6.4000000000000001e-06    evaluation reward: 4.92
episode: 1962  score: 5.0    memory length: 424078    epsilon: 0.6791617900069651
steps: 308    lr: 6.4000000000000001e-06    evaluation reward: 4.94
episode: 1963  score: 2.0    memory length: 424297    epsilon: 0.6789449800069698
steps: 219    lr: 6.4000000000000001e-06    evaluation reward: 4.94
episode: 1964  score: 7.0    memory length: 424703    epsilon: 0.6785430400069785
steps: 406    lr: 6.4000000000000001e-06    evaluation reward: 4.96
episode: 1965  score: 1.0    memory length: 424854    epsilon: 0.6783935500069818
steps: 151    lr: 6.4000000000000001e-06    evaluation reward: 4.94
episode: 1966  score: 8.0    memory length: 425329    epsilon: 0.677923300006992
steps: 475    lr: 6.4000000000000001e-06    evaluation reward: 4.98
episode: 1967  score: 3.0    memory length: 425574    epsilon: 0.6776807500069972
steps: 245    lr: 6.4000000000000001e-06    evaluation reward: 4.97
episode: 1968  score: 2.0    memory length: 425772    epsilon: 0.6774847300070015
steps: 198    lr: 6.4000000000000001e-06    evaluation reward: 4.98
episode: 1969  score: 3.0    memory length: 425998    epsilon: 0.6772609900070063
steps: 226    lr: 6.4000000000000001e-06    evaluation reward: 4.98
episode: 1970  score: 4.0    memory length: 426293    epsilon: 0.6769689400070127
steps: 295    lr: 6.4000000000000001e-06    evaluation reward: 4.98
episode: 1971  score: 3.0    memory length: 426504    epsilon: 0.6767600500070172
steps: 211    lr: 6.4000000000000001e-06    evaluation reward: 4.94
episode: 1972  score: 4.0    memory length: 426802    epsilon: 0.6764650300070236
steps: 298    lr: 6.4000000000000001e-06    evaluation reward: 4.95
episode: 1973  score: 6.0    memory length: 427158    epsilon: 0.6761125900070313
steps: 356    lr: 6.4000000000000001e-06    evaluation reward: 4.99
episode: 1974  score: 3.0    memory length: 427427    epsilon: 0.6758462800070371
steps: 269    lr: 6.4000000000000001e-06    evaluation reward: 4.99
episode: 1975  score: 10.0   memory length: 427908    epsilon:
0.6753700900070474    steps: 481    lr: 6.4000000000000001e-06    evaluation
reward: 5.05
episode: 1976  score: 6.0    memory length: 428267    epsilon: 0.6750146800070551
steps: 359    lr: 6.4000000000000001e-06    evaluation reward: 5.06

```

episode: 1977 score: 7.0 memory length: 428647 epsilon: 0.6746384800070633
 steps: 380 lr: 6.4000000000000001e-06 evaluation reward: 5.1
 episode: 1978 score: 6.0 memory length: 429011 epsilon: 0.6742781200070711
 steps: 364 lr: 6.4000000000000001e-06 evaluation reward: 5.12
 episode: 1979 score: 6.0 memory length: 429368 epsilon: 0.6739246900070788
 steps: 357 lr: 6.4000000000000001e-06 evaluation reward: 5.14
 episode: 1980 score: 1.0 memory length: 429518 epsilon: 0.673776190007082
 steps: 150 lr: 6.4000000000000001e-06 evaluation reward: 5.12
 episode: 1981 score: 2.0 memory length: 429716 epsilon: 0.6735801700070863
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 5.06
 episode: 1982 score: 5.0 memory length: 430042 epsilon: 0.6732574300070933
 steps: 326 lr: 6.4000000000000001e-06 evaluation reward: 4.99
 episode: 1983 score: 6.0 memory length: 430407 epsilon: 0.6728960800071011
 steps: 365 lr: 6.4000000000000001e-06 evaluation reward: 5.02
 episode: 1984 score: 3.0 memory length: 430633 epsilon: 0.672672340007106
 steps: 226 lr: 6.4000000000000001e-06 evaluation reward: 5.01
 episode: 1985 score: 4.0 memory length: 430913 epsilon: 0.672395140007112
 steps: 280 lr: 6.4000000000000001e-06 evaluation reward: 5.04
 episode: 1986 score: 9.0 memory length: 431388 epsilon: 0.6719248900071222
 steps: 475 lr: 6.4000000000000001e-06 evaluation reward: 5.08
 episode: 1987 score: 6.0 memory length: 431709 epsilon: 0.6716071000071291
 steps: 321 lr: 6.4000000000000001e-06 evaluation reward: 5.14
 episode: 1988 score: 9.0 memory length: 432144 epsilon: 0.6711764500071384
 steps: 435 lr: 6.4000000000000001e-06 evaluation reward: 5.16
 episode: 1989 score: 3.0 memory length: 432354 epsilon: 0.670968550007143
 steps: 210 lr: 6.4000000000000001e-06 evaluation reward: 5.15
 episode: 1990 score: 8.0 memory length: 432781 epsilon: 0.6705458200071521
 steps: 427 lr: 6.4000000000000001e-06 evaluation reward: 5.17
 episode: 1991 score: 9.0 memory length: 433284 epsilon: 0.6700478500071629
 steps: 503 lr: 6.4000000000000001e-06 evaluation reward: 5.23
 episode: 1992 score: 4.0 memory length: 433559 epsilon: 0.6697756000071688
 steps: 275 lr: 6.4000000000000001e-06 evaluation reward: 5.24
 episode: 1993 score: 8.0 memory length: 434015 epsilon: 0.6693241600071786
 steps: 456 lr: 6.4000000000000001e-06 evaluation reward: 5.26
 episode: 1994 score: 4.0 memory length: 434309 epsilon: 0.669033100007185
 steps: 294 lr: 6.4000000000000001e-06 evaluation reward: 5.26
 episode: 1995 score: 3.0 memory length: 434535 epsilon: 0.6688093600071898
 steps: 226 lr: 6.4000000000000001e-06 evaluation reward: 5.17
 episode: 1996 score: 3.0 memory length: 434746 epsilon: 0.6686004700071944
 steps: 211 lr: 6.4000000000000001e-06 evaluation reward: 5.17
 episode: 1997 score: 4.0 memory length: 435004 epsilon: 0.6683450500071999
 steps: 258 lr: 6.4000000000000001e-06 evaluation reward: 5.18
 episode: 1998 score: 3.0 memory length: 435231 epsilon: 0.6681203200072048
 steps: 227 lr: 6.4000000000000001e-06 evaluation reward: 5.16
 episode: 1999 score: 7.0 memory length: 435636 epsilon: 0.6677193700072135
 steps: 405 lr: 6.4000000000000001e-06 evaluation reward: 5.2
 episode: 2000 score: 1.0 memory length: 435787 epsilon: 0.6675698800072167
 steps: 151 lr: 6.4000000000000001e-06 evaluation reward: 5.17

episode: 2001 score: 5.0 memory length: 436096 epsilon: 0.6672639700072234
 steps: 309 lr: 6.4000000000000001e-06 evaluation reward: 5.21
 episode: 2002 score: 4.0 memory length: 436340 epsilon: 0.6670224100072286
 steps: 244 lr: 6.4000000000000001e-06 evaluation reward: 5.2
 episode: 2003 score: 9.0 memory length: 436840 epsilon: 0.6665274100072394
 steps: 500 lr: 6.4000000000000001e-06 evaluation reward: 5.22
 episode: 2004 score: 4.0 memory length: 437101 epsilon: 0.666269020007245
 steps: 261 lr: 6.4000000000000001e-06 evaluation reward: 5.14
 episode: 2005 score: 5.0 memory length: 437428 epsilon: 0.665945290007252
 steps: 327 lr: 6.4000000000000001e-06 evaluation reward: 5.16
 episode: 2006 score: 5.0 memory length: 437743 epsilon: 0.6656334400072588
 steps: 315 lr: 6.4000000000000001e-06 evaluation reward: 5.13
 episode: 2007 score: 7.0 memory length: 438187 epsilon: 0.6651938800072683
 steps: 444 lr: 6.4000000000000001e-06 evaluation reward: 5.17
 episode: 2008 score: 3.0 memory length: 438416 epsilon: 0.6649671700072732
 steps: 229 lr: 6.4000000000000001e-06 evaluation reward: 5.17
 episode: 2009 score: 3.0 memory length: 438646 epsilon: 0.6647394700072782
 steps: 230 lr: 6.4000000000000001e-06 evaluation reward: 5.14
 episode: 2010 score: 4.0 memory length: 438962 epsilon: 0.664426630007285
 steps: 316 lr: 6.4000000000000001e-06 evaluation reward: 5.11
 episode: 2011 score: 7.0 memory length: 439382 epsilon: 0.664010830007294
 steps: 420 lr: 6.4000000000000001e-06 evaluation reward: 5.12
 episode: 2012 score: 9.0 memory length: 439851 epsilon: 0.6635465200073041
 steps: 469 lr: 6.4000000000000001e-06 evaluation reward: 5.17
 episode: 2013 score: 7.0 memory length: 440216 epsilon: 0.6631851700073119
 steps: 365 lr: 6.4000000000000001e-06 evaluation reward: 5.15
 episode: 2014 score: 5.0 memory length: 440523 epsilon: 0.6628812400073185
 steps: 307 lr: 6.4000000000000001e-06 evaluation reward: 5.08
 episode: 2015 score: 6.0 memory length: 440913 epsilon: 0.6624951400073269
 steps: 390 lr: 6.4000000000000001e-06 evaluation reward: 5.09
 episode: 2016 score: 5.0 memory length: 441228 epsilon: 0.6621832900073337
 steps: 315 lr: 6.4000000000000001e-06 evaluation reward: 5.06
 episode: 2017 score: 7.0 memory length: 441628 epsilon: 0.6617872900073423
 steps: 400 lr: 6.4000000000000001e-06 evaluation reward: 5.09
 episode: 2018 score: 6.0 memory length: 441991 epsilon: 0.6614279200073501
 steps: 363 lr: 6.4000000000000001e-06 evaluation reward: 5.11
 episode: 2019 score: 2.0 memory length: 442171 epsilon: 0.6612497200073539
 steps: 180 lr: 6.4000000000000001e-06 evaluation reward: 5.09
 episode: 2020 score: 4.0 memory length: 442428 epsilon: 0.6609952900073595
 steps: 257 lr: 6.4000000000000001e-06 evaluation reward: 5.05
 episode: 2021 score: 5.0 memory length: 442773 epsilon: 0.6606537400073669
 steps: 345 lr: 6.4000000000000001e-06 evaluation reward: 5.07
 episode: 2022 score: 7.0 memory length: 443180 epsilon: 0.6602508100073756
 steps: 407 lr: 6.4000000000000001e-06 evaluation reward: 5.01
 episode: 2023 score: 2.0 memory length: 443378 epsilon: 0.6600547900073799
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 5.01
 episode: 2024 score: 5.0 memory length: 443683 epsilon: 0.6597528400073864
 steps: 305 lr: 6.4000000000000001e-06 evaluation reward: 5.0

episode: 2025 score: 8.0 memory length: 444116 epsilon: 0.6593241700073957
 steps: 433 lr: 6.4000000000000001e-06 evaluation reward: 5.05
 episode: 2026 score: 6.0 memory length: 444469 epsilon: 0.6589747000074033
 steps: 353 lr: 6.4000000000000001e-06 evaluation reward: 5.03
 episode: 2027 score: 7.0 memory length: 444864 epsilon: 0.6585836500074118
 steps: 395 lr: 6.4000000000000001e-06 evaluation reward: 5.08
 episode: 2028 score: 3.0 memory length: 445076 epsilon: 0.6583737700074164
 steps: 212 lr: 6.4000000000000001e-06 evaluation reward: 5.07
 episode: 2029 score: 4.0 memory length: 445342 epsilon: 0.6581104300074221
 steps: 266 lr: 6.4000000000000001e-06 evaluation reward: 5.07
 episode: 2030 score: 5.0 memory length: 445650 epsilon: 0.6578055100074287
 steps: 308 lr: 6.4000000000000001e-06 evaluation reward: 5.05
 episode: 2031 score: 6.0 memory length: 445990 epsilon: 0.657468910007436
 steps: 340 lr: 6.4000000000000001e-06 evaluation reward: 5.04
 episode: 2032 score: 12.0 memory length: 446451 epsilon:
 0.6570125200074459 steps: 461 lr: 6.4000000000000001e-06 evaluation
 reward: 5.12
 episode: 2033 score: 6.0 memory length: 446791 epsilon: 0.6566759200074532
 steps: 340 lr: 6.4000000000000001e-06 evaluation reward: 5.16
 episode: 2034 score: 3.0 memory length: 447043 epsilon: 0.6564264400074586
 steps: 252 lr: 6.4000000000000001e-06 evaluation reward: 5.16
 episode: 2035 score: 9.0 memory length: 447546 epsilon: 0.6559284700074695
 steps: 503 lr: 6.4000000000000001e-06 evaluation reward: 5.2
 episode: 2036 score: 7.0 memory length: 447951 epsilon: 0.6555275200074782
 steps: 405 lr: 6.4000000000000001e-06 evaluation reward: 5.25
 episode: 2037 score: 5.0 memory length: 448258 epsilon: 0.6552235900074848
 steps: 307 lr: 6.4000000000000001e-06 evaluation reward: 5.24
 episode: 2038 score: 7.0 memory length: 448698 epsilon: 0.6547879900074942
 steps: 440 lr: 6.4000000000000001e-06 evaluation reward: 5.25
 episode: 2039 score: 4.0 memory length: 448954 epsilon: 0.6545345500074997
 steps: 256 lr: 6.4000000000000001e-06 evaluation reward: 5.27
 episode: 2040 score: 3.0 memory length: 449180 epsilon: 0.6543108100075046
 steps: 226 lr: 6.4000000000000001e-06 evaluation reward: 5.28
 episode: 2041 score: 16.0 memory length: 449755 epsilon:
 0.6537415600075169 steps: 575 lr: 6.4000000000000001e-06 evaluation
 reward: 5.4
 episode: 2042 score: 4.0 memory length: 450029 epsilon: 0.6534703000075228
 steps: 274 lr: 6.4000000000000001e-06 evaluation reward: 5.38
 episode: 2043 score: 14.0 memory length: 450683 epsilon:
 0.6528228400075369 steps: 654 lr: 6.4000000000000001e-06 evaluation
 reward: 5.46
 episode: 2044 score: 4.0 memory length: 450924 epsilon: 0.652584250007542
 steps: 241 lr: 6.4000000000000001e-06 evaluation reward: 5.45
 episode: 2045 score: 3.0 memory length: 451191 epsilon: 0.6523199200075478
 steps: 267 lr: 6.4000000000000001e-06 evaluation reward: 5.46
 episode: 2046 score: 6.0 memory length: 451528 epsilon: 0.651986290007555
 steps: 337 lr: 6.4000000000000001e-06 evaluation reward: 5.47
 episode: 2047 score: 2.0 memory length: 451726 epsilon: 0.6517902700075593

```

steps: 198    lr: 6.4000000000000001e-06    evaluation reward: 5.43
episode: 2048  score: 4.0    memory length: 452010    epsilon: 0.6515091100075654
steps: 284    lr: 6.4000000000000001e-06    evaluation reward: 5.38
episode: 2049  score: 5.0    memory length: 452318    epsilon: 0.651204190007572
steps: 308    lr: 6.4000000000000001e-06    evaluation reward: 5.3
episode: 2050  score: 7.0    memory length: 452706    epsilon: 0.6508200700075804
steps: 388    lr: 6.4000000000000001e-06    evaluation reward: 5.35
episode: 2051  score: 4.0    memory length: 452967    epsilon: 0.650561680007586
steps: 261    lr: 6.4000000000000001e-06    evaluation reward: 5.32
episode: 2052  score: 5.0    memory length: 453274    epsilon: 0.6502577500075926
steps: 307    lr: 6.4000000000000001e-06    evaluation reward: 5.3
episode: 2053  score: 2.0    memory length: 453475    epsilon: 0.6500587600075969
steps: 201    lr: 6.4000000000000001e-06    evaluation reward: 5.25
episode: 2054  score: 6.0    memory length: 453796    epsilon: 0.6497409700076038
steps: 321    lr: 6.4000000000000001e-06    evaluation reward: 5.26
episode: 2055  score: 10.0   memory length: 454320    epsilon: 0.649222210007615
steps: 524    lr: 6.4000000000000001e-06    evaluation reward: 5.33
episode: 2056  score: 5.0    memory length: 454646    epsilon: 0.648899470007622
steps: 326    lr: 6.4000000000000001e-06    evaluation reward: 5.33
episode: 2057  score: 2.0    memory length: 454844    epsilon: 0.6487034500076263
steps: 198    lr: 6.4000000000000001e-06    evaluation reward: 5.31
episode: 2058  score: 3.0    memory length: 455057    epsilon: 0.6484925800076309
steps: 213    lr: 6.4000000000000001e-06    evaluation reward: 5.28
episode: 2059  score: 2.0    memory length: 455255    epsilon: 0.6482965600076351
steps: 198    lr: 6.4000000000000001e-06    evaluation reward: 5.23
episode: 2060  score: 5.0    memory length: 455565    epsilon: 0.6479896600076418
steps: 310    lr: 6.4000000000000001e-06    evaluation reward: 5.24
episode: 2061  score: 1.0    memory length: 455736    epsilon: 0.6478203700076455
steps: 171    lr: 6.4000000000000001e-06    evaluation reward: 5.19
episode: 2062  score: 5.0    memory length: 456044    epsilon: 0.6475154500076521
steps: 308    lr: 6.4000000000000001e-06    evaluation reward: 5.19
episode: 2063  score: 4.0    memory length: 456313    epsilon: 0.6472491400076579
steps: 269    lr: 6.4000000000000001e-06    evaluation reward: 5.21
episode: 2064  score: 6.0    memory length: 456667    epsilon: 0.6468986800076655
steps: 354    lr: 6.4000000000000001e-06    evaluation reward: 5.2
episode: 2065  score: 4.0    memory length: 456926    epsilon: 0.646642270007671
steps: 259    lr: 6.4000000000000001e-06    evaluation reward: 5.23
episode: 2066  score: 1.0    memory length: 457098    epsilon: 0.6464719900076747
steps: 172    lr: 6.4000000000000001e-06    evaluation reward: 5.16
episode: 2067  score: 4.0    memory length: 457375    epsilon: 0.6461977600076807
steps: 277    lr: 6.4000000000000001e-06    evaluation reward: 5.17
episode: 2068  score: 4.0    memory length: 457636    epsilon: 0.6459393700076863
steps: 261    lr: 6.4000000000000001e-06    evaluation reward: 5.19
episode: 2069  score: 6.0    memory length: 457976    epsilon: 0.6456027700076936
steps: 340    lr: 6.4000000000000001e-06    evaluation reward: 5.22
episode: 2070  score: 2.0    memory length: 458176    epsilon: 0.6454047700076979
steps: 200    lr: 6.4000000000000001e-06    evaluation reward: 5.2
episode: 2071  score: 7.0    memory length: 458563    epsilon: 0.6450216400077062

```



```

steps: 387    lr: 6.4000000000000001e-06    evaluation reward: 5.24
episode: 2072  score: 8.0    memory length: 459009    epsilon: 0.6445801000077158
steps: 446    lr: 6.4000000000000001e-06    evaluation reward: 5.28
episode: 2073  score: 4.0    memory length: 459322    epsilon: 0.6442702300077225
steps: 313    lr: 6.4000000000000001e-06    evaluation reward: 5.26
episode: 2074  score: 3.0    memory length: 459535    epsilon: 0.6440593600077271
steps: 213    lr: 6.4000000000000001e-06    evaluation reward: 5.26
episode: 2075  score: 2.0    memory length: 459733    epsilon: 0.6438633400077314
steps: 198    lr: 6.4000000000000001e-06    evaluation reward: 5.18
episode: 2076  score: 4.0    memory length: 460026    epsilon: 0.6435732700077377
steps: 293    lr: 6.4000000000000001e-06    evaluation reward: 5.16
episode: 2077  score: 10.0   memory length: 460553    epsilon: 0.643051540007749
steps: 527    lr: 6.4000000000000001e-06    evaluation reward: 5.19
episode: 2078  score: 3.0    memory length: 460783    epsilon: 0.6428238400077539
steps: 230    lr: 6.4000000000000001e-06    evaluation reward: 5.16
episode: 2079  score: 7.0    memory length: 461158    epsilon: 0.642452590007762
steps: 375    lr: 6.4000000000000001e-06    evaluation reward: 5.17
episode: 2080  score: 4.0    memory length: 461431    epsilon: 0.6421823200077679
steps: 273    lr: 6.4000000000000001e-06    evaluation reward: 5.2
episode: 2081  score: 9.0    memory length: 461893    epsilon: 0.6417249400077778
steps: 462    lr: 6.4000000000000001e-06    evaluation reward: 5.27
episode: 2082  score: 6.0    memory length: 462214    epsilon: 0.6414071500077847
steps: 321    lr: 6.4000000000000001e-06    evaluation reward: 5.28
episode: 2083  score: 6.0    memory length: 462549    epsilon: 0.6410755000077919
steps: 335    lr: 6.4000000000000001e-06    evaluation reward: 5.28
episode: 2084  score: 6.0    memory length: 462913    epsilon: 0.6407151400077997
steps: 364    lr: 6.4000000000000001e-06    evaluation reward: 5.31
episode: 2085  score: 6.0    memory length: 463274    epsilon: 0.6403577500078075
steps: 361    lr: 6.4000000000000001e-06    evaluation reward: 5.33
episode: 2086  score: 5.0    memory length: 463584    epsilon: 0.6400508500078141
steps: 310    lr: 6.4000000000000001e-06    evaluation reward: 5.29
episode: 2087  score: 6.0    memory length: 463938    epsilon: 0.6397003900078218
steps: 354    lr: 6.4000000000000001e-06    evaluation reward: 5.29
episode: 2088  score: 11.0   memory length: 464495    epsilon:
0.6391489600078337    steps: 557    lr: 6.4000000000000001e-06    evaluation
reward: 5.31
episode: 2089  score: 2.0    memory length: 464693    epsilon: 0.638952940007838
steps: 198    lr: 6.4000000000000001e-06    evaluation reward: 5.3
episode: 2090  score: 7.0    memory length: 465074    epsilon: 0.6385757500078462
steps: 381    lr: 6.4000000000000001e-06    evaluation reward: 5.29
episode: 2091  score: 4.0    memory length: 465370    epsilon: 0.6382827100078525
steps: 296    lr: 6.4000000000000001e-06    evaluation reward: 5.24
episode: 2092  score: 6.0    memory length: 465751    epsilon: 0.6379055200078607
steps: 381    lr: 6.4000000000000001e-06    evaluation reward: 5.26
episode: 2093  score: 5.0    memory length: 466056    epsilon: 0.6376035700078673
steps: 305    lr: 6.4000000000000001e-06    evaluation reward: 5.23
episode: 2094  score: 3.0    memory length: 466264    epsilon: 0.6373976500078717
steps: 208    lr: 6.4000000000000001e-06    evaluation reward: 5.22

```

episode: 2095 score: 4.0 memory length: 466520 epsilon: 0.6371442100078772
 steps: 256 lr: 6.4000000000000001e-06 evaluation reward: 5.23
 episode: 2096 score: 5.0 memory length: 466816 epsilon: 0.6368511700078836
 steps: 296 lr: 6.4000000000000001e-06 evaluation reward: 5.25
 episode: 2097 score: 7.0 memory length: 467213 epsilon: 0.6364581400078921
 steps: 397 lr: 6.4000000000000001e-06 evaluation reward: 5.28
 episode: 2098 score: 5.0 memory length: 467500 epsilon: 0.6361740100078983
 steps: 287 lr: 6.4000000000000001e-06 evaluation reward: 5.3
 episode: 2099 score: 3.0 memory length: 467731 epsilon: 0.6359453200079033
 steps: 231 lr: 6.4000000000000001e-06 evaluation reward: 5.26
 episode: 2100 score: 4.0 memory length: 468025 epsilon: 0.6356542600079096
 steps: 294 lr: 6.4000000000000001e-06 evaluation reward: 5.29
 episode: 2101 score: 7.0 memory length: 468275 epsilon: 0.635406760007915
 steps: 250 lr: 6.4000000000000001e-06 evaluation reward: 5.31
 episode: 2102 score: 5.0 memory length: 468591 epsilon: 0.6350939200079218
 steps: 316 lr: 6.4000000000000001e-06 evaluation reward: 5.32
 episode: 2103 score: 8.0 memory length: 469044 epsilon: 0.6346454500079315
 steps: 453 lr: 6.4000000000000001e-06 evaluation reward: 5.31
 episode: 2104 score: 3.0 memory length: 469254 epsilon: 0.634437550007936
 steps: 210 lr: 6.4000000000000001e-06 evaluation reward: 5.3
 episode: 2105 score: 4.0 memory length: 469513 epsilon: 0.6341811400079416
 steps: 259 lr: 6.4000000000000001e-06 evaluation reward: 5.29
 episode: 2106 score: 5.0 memory length: 469803 epsilon: 0.6338940400079478
 steps: 290 lr: 6.4000000000000001e-06 evaluation reward: 5.29
 episode: 2107 score: 3.0 memory length: 470015 epsilon: 0.6336841600079524
 steps: 212 lr: 6.4000000000000001e-06 evaluation reward: 5.25
 episode: 2108 score: 5.0 memory length: 470324 epsilon: 0.633378250007959
 steps: 309 lr: 6.4000000000000001e-06 evaluation reward: 5.27
 episode: 2109 score: 7.0 memory length: 470730 epsilon: 0.6329763100079677
 steps: 406 lr: 6.4000000000000001e-06 evaluation reward: 5.31
 episode: 2110 score: 8.0 memory length: 471139 epsilon: 0.6325714000079765
 steps: 409 lr: 6.4000000000000001e-06 evaluation reward: 5.35
 episode: 2111 score: 12.0 memory length: 471575 epsilon:
 0.6321397600079859 steps: 436 lr: 6.4000000000000001e-06 evaluation
 reward: 5.4
 episode: 2112 score: 2.0 memory length: 471792 epsilon: 0.6319249300079905
 steps: 217 lr: 6.4000000000000001e-06 evaluation reward: 5.33
 episode: 2113 score: 4.0 memory length: 472065 epsilon: 0.6316546600079964
 steps: 273 lr: 6.4000000000000001e-06 evaluation reward: 5.3
 episode: 2114 score: 5.0 memory length: 472381 epsilon: 0.6313418200080032
 steps: 316 lr: 6.4000000000000001e-06 evaluation reward: 5.3
 episode: 2115 score: 9.0 memory length: 472866 epsilon: 0.6308616700080136
 steps: 485 lr: 6.4000000000000001e-06 evaluation reward: 5.33
 episode: 2116 score: 6.0 memory length: 473233 epsilon: 0.6304983400080215
 steps: 367 lr: 6.4000000000000001e-06 evaluation reward: 5.34
 episode: 2117 score: 9.0 memory length: 473744 epsilon: 0.6299924500080325
 steps: 511 lr: 6.4000000000000001e-06 evaluation reward: 5.36
 episode: 2118 score: 3.0 memory length: 473991 epsilon: 0.6297479200080378

```

steps: 247    lr: 6.4000000000000001e-06    evaluation reward: 5.33
episode: 2119  score: 5.0    memory length: 474296    epsilon: 0.6294459700080444
steps: 305    lr: 6.4000000000000001e-06    evaluation reward: 5.36
episode: 2120  score: 8.0    memory length: 474742    epsilon: 0.629004430008054
steps: 446    lr: 6.4000000000000001e-06    evaluation reward: 5.4
episode: 2121  score: 3.0    memory length: 475010    epsilon: 0.6287391100080597
steps: 268    lr: 6.4000000000000001e-06    evaluation reward: 5.38
episode: 2122  score: 4.0    memory length: 475303    epsilon: 0.628449040008066
steps: 293    lr: 6.4000000000000001e-06    evaluation reward: 5.35
episode: 2123  score: 8.0    memory length: 475705    epsilon: 0.6280510600080746
steps: 402    lr: 6.4000000000000001e-06    evaluation reward: 5.41
episode: 2124  score: 1.0    memory length: 475856    epsilon: 0.6279015700080779
steps: 151    lr: 6.4000000000000001e-06    evaluation reward: 5.37
episode: 2125  score: 3.0    memory length: 476068    epsilon: 0.6276916900080824
steps: 212    lr: 6.4000000000000001e-06    evaluation reward: 5.32
episode: 2126  score: 4.0    memory length: 476325    epsilon: 0.627437260008088
steps: 257    lr: 6.4000000000000001e-06    evaluation reward: 5.3
episode: 2127  score: 9.0    memory length: 476818    epsilon: 0.6269491900080986
steps: 493    lr: 6.4000000000000001e-06    evaluation reward: 5.32
episode: 2128  score: 4.0    memory length: 477077    epsilon: 0.6266927800081041
steps: 259    lr: 6.4000000000000001e-06    evaluation reward: 5.33
episode: 2129  score: 9.0    memory length: 477565    epsilon: 0.6262096600081146
steps: 488    lr: 6.4000000000000001e-06    evaluation reward: 5.38
episode: 2130  score: 8.0    memory length: 477978    epsilon: 0.6258007900081235
steps: 413    lr: 6.4000000000000001e-06    evaluation reward: 5.41
episode: 2131  score: 6.0    memory length: 478353    epsilon: 0.6254295400081316
steps: 375    lr: 6.4000000000000001e-06    evaluation reward: 5.41
episode: 2132  score: 5.0    memory length: 478677    epsilon: 0.6251087800081385
steps: 324    lr: 6.4000000000000001e-06    evaluation reward: 5.34
episode: 2133  score: 3.0    memory length: 478891    epsilon: 0.6248969200081431
steps: 214    lr: 6.4000000000000001e-06    evaluation reward: 5.31
episode: 2134  score: 6.0    memory length: 479219    epsilon: 0.6245722000081502
steps: 328    lr: 6.4000000000000001e-06    evaluation reward: 5.34
episode: 2135  score: 11.0   memory length: 479651    epsilon:
0.6241445200081595    steps: 432    lr: 6.4000000000000001e-06    evaluation
reward: 5.36
episode: 2136  score: 6.0    memory length: 480025    epsilon: 0.6237742600081675
steps: 374    lr: 6.4000000000000001e-06    evaluation reward: 5.35
episode: 2137  score: 7.0    memory length: 480446    epsilon: 0.6233574700081765
steps: 421    lr: 6.4000000000000001e-06    evaluation reward: 5.37
episode: 2138  score: 8.0    memory length: 480869    epsilon: 0.6229387000081856
steps: 423    lr: 6.4000000000000001e-06    evaluation reward: 5.38
episode: 2139  score: 9.0    memory length: 481372    epsilon: 0.6224407300081964
steps: 503    lr: 6.4000000000000001e-06    evaluation reward: 5.43
episode: 2140  score: 6.0    memory length: 481694    epsilon: 0.6221219500082034
steps: 322    lr: 6.4000000000000001e-06    evaluation reward: 5.46
episode: 2141  score: 9.0    memory length: 482146    epsilon: 0.6216744700082131
steps: 452    lr: 6.4000000000000001e-06    evaluation reward: 5.39

```

episode: 2142 score: 6.0 memory length: 482546 epsilon: 0.6212784700082217
 steps: 400 lr: 6.4000000000000001e-06 evaluation reward: 5.41
 episode: 2143 score: 8.0 memory length: 483010 epsilon: 0.6208191100082316
 steps: 464 lr: 6.4000000000000001e-06 evaluation reward: 5.35
 episode: 2144 score: 15.0 memory length: 483606 epsilon:
 0.6202290700082445 steps: 596 lr: 6.4000000000000001e-06 evaluation
 reward: 5.46
 episode: 2145 score: 7.0 memory length: 484005 epsilon: 0.619834060008253
 steps: 399 lr: 6.4000000000000001e-06 evaluation reward: 5.5
 episode: 2146 score: 3.0 memory length: 484252 epsilon: 0.6195895300082583
 steps: 247 lr: 6.4000000000000001e-06 evaluation reward: 5.47
 episode: 2147 score: 8.0 memory length: 484689 epsilon: 0.6191569000082677
 steps: 437 lr: 6.4000000000000001e-06 evaluation reward: 5.53
 episode: 2148 score: 6.0 memory length: 485043 epsilon: 0.6188064400082753
 steps: 354 lr: 6.4000000000000001e-06 evaluation reward: 5.55
 episode: 2149 score: 12.0 memory length: 485522 epsilon:
 0.6183322300082856 steps: 479 lr: 6.4000000000000001e-06 evaluation
 reward: 5.62
 episode: 2150 score: 3.0 memory length: 485767 epsilon: 0.6180896800082909
 steps: 245 lr: 6.4000000000000001e-06 evaluation reward: 5.58
 episode: 2151 score: 6.0 memory length: 486098 epsilon: 0.617761990008298
 steps: 331 lr: 6.4000000000000001e-06 evaluation reward: 5.6
 episode: 2152 score: 7.0 memory length: 486455 epsilon: 0.6174085600083057
 steps: 357 lr: 6.4000000000000001e-06 evaluation reward: 5.62
 episode: 2153 score: 8.0 memory length: 486855 epsilon: 0.6170125600083143
 steps: 400 lr: 6.4000000000000001e-06 evaluation reward: 5.68
 episode: 2154 score: 5.0 memory length: 487161 epsilon: 0.6167096200083209
 steps: 306 lr: 6.4000000000000001e-06 evaluation reward: 5.67
 episode: 2155 score: 3.0 memory length: 487409 epsilon: 0.6164641000083262
 steps: 248 lr: 6.4000000000000001e-06 evaluation reward: 5.6
 episode: 2156 score: 6.0 memory length: 487784 epsilon: 0.6160928500083342
 steps: 375 lr: 6.4000000000000001e-06 evaluation reward: 5.61
 episode: 2157 score: 8.0 memory length: 488186 epsilon: 0.6156948700083429
 steps: 402 lr: 6.4000000000000001e-06 evaluation reward: 5.67
 episode: 2158 score: 6.0 memory length: 488560 epsilon: 0.6153246100083509
 steps: 374 lr: 6.4000000000000001e-06 evaluation reward: 5.7
 episode: 2159 score: 5.0 memory length: 488834 epsilon: 0.6150533500083568
 steps: 274 lr: 6.4000000000000001e-06 evaluation reward: 5.73
 episode: 2160 score: 5.0 memory length: 489138 epsilon: 0.6147523900083633
 steps: 304 lr: 6.4000000000000001e-06 evaluation reward: 5.73
 episode: 2161 score: 3.0 memory length: 489349 epsilon: 0.6145435000083679
 steps: 211 lr: 6.4000000000000001e-06 evaluation reward: 5.75
 episode: 2162 score: 7.0 memory length: 489710 epsilon: 0.6141861100083756
 steps: 361 lr: 6.4000000000000001e-06 evaluation reward: 5.77
 episode: 2163 score: 6.0 memory length: 490067 epsilon: 0.6138326800083833
 steps: 357 lr: 6.4000000000000001e-06 evaluation reward: 5.79
 episode: 2164 score: 4.0 memory length: 490329 epsilon: 0.6135733000083889
 steps: 262 lr: 6.4000000000000001e-06 evaluation reward: 5.77

episode: 2165 score: 7.0 memory length: 490743 epsilon: 0.6131634400083978
 steps: 414 lr: 6.4000000000000001e-06 evaluation reward: 5.8
 episode: 2166 score: 3.0 memory length: 490968 epsilon: 0.6129406900084027
 steps: 225 lr: 6.4000000000000001e-06 evaluation reward: 5.82
 episode: 2167 score: 3.0 memory length: 491198 epsilon: 0.6127129900084076
 steps: 230 lr: 6.4000000000000001e-06 evaluation reward: 5.81
 episode: 2168 score: 2.0 memory length: 491379 epsilon: 0.6125338000084115
 steps: 181 lr: 6.4000000000000001e-06 evaluation reward: 5.79
 episode: 2169 score: 12.0 memory length: 491931 epsilon:
 0.6119873200084234 steps: 552 lr: 6.4000000000000001e-06 evaluation
 reward: 5.85
 episode: 2170 score: 4.0 memory length: 492210 epsilon: 0.6117111100084294
 steps: 279 lr: 6.4000000000000001e-06 evaluation reward: 5.87
 episode: 2171 score: 5.0 memory length: 492532 epsilon: 0.6113923300084363
 steps: 322 lr: 6.4000000000000001e-06 evaluation reward: 5.85
 episode: 2172 score: 3.0 memory length: 492761 epsilon: 0.6111656200084412
 steps: 229 lr: 6.4000000000000001e-06 evaluation reward: 5.8
 episode: 2173 score: 8.0 memory length: 493159 epsilon: 0.6107716000084498
 steps: 398 lr: 6.4000000000000001e-06 evaluation reward: 5.84
 episode: 2174 score: 4.0 memory length: 493454 epsilon: 0.6104795500084561
 steps: 295 lr: 6.4000000000000001e-06 evaluation reward: 5.85
 episode: 2175 score: 7.0 memory length: 493847 epsilon: 0.6100904800084646
 steps: 393 lr: 6.4000000000000001e-06 evaluation reward: 5.9
 episode: 2176 score: 1.0 memory length: 493997 epsilon: 0.6099419800084678
 steps: 150 lr: 6.4000000000000001e-06 evaluation reward: 5.87
 episode: 2177 score: 2.0 memory length: 494219 epsilon: 0.6097222000084725
 steps: 222 lr: 6.4000000000000001e-06 evaluation reward: 5.79
 episode: 2178 score: 5.0 memory length: 494498 epsilon: 0.6094459900084785
 steps: 279 lr: 6.4000000000000001e-06 evaluation reward: 5.81
 episode: 2179 score: 1.0 memory length: 494669 epsilon: 0.6092767000084822
 steps: 171 lr: 6.4000000000000001e-06 evaluation reward: 5.75
 episode: 2180 score: 4.0 memory length: 494930 epsilon: 0.6090183100084878
 steps: 261 lr: 6.4000000000000001e-06 evaluation reward: 5.75
 episode: 2181 score: 6.0 memory length: 495280 epsilon: 0.6086718100084954
 steps: 350 lr: 6.4000000000000001e-06 evaluation reward: 5.72
 episode: 2182 score: 4.0 memory length: 495547 epsilon: 0.6084074800085011
 steps: 267 lr: 6.4000000000000001e-06 evaluation reward: 5.7
 episode: 2183 score: 7.0 memory length: 495954 epsilon: 0.6080045500085098
 steps: 407 lr: 6.4000000000000001e-06 evaluation reward: 5.71
 episode: 2184 score: 4.0 memory length: 496233 epsilon: 0.6077283400085158
 steps: 279 lr: 6.4000000000000001e-06 evaluation reward: 5.69
 episode: 2185 score: 4.0 memory length: 496473 epsilon: 0.607490740008521
 steps: 240 lr: 6.4000000000000001e-06 evaluation reward: 5.67
 episode: 2186 score: 2.0 memory length: 496671 epsilon: 0.6072947200085252
 steps: 198 lr: 6.4000000000000001e-06 evaluation reward: 5.64
 episode: 2187 score: 2.0 memory length: 496854 epsilon: 0.6071135500085292
 steps: 183 lr: 6.4000000000000001e-06 evaluation reward: 5.6
 episode: 2188 score: 7.0 memory length: 497206 epsilon: 0.6067650700085367

```

steps: 352    lr: 6.4000000000000001e-06    evaluation reward: 5.56
episode: 2189  score: 8.0    memory length: 497684    epsilon: 0.606291850008547
steps: 478    lr: 6.4000000000000001e-06    evaluation reward: 5.62
episode: 2190  score: 5.0    memory length: 498046    epsilon: 0.6059334700085548
steps: 362    lr: 6.4000000000000001e-06    evaluation reward: 5.6
episode: 2191  score: 4.0    memory length: 498344    epsilon: 0.6056384500085612
steps: 298    lr: 6.4000000000000001e-06    evaluation reward: 5.6
episode: 2192  score: 8.0    memory length: 498639    epsilon: 0.6053464000085675
steps: 295    lr: 6.4000000000000001e-06    evaluation reward: 5.62
episode: 2193  score: 8.0    memory length: 499070    epsilon: 0.6049197100085768
steps: 431    lr: 6.4000000000000001e-06    evaluation reward: 5.65
episode: 2194  score: 7.0    memory length: 499415    epsilon: 0.6045781600085842
steps: 345    lr: 6.4000000000000001e-06    evaluation reward: 5.69
episode: 2195  score: 2.0    memory length: 499617    epsilon: 0.6043781800085886
steps: 202    lr: 6.4000000000000001e-06    evaluation reward: 5.67
episode: 2196  score: 6.0    memory length: 499957    epsilon: 0.6040415800085959
steps: 340    lr: 6.4000000000000001e-06    evaluation reward: 5.68
episode: 2197  score: 6.0    memory length: 500303    epsilon: 0.6036990400086033
steps: 346    lr: 2.5600000000000001e-06    evaluation reward: 5.67
episode: 2198  score: 6.0    memory length: 500653    epsilon: 0.6033525400086108
steps: 350    lr: 2.5600000000000001e-06    evaluation reward: 5.68
episode: 2199  score: 12.0   memory length: 501165    epsilon:
0.6028456600086218    steps: 512    lr: 2.5600000000000001e-06    evaluation
reward: 5.77
episode: 2200  score: 8.0    memory length: 501572    epsilon: 0.6024427300086306
steps: 407    lr: 2.5600000000000001e-06    evaluation reward: 5.81
episode: 2201  score: 5.0    memory length: 501909    epsilon: 0.6021091000086378
steps: 337    lr: 2.5600000000000001e-06    evaluation reward: 5.79
episode: 2202  score: 5.0    memory length: 502256    epsilon: 0.6017655700086453
steps: 347    lr: 2.5600000000000001e-06    evaluation reward: 5.79
episode: 2203  score: 4.0    memory length: 502514    epsilon: 0.6015101500086508
steps: 258    lr: 2.5600000000000001e-06    evaluation reward: 5.75
episode: 2204  score: 4.0    memory length: 502786    epsilon: 0.6012408700086567
steps: 272    lr: 2.5600000000000001e-06    evaluation reward: 5.76
episode: 2205  score: 7.0    memory length: 503229    epsilon: 0.6008023000086662
steps: 443    lr: 2.5600000000000001e-06    evaluation reward: 5.79
episode: 2206  score: 2.0    memory length: 503411    epsilon: 0.6006221200086701
steps: 182    lr: 2.5600000000000001e-06    evaluation reward: 5.76
episode: 2207  score: 9.0    memory length: 503899    epsilon: 0.6001390000086806
steps: 488    lr: 2.5600000000000001e-06    evaluation reward: 5.82
episode: 2208  score: 6.0    memory length: 504262    epsilon: 0.5997796300086884
steps: 363    lr: 2.5600000000000001e-06    evaluation reward: 5.83
episode: 2209  score: 6.0    memory length: 504626    epsilon: 0.5994192700086962
steps: 364    lr: 2.5600000000000001e-06    evaluation reward: 5.82
episode: 2210  score: 3.0    memory length: 504858    epsilon: 0.5991895900087012
steps: 232    lr: 2.5600000000000001e-06    evaluation reward: 5.77
episode: 2211  score: 5.0    memory length: 505145    epsilon: 0.5989054600087074
steps: 287    lr: 2.5600000000000001e-06    evaluation reward: 5.7

```

episode: 2212 score: 8.0 memory length: 505569 epsilon: 0.5984857000087165
 steps: 424 lr: 2.5600000000000001e-06 evaluation reward: 5.76
 episode: 2213 score: 7.0 memory length: 505940 epsilon: 0.5981184100087245
 steps: 371 lr: 2.5600000000000001e-06 evaluation reward: 5.79
 episode: 2214 score: 3.0 memory length: 506168 epsilon: 0.5978926900087294
 steps: 228 lr: 2.5600000000000001e-06 evaluation reward: 5.77
 episode: 2215 score: 8.0 memory length: 506578 epsilon: 0.5974867900087382
 steps: 410 lr: 2.5600000000000001e-06 evaluation reward: 5.76
 episode: 2216 score: 6.0 memory length: 506933 epsilon: 0.5971353400087458
 steps: 355 lr: 2.5600000000000001e-06 evaluation reward: 5.76
 episode: 2217 score: 6.0 memory length: 507297 epsilon: 0.5967749800087536
 steps: 364 lr: 2.5600000000000001e-06 evaluation reward: 5.73
 episode: 2218 score: 7.0 memory length: 507545 epsilon: 0.596529460008759
 steps: 248 lr: 2.5600000000000001e-06 evaluation reward: 5.77
 episode: 2219 score: 4.0 memory length: 507786 epsilon: 0.5962908700087641
 steps: 241 lr: 2.5600000000000001e-06 evaluation reward: 5.76
 episode: 2220 score: 10.0 memory length: 508299 epsilon:
 0.5957830000087752 steps: 513 lr: 2.5600000000000001e-06 evaluation
 reward: 5.78
 episode: 2221 score: 2.0 memory length: 508499 epsilon: 0.5955850000087795
 steps: 200 lr: 2.5600000000000001e-06 evaluation reward: 5.77
 episode: 2222 score: 7.0 memory length: 508872 epsilon: 0.5952157300087875
 steps: 373 lr: 2.5600000000000001e-06 evaluation reward: 5.8
 episode: 2223 score: 13.0 memory length: 509420 epsilon:
 0.5946732100087992 steps: 548 lr: 2.5600000000000001e-06 evaluation
 reward: 5.85
 episode: 2224 score: 7.0 memory length: 509789 epsilon: 0.5943079000088072
 steps: 369 lr: 2.5600000000000001e-06 evaluation reward: 5.91
 episode: 2225 score: 4.0 memory length: 510049 epsilon: 0.5940505000088128
 steps: 260 lr: 2.5600000000000001e-06 evaluation reward: 5.92
 episode: 2226 score: 3.0 memory length: 510302 epsilon: 0.5938000300088182
 steps: 253 lr: 2.5600000000000001e-06 evaluation reward: 5.91
 episode: 2227 score: 7.0 memory length: 510689 epsilon: 0.5934169000088265
 steps: 387 lr: 2.5600000000000001e-06 evaluation reward: 5.89
 episode: 2228 score: 10.0 memory length: 511215 epsilon:
 0.5928961600088378 steps: 526 lr: 2.5600000000000001e-06 evaluation
 reward: 5.95
 episode: 2229 score: 4.0 memory length: 511473 epsilon: 0.5926407400088434
 steps: 258 lr: 2.5600000000000001e-06 evaluation reward: 5.9
 episode: 2230 score: 7.0 memory length: 511874 epsilon: 0.592243750008852
 steps: 401 lr: 2.5600000000000001e-06 evaluation reward: 5.89
 episode: 2231 score: 7.0 memory length: 512280 epsilon: 0.5918418100088607
 steps: 406 lr: 2.5600000000000001e-06 evaluation reward: 5.9
 episode: 2232 score: 10.0 memory length: 512755 epsilon:
 0.5913715600088709 steps: 475 lr: 2.5600000000000001e-06 evaluation
 reward: 5.95
 episode: 2233 score: 7.0 memory length: 513126 epsilon: 0.5910042700088789
 steps: 371 lr: 2.5600000000000001e-06 evaluation reward: 5.99

episode: 2234 score: 5.0 memory length: 513432 epsilon: 0.5907013300088855
 steps: 306 lr: 2.5600000000000001e-06 evaluation reward: 5.98
 episode: 2235 score: 4.0 memory length: 513715 epsilon: 0.5904211600088916
 steps: 283 lr: 2.5600000000000001e-06 evaluation reward: 5.91
 episode: 2236 score: 6.0 memory length: 514084 epsilon: 0.5900558500088995
 steps: 369 lr: 2.5600000000000001e-06 evaluation reward: 5.91
 episode: 2237 score: 3.0 memory length: 514297 epsilon: 0.5898449800089041
 steps: 213 lr: 2.5600000000000001e-06 evaluation reward: 5.87
 episode: 2238 score: 5.0 memory length: 514607 epsilon: 0.5895380800089107
 steps: 310 lr: 2.5600000000000001e-06 evaluation reward: 5.84
 episode: 2239 score: 8.0 memory length: 515054 epsilon: 0.5890955500089203
 steps: 447 lr: 2.5600000000000001e-06 evaluation reward: 5.83
 episode: 2240 score: 8.0 memory length: 515502 epsilon: 0.58865203000893
 steps: 448 lr: 2.5600000000000001e-06 evaluation reward: 5.85
 episode: 2241 score: 7.0 memory length: 515889 epsilon: 0.5882689000089383
 steps: 387 lr: 2.5600000000000001e-06 evaluation reward: 5.83
 episode: 2242 score: 7.0 memory length: 516258 epsilon: 0.5879035900089462
 steps: 369 lr: 2.5600000000000001e-06 evaluation reward: 5.84
 episode: 2243 score: 3.0 memory length: 516469 epsilon: 0.5876947000089507
 steps: 211 lr: 2.5600000000000001e-06 evaluation reward: 5.79
 episode: 2244 score: 11.0 memory length: 516984 epsilon:
 0.5871848500089618 steps: 515 lr: 2.5600000000000001e-06 evaluation
 reward: 5.75
 episode: 2245 score: 7.0 memory length: 517377 epsilon: 0.5867957800089703
 steps: 393 lr: 2.5600000000000001e-06 evaluation reward: 5.75
 episode: 2246 score: 6.0 memory length: 517764 epsilon: 0.5864126500089786
 steps: 387 lr: 2.5600000000000001e-06 evaluation reward: 5.78
 episode: 2247 score: 7.0 memory length: 518119 epsilon: 0.5860612000089862
 steps: 355 lr: 2.5600000000000001e-06 evaluation reward: 5.77
 episode: 2248 score: 10.0 memory length: 518609 epsilon:
 0.5855761000089967 steps: 490 lr: 2.5600000000000001e-06 evaluation
 reward: 5.81
 episode: 2249 score: 10.0 memory length: 519162 epsilon:
 0.5850286300090086 steps: 553 lr: 2.5600000000000001e-06 evaluation
 reward: 5.79
 episode: 2250 score: 7.0 memory length: 519564 epsilon: 0.5846306500090173
 steps: 402 lr: 2.5600000000000001e-06 evaluation reward: 5.83
 episode: 2251 score: 9.0 memory length: 520038 epsilon: 0.5841613900090274
 steps: 474 lr: 2.5600000000000001e-06 evaluation reward: 5.86
 episode: 2252 score: 15.0 memory length: 520620 epsilon: 0.58358521000904
 steps: 582 lr: 2.5600000000000001e-06 evaluation reward: 5.94
 episode: 2253 score: 5.0 memory length: 520950 epsilon: 0.583258510009047
 steps: 330 lr: 2.5600000000000001e-06 evaluation reward: 5.91
 episode: 2254 score: 5.0 memory length: 521294 epsilon: 0.5829179500090544
 steps: 344 lr: 2.5600000000000001e-06 evaluation reward: 5.91
 episode: 2255 score: 10.0 memory length: 521738 epsilon: 0.582478390009064
 steps: 444 lr: 2.5600000000000001e-06 evaluation reward: 5.98
 episode: 2256 score: 8.0 memory length: 522178 epsilon: 0.5820427900090734


```

steps: 440    lr: 2.5600000000000001e-06    evaluation reward: 6.0
episode: 2257  score: 4.0    memory length: 522451    epsilon: 0.5817725200090793
steps: 273    lr: 2.5600000000000001e-06    evaluation reward: 5.96
episode: 2258  score: 3.0    memory length: 522702    epsilon: 0.5815240300090847
steps: 251    lr: 2.5600000000000001e-06    evaluation reward: 5.93
episode: 2259  score: 4.0    memory length: 522977    epsilon: 0.5812517800090906
steps: 275    lr: 2.5600000000000001e-06    evaluation reward: 5.92
episode: 2260  score: 5.0    memory length: 523289    epsilon: 0.5809429000090973
steps: 312    lr: 2.5600000000000001e-06    evaluation reward: 5.92
episode: 2261  score: 9.0    memory length: 523811    epsilon: 0.5804261200091085
steps: 522    lr: 2.5600000000000001e-06    evaluation reward: 5.98
episode: 2262  score: 12.0   memory length: 524369    epsilon:
0.5798737000091205    steps: 558    lr: 2.5600000000000001e-06    evaluation
reward: 6.03
episode: 2263  score: 6.0    memory length: 524713    epsilon: 0.5795331400091279
steps: 344    lr: 2.5600000000000001e-06    evaluation reward: 6.03
episode: 2264  score: 10.0   memory length: 525087    epsilon: 0.579162880009136
steps: 374    lr: 2.5600000000000001e-06    evaluation reward: 6.09
episode: 2265  score: 6.0    memory length: 525463    epsilon: 0.578790640009144
steps: 376    lr: 2.5600000000000001e-06    evaluation reward: 6.08
episode: 2266  score: 4.0    memory length: 525739    epsilon: 0.57851740000915
steps: 276    lr: 2.5600000000000001e-06    evaluation reward: 6.09
episode: 2267  score: 5.0    memory length: 526081    epsilon: 0.5781788200091573
steps: 342    lr: 2.5600000000000001e-06    evaluation reward: 6.11
episode: 2268  score: 3.0    memory length: 526327    epsilon: 0.5779352800091626
steps: 246    lr: 2.5600000000000001e-06    evaluation reward: 6.12
episode: 2269  score: 9.0    memory length: 526652    epsilon: 0.5776135300091696
steps: 325    lr: 2.5600000000000001e-06    evaluation reward: 6.09
episode: 2270  score: 2.0    memory length: 526832    epsilon: 0.5774353300091735
steps: 180    lr: 2.5600000000000001e-06    evaluation reward: 6.07
episode: 2271  score: 8.0    memory length: 527251    epsilon: 0.5770205200091825
steps: 419    lr: 2.5600000000000001e-06    evaluation reward: 6.1
episode: 2272  score: 3.0    memory length: 527464    epsilon: 0.576809650009187
steps: 213    lr: 2.5600000000000001e-06    evaluation reward: 6.1
episode: 2273  score: 10.0   memory length: 527901    epsilon:
0.5763770200091964    steps: 437    lr: 2.5600000000000001e-06    evaluation
reward: 6.12
episode: 2274  score: 8.0    memory length: 528360    epsilon: 0.5759226100092063
steps: 459    lr: 2.5600000000000001e-06    evaluation reward: 6.16
episode: 2275  score: 11.0   memory length: 528817    epsilon:
0.5754701800092161    steps: 457    lr: 2.5600000000000001e-06    evaluation
reward: 6.2
episode: 2276  score: 8.0    memory length: 529253    epsilon: 0.5750385400092255
steps: 436    lr: 2.5600000000000001e-06    evaluation reward: 6.27
episode: 2277  score: 7.0    memory length: 529696    epsilon: 0.574599970009235
steps: 443    lr: 2.5600000000000001e-06    evaluation reward: 6.32
episode: 2278  score: 6.0    memory length: 530084    epsilon: 0.5742158500092434
steps: 388    lr: 2.5600000000000001e-06    evaluation reward: 6.33

```

episode: 2279 score: 6.0 memory length: 530464 epsilon: 0.5738396500092515
 steps: 380 lr: 2.5600000000000001e-06 evaluation reward: 6.38
 episode: 2280 score: 5.0 memory length: 530807 epsilon: 0.5735000800092589
 steps: 343 lr: 2.5600000000000001e-06 evaluation reward: 6.39
 episode: 2281 score: 6.0 memory length: 531151 epsilon: 0.5731595200092663
 steps: 344 lr: 2.5600000000000001e-06 evaluation reward: 6.39
 episode: 2282 score: 8.0 memory length: 531575 epsilon: 0.5727397600092754
 steps: 424 lr: 2.5600000000000001e-06 evaluation reward: 6.43
 episode: 2283 score: 7.0 memory length: 531967 epsilon: 0.5723516800092838
 steps: 392 lr: 2.5600000000000001e-06 evaluation reward: 6.43
 episode: 2284 score: 5.0 memory length: 532252 epsilon: 0.57206953000929
 steps: 285 lr: 2.5600000000000001e-06 evaluation reward: 6.44
 episode: 2285 score: 6.0 memory length: 532606 epsilon: 0.5717190700092976
 steps: 354 lr: 2.5600000000000001e-06 evaluation reward: 6.46
 episode: 2286 score: 6.0 memory length: 532940 epsilon: 0.5713884100093047
 steps: 334 lr: 2.5600000000000001e-06 evaluation reward: 6.5
 episode: 2287 score: 9.0 memory length: 533429 epsilon: 0.5709043000093152
 steps: 489 lr: 2.5600000000000001e-06 evaluation reward: 6.57
 episode: 2288 score: 5.0 memory length: 533737 epsilon: 0.5705993800093219
 steps: 308 lr: 2.5600000000000001e-06 evaluation reward: 6.55
 episode: 2289 score: 3.0 memory length: 533963 epsilon: 0.5703756400093267
 steps: 226 lr: 2.5600000000000001e-06 evaluation reward: 6.5
 episode: 2290 score: 5.0 memory length: 534308 epsilon: 0.5700340900093341
 steps: 345 lr: 2.5600000000000001e-06 evaluation reward: 6.5
 episode: 2291 score: 6.0 memory length: 534631 epsilon: 0.5697143200093411
 steps: 323 lr: 2.5600000000000001e-06 evaluation reward: 6.52
 episode: 2292 score: 8.0 memory length: 535075 epsilon: 0.5692747600093506
 steps: 444 lr: 2.5600000000000001e-06 evaluation reward: 6.52
 episode: 2293 score: 4.0 memory length: 535353 epsilon: 0.5689995400093566
 steps: 278 lr: 2.5600000000000001e-06 evaluation reward: 6.48
 episode: 2294 score: 11.0 memory length: 535864 epsilon:
 0.5684936500093676 steps: 511 lr: 2.5600000000000001e-06 evaluation
 reward: 6.52
 episode: 2295 score: 6.0 memory length: 536220 epsilon: 0.5681412100093752
 steps: 356 lr: 2.5600000000000001e-06 evaluation reward: 6.56
 episode: 2296 score: 6.0 memory length: 536563 epsilon: 0.5678016400093826
 steps: 343 lr: 2.5600000000000001e-06 evaluation reward: 6.56
 episode: 2297 score: 7.0 memory length: 536988 epsilon: 0.5673808900093917
 steps: 425 lr: 2.5600000000000001e-06 evaluation reward: 6.57
 episode: 2298 score: 7.0 memory length: 537375 epsilon: 0.5669977600094
 steps: 387 lr: 2.5600000000000001e-06 evaluation reward: 6.58
 episode: 2299 score: 6.0 memory length: 537720 epsilon: 0.5666562100094075
 steps: 345 lr: 2.5600000000000001e-06 evaluation reward: 6.52
 episode: 2300 score: 6.0 memory length: 538054 epsilon: 0.5663255500094146
 steps: 334 lr: 2.5600000000000001e-06 evaluation reward: 6.5
 episode: 2301 score: 6.0 memory length: 538400 epsilon: 0.5659830100094221
 steps: 346 lr: 2.5600000000000001e-06 evaluation reward: 6.51
 episode: 2302 score: 4.0 memory length: 538661 epsilon: 0.5657246200094277

```

steps: 261    lr: 2.5600000000000001e-06    evaluation reward: 6.5
episode: 2303  score: 4.0    memory length: 538947    epsilon: 0.5654414800094338
steps: 286    lr: 2.5600000000000001e-06    evaluation reward: 6.5
episode: 2304  score: 7.0    memory length: 539356    epsilon: 0.5650365700094426
steps: 409    lr: 2.5600000000000001e-06    evaluation reward: 6.53
episode: 2305  score: 4.0    memory length: 539617    epsilon: 0.5647781800094482
steps: 261    lr: 2.5600000000000001e-06    evaluation reward: 6.5
episode: 2306  score: 7.0    memory length: 539987    epsilon: 0.5644118800094562
steps: 370    lr: 2.5600000000000001e-06    evaluation reward: 6.55
episode: 2307  score: 9.0    memory length: 540482    epsilon: 0.5639218300094668
steps: 495    lr: 2.5600000000000001e-06    evaluation reward: 6.55
episode: 2308  score: 5.0    memory length: 540795    epsilon: 0.5636119600094736
steps: 313    lr: 2.5600000000000001e-06    evaluation reward: 6.54
episode: 2309  score: 10.0   memory length: 541312    epsilon:
0.5631001300094847    steps: 517    lr: 2.5600000000000001e-06    evaluation
reward: 6.58
episode: 2310  score: 5.0    memory length: 541619    epsilon: 0.5627962000094913
steps: 307    lr: 2.5600000000000001e-06    evaluation reward: 6.6
episode: 2311  score: 4.0    memory length: 541879    epsilon: 0.5625388000094969
steps: 260    lr: 2.5600000000000001e-06    evaluation reward: 6.59
episode: 2312  score: 6.0    memory length: 542241    epsilon: 0.5621804200095046
steps: 362    lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 2313  score: 7.0    memory length: 542615    epsilon: 0.5618101600095127
steps: 374    lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 2314  score: 3.0    memory length: 542844    epsilon: 0.5615834500095176
steps: 229    lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 2315  score: 8.0    memory length: 543280    epsilon: 0.561151810009527
steps: 436    lr: 2.5600000000000001e-06    evaluation reward: 6.57
episode: 2316  score: 9.0    memory length: 543735    epsilon: 0.5607013600095367
steps: 455    lr: 2.5600000000000001e-06    evaluation reward: 6.6
episode: 2317  score: 4.0    memory length: 544010    epsilon: 0.5604291100095427
steps: 275    lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 2318  score: 3.0    memory length: 544236    epsilon: 0.5602053700095475
steps: 226    lr: 2.5600000000000001e-06    evaluation reward: 6.54
episode: 2319  score: 4.0    memory length: 544536    epsilon: 0.559908370009554
steps: 300    lr: 2.5600000000000001e-06    evaluation reward: 6.54
episode: 2320  score: 7.0    memory length: 544939    epsilon: 0.5595094000095626
steps: 403    lr: 2.5600000000000001e-06    evaluation reward: 6.51
episode: 2321  score: 9.0    memory length: 545411    epsilon: 0.5590421200095728
steps: 472    lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 2322  score: 5.0    memory length: 545705    epsilon: 0.5587510600095791
steps: 294    lr: 2.5600000000000001e-06    evaluation reward: 6.56
episode: 2323  score: 4.0    memory length: 545963    epsilon: 0.5584956400095846
steps: 258    lr: 2.5600000000000001e-06    evaluation reward: 6.47
episode: 2324  score: 6.0    memory length: 546309    epsilon: 0.5581531000095921
steps: 346    lr: 2.5600000000000001e-06    evaluation reward: 6.46
episode: 2325  score: 7.0    memory length: 546719    epsilon: 0.5577472000096009
steps: 410    lr: 2.5600000000000001e-06    evaluation reward: 6.49

```

episode: 2326 score: 9.0 memory length: 547172 epsilon: 0.5572987300096106
 steps: 453 lr: 2.5600000000000001e-06 evaluation reward: 6.55
 episode: 2327 score: 8.0 memory length: 547589 epsilon: 0.5568859000096196
 steps: 417 lr: 2.5600000000000001e-06 evaluation reward: 6.56
 episode: 2328 score: 4.0 memory length: 547848 epsilon: 0.5566294900096251
 steps: 259 lr: 2.5600000000000001e-06 evaluation reward: 6.5
 episode: 2329 score: 5.0 memory length: 548154 epsilon: 0.5563265500096317
 steps: 306 lr: 2.5600000000000001e-06 evaluation reward: 6.51
 episode: 2330 score: 9.0 memory length: 548605 epsilon: 0.5558800600096414
 steps: 451 lr: 2.5600000000000001e-06 evaluation reward: 6.53
 episode: 2331 score: 4.0 memory length: 548883 epsilon: 0.5556048400096474
 steps: 278 lr: 2.5600000000000001e-06 evaluation reward: 6.5
 episode: 2332 score: 11.0 memory length: 549416 epsilon:
 0.5550771700096588 steps: 533 lr: 2.5600000000000001e-06 evaluation
 reward: 6.51
 episode: 2333 score: 8.0 memory length: 549862 epsilon: 0.5546356300096684
 steps: 446 lr: 2.5600000000000001e-06 evaluation reward: 6.52
 episode: 2334 score: 8.0 memory length: 550320 epsilon: 0.5541822100096783
 steps: 458 lr: 2.5600000000000001e-06 evaluation reward: 6.55
 episode: 2335 score: 4.0 memory length: 550577 epsilon: 0.5539277800096838
 steps: 257 lr: 2.5600000000000001e-06 evaluation reward: 6.55
 episode: 2336 score: 7.0 memory length: 550956 epsilon: 0.5535525700096919
 steps: 379 lr: 2.5600000000000001e-06 evaluation reward: 6.56
 episode: 2337 score: 5.0 memory length: 551237 epsilon: 0.553274380009698
 steps: 281 lr: 2.5600000000000001e-06 evaluation reward: 6.58
 episode: 2338 score: 10.0 memory length: 551718 epsilon:
 0.5527981900097083 steps: 481 lr: 2.5600000000000001e-06 evaluation
 reward: 6.63
 episode: 2339 score: 7.0 memory length: 552162 epsilon: 0.5523586300097179
 steps: 444 lr: 2.5600000000000001e-06 evaluation reward: 6.62
 episode: 2340 score: 4.0 memory length: 552437 epsilon: 0.5520863800097238
 steps: 275 lr: 2.5600000000000001e-06 evaluation reward: 6.58
 episode: 2341 score: 8.0 memory length: 552843 epsilon: 0.5516844400097325
 steps: 406 lr: 2.5600000000000001e-06 evaluation reward: 6.59
 episode: 2342 score: 4.0 memory length: 553138 epsilon: 0.5513923900097388
 steps: 295 lr: 2.5600000000000001e-06 evaluation reward: 6.56
 episode: 2343 score: 6.0 memory length: 553510 epsilon: 0.5510241100097468
 steps: 372 lr: 2.5600000000000001e-06 evaluation reward: 6.59
 episode: 2344 score: 5.0 memory length: 553806 epsilon: 0.5507310700097532
 steps: 296 lr: 2.5600000000000001e-06 evaluation reward: 6.53
 episode: 2345 score: 8.0 memory length: 554230 epsilon: 0.5503113100097623
 steps: 424 lr: 2.5600000000000001e-06 evaluation reward: 6.54
 episode: 2346 score: 7.0 memory length: 554636 epsilon: 0.549909370009771
 steps: 406 lr: 2.5600000000000001e-06 evaluation reward: 6.55
 episode: 2347 score: 7.0 memory length: 555028 epsilon: 0.5495212900097795
 steps: 392 lr: 2.5600000000000001e-06 evaluation reward: 6.55
 episode: 2348 score: 11.0 memory length: 555585 epsilon:
 0.5489698600097914 steps: 557 lr: 2.5600000000000001e-06 evaluation

```

reward: 6.56
episode: 2349  score: 8.0  memory length: 556038  epsilon: 0.5485213900098012
steps: 453  lr: 2.5600000000000001e-06  evaluation reward: 6.54
episode: 2350  score: 5.0  memory length: 556329  epsilon: 0.5482333000098074
steps: 291  lr: 2.5600000000000001e-06  evaluation reward: 6.52
episode: 2351  score: 5.0  memory length: 556641  epsilon: 0.5479244200098141
steps: 312  lr: 2.5600000000000001e-06  evaluation reward: 6.48
episode: 2352  score: 7.0  memory length: 557050  epsilon: 0.5475195100098229
steps: 409  lr: 2.5600000000000001e-06  evaluation reward: 6.4
episode: 2353  score: 8.0  memory length: 557458  epsilon: 0.5471155900098317
steps: 408  lr: 2.5600000000000001e-06  evaluation reward: 6.43
episode: 2354  score: 8.0  memory length: 557899  epsilon: 0.5466790000098412
steps: 441  lr: 2.5600000000000001e-06  evaluation reward: 6.46
episode: 2355  score: 12.0  memory length: 558390  epsilon:
0.5461929100098517  steps: 491  lr: 2.5600000000000001e-06  evaluation
reward: 6.48
episode: 2356  score: 6.0  memory length: 558762  epsilon: 0.5458246300098597
steps: 372  lr: 2.5600000000000001e-06  evaluation reward: 6.46
episode: 2357  score: 6.0  memory length: 559153  epsilon: 0.5454375400098681
steps: 391  lr: 2.5600000000000001e-06  evaluation reward: 6.48
episode: 2358  score: 7.0  memory length: 559526  epsilon: 0.5450682700098761
steps: 373  lr: 2.5600000000000001e-06  evaluation reward: 6.52
episode: 2359  score: 5.0  memory length: 559818  epsilon: 0.5447791900098824
steps: 292  lr: 2.5600000000000001e-06  evaluation reward: 6.53
episode: 2360  score: 3.0  memory length: 560029  epsilon: 0.5445703000098869
steps: 211  lr: 2.5600000000000001e-06  evaluation reward: 6.51
episode: 2361  score: 7.0  memory length: 560452  epsilon: 0.544151530009896
steps: 423  lr: 2.5600000000000001e-06  evaluation reward: 6.49
episode: 2362  score: 5.0  memory length: 560783  epsilon: 0.5438238400099031
steps: 331  lr: 2.5600000000000001e-06  evaluation reward: 6.42
episode: 2363  score: 10.0  memory length: 561301  epsilon:
0.5433110200099143  steps: 518  lr: 2.5600000000000001e-06  evaluation
reward: 6.46
episode: 2364  score: 3.0  memory length: 561531  epsilon: 0.5430833200099192
steps: 230  lr: 2.5600000000000001e-06  evaluation reward: 6.39
episode: 2365  score: 3.0  memory length: 561798  epsilon: 0.542818990009925
steps: 267  lr: 2.5600000000000001e-06  evaluation reward: 6.36
episode: 2366  score: 9.0  memory length: 562249  epsilon: 0.5423725000099346
steps: 451  lr: 2.5600000000000001e-06  evaluation reward: 6.41
episode: 2367  score: 5.0  memory length: 562538  epsilon: 0.5420863900099409
steps: 289  lr: 2.5600000000000001e-06  evaluation reward: 6.41
episode: 2368  score: 3.0  memory length: 562786  epsilon: 0.5418408700099462
steps: 248  lr: 2.5600000000000001e-06  evaluation reward: 6.41
episode: 2369  score: 5.0  memory length: 563094  epsilon: 0.5415359500099528
steps: 308  lr: 2.5600000000000001e-06  evaluation reward: 6.37
episode: 2370  score: 7.0  memory length: 563522  epsilon: 0.541112230009962
steps: 428  lr: 2.5600000000000001e-06  evaluation reward: 6.42
episode: 2371  score: 9.0  memory length: 563961  epsilon: 0.5406776200099714

```

```

steps: 439    lr: 2.5600000000000001e-06    evaluation reward: 6.43
episode: 2372  score: 9.0    memory length: 564284    epsilon: 0.5403578500099784
steps: 323    lr: 2.5600000000000001e-06    evaluation reward: 6.49
episode: 2373  score: 3.0    memory length: 564511    epsilon: 0.5401331200099833
steps: 227    lr: 2.5600000000000001e-06    evaluation reward: 6.42
episode: 2374  score: 8.0    memory length: 564983    epsilon: 0.5396658400099934
steps: 472    lr: 2.5600000000000001e-06    evaluation reward: 6.42
episode: 2375  score: 4.0    memory length: 565244    epsilon: 0.539407450009999
steps: 261    lr: 2.5600000000000001e-06    evaluation reward: 6.35
episode: 2376  score: 6.0    memory length: 565581    epsilon: 0.5390738200100063
steps: 337    lr: 2.5600000000000001e-06    evaluation reward: 6.33
episode: 2377  score: 6.0    memory length: 565960    epsilon: 0.5386986100100144
steps: 379    lr: 2.5600000000000001e-06    evaluation reward: 6.32
episode: 2378  score: 9.0    memory length: 566460    epsilon: 0.5382036100100251
steps: 500    lr: 2.5600000000000001e-06    evaluation reward: 6.35
episode: 2379  score: 14.0   memory length: 567009    epsilon: 0.537660100010037
steps: 549    lr: 2.5600000000000001e-06    evaluation reward: 6.43
episode: 2380  score: 6.0    memory length: 567354    epsilon: 0.5373185500100444
steps: 345    lr: 2.5600000000000001e-06    evaluation reward: 6.44
episode: 2381  score: 9.0    memory length: 567802    epsilon: 0.536875030010054
steps: 448    lr: 2.5600000000000001e-06    evaluation reward: 6.47
episode: 2382  score: 3.0    memory length: 568013    epsilon: 0.5366661400100585
steps: 211    lr: 2.5600000000000001e-06    evaluation reward: 6.42
episode: 2383  score: 6.0    memory length: 568357    epsilon: 0.5363255800100659
steps: 344    lr: 2.5600000000000001e-06    evaluation reward: 6.41
episode: 2384  score: 14.0   memory length: 568876    epsilon:
0.5358117700100771    steps: 519    lr: 2.5600000000000001e-06    evaluation
reward: 6.5
episode: 2385  score: 6.0    memory length: 569215    epsilon: 0.5354761600100844
steps: 339    lr: 2.5600000000000001e-06    evaluation reward: 6.5
episode: 2386  score: 10.0   memory length: 569716    epsilon:
0.5349801700100951    steps: 501    lr: 2.5600000000000001e-06    evaluation
reward: 6.54
episode: 2387  score: 2.0    memory length: 569898    epsilon: 0.534799990010099
steps: 182    lr: 2.5600000000000001e-06    evaluation reward: 6.47
episode: 2388  score: 7.0    memory length: 570323    epsilon: 0.5343792400101082
steps: 425    lr: 2.5600000000000001e-06    evaluation reward: 6.49
episode: 2389  score: 7.0    memory length: 570739    epsilon: 0.5339674000101171
steps: 416    lr: 2.5600000000000001e-06    evaluation reward: 6.53
episode: 2390  score: 6.0    memory length: 571074    epsilon: 0.5336357500101243
steps: 335    lr: 2.5600000000000001e-06    evaluation reward: 6.54
episode: 2391  score: 11.0   memory length: 571597    epsilon:
0.5331179800101356    steps: 523    lr: 2.5600000000000001e-06    evaluation
reward: 6.59
episode: 2392  score: 14.0   memory length: 572154    epsilon:
0.5325665500101475    steps: 557    lr: 2.5600000000000001e-06    evaluation
reward: 6.65
episode: 2393  score: 6.0    memory length: 572474    epsilon: 0.5322497500101544

```

```

steps: 320    lr: 2.5600000000000001e-06    evaluation reward: 6.67
episode: 2394    score: 7.0    memory length: 572901    epsilon: 0.5318270200101636
steps: 427    lr: 2.5600000000000001e-06    evaluation reward: 6.63
episode: 2395    score: 3.0    memory length: 573127    epsilon: 0.5316032800101684
steps: 226    lr: 2.5600000000000001e-06    evaluation reward: 6.6
episode: 2396    score: 6.0    memory length: 573519    epsilon: 0.5312152000101769
steps: 392    lr: 2.5600000000000001e-06    evaluation reward: 6.6
episode: 2397    score: 7.0    memory length: 573923    epsilon: 0.5308152400101855
steps: 404    lr: 2.5600000000000001e-06    evaluation reward: 6.6
episode: 2398    score: 5.0    memory length: 574249    epsilon: 0.5304925000101925
steps: 326    lr: 2.5600000000000001e-06    evaluation reward: 6.58
episode: 2399    score: 9.0    memory length: 574730    epsilon: 0.5300163100102029
steps: 481    lr: 2.5600000000000001e-06    evaluation reward: 6.61
episode: 2400    score: 11.0    memory length: 575286    epsilon:
0.5294658700102148    steps: 556    lr: 2.5600000000000001e-06    evaluation
reward: 6.66
episode: 2401    score: 6.0    memory length: 575648    epsilon: 0.5291074900102226
steps: 362    lr: 2.5600000000000001e-06    evaluation reward: 6.66
episode: 2402    score: 6.0    memory length: 576012    epsilon: 0.5287471300102304
steps: 364    lr: 2.5600000000000001e-06    evaluation reward: 6.68
episode: 2403    score: 20.0    memory length: 576523    epsilon:
0.5282412400102414    steps: 511    lr: 2.5600000000000001e-06    evaluation
reward: 6.84
episode: 2404    score: 7.0    memory length: 576894    epsilon: 0.5278739500102494
steps: 371    lr: 2.5600000000000001e-06    evaluation reward: 6.84
episode: 2405    score: 9.0    memory length: 577367    epsilon: 0.5274056800102596
steps: 473    lr: 2.5600000000000001e-06    evaluation reward: 6.89
episode: 2406    score: 4.0    memory length: 577643    epsilon: 0.5271324400102655
steps: 276    lr: 2.5600000000000001e-06    evaluation reward: 6.86
episode: 2407    score: 3.0    memory length: 577912    epsilon: 0.5268661300102713
steps: 269    lr: 2.5600000000000001e-06    evaluation reward: 6.8
episode: 2408    score: 16.0    memory length: 578406    epsilon:
0.5263770700102819    steps: 494    lr: 2.5600000000000001e-06    evaluation
reward: 6.91
episode: 2409    score: 7.0    memory length: 578811    epsilon: 0.5259761200102906
steps: 405    lr: 2.5600000000000001e-06    evaluation reward: 6.88
episode: 2410    score: 8.0    memory length: 579211    epsilon: 0.5255801200102992
steps: 400    lr: 2.5600000000000001e-06    evaluation reward: 6.91
episode: 2411    score: 9.0    memory length: 579686    epsilon: 0.5251098700103094
steps: 475    lr: 2.5600000000000001e-06    evaluation reward: 6.96
episode: 2412    score: 8.0    memory length: 580154    epsilon: 0.5246465500103195
steps: 468    lr: 2.5600000000000001e-06    evaluation reward: 6.98
episode: 2413    score: 10.0    memory length: 580672    epsilon:
0.5241337300103306    steps: 518    lr: 2.5600000000000001e-06    evaluation
reward: 7.01
episode: 2414    score: 8.0    memory length: 581109    epsilon: 0.52370110001034
steps: 437    lr: 2.5600000000000001e-06    evaluation reward: 7.06
episode: 2415    score: 8.0    memory length: 581548    epsilon: 0.5232664900103494

```

```

steps: 439    lr: 2.5600000000000001e-06    evaluation reward: 7.06
episode: 2416  score: 3.0    memory length: 581774    epsilon: 0.5230427500103543
steps: 226    lr: 2.5600000000000001e-06    evaluation reward: 7.0
episode: 2417  score: 5.0    memory length: 582094    epsilon: 0.5227259500103612
steps: 320    lr: 2.5600000000000001e-06    evaluation reward: 7.01
episode: 2418  score: 6.0    memory length: 582405    epsilon: 0.5224180600103678
steps: 311    lr: 2.5600000000000001e-06    evaluation reward: 7.04
episode: 2419  score: 5.0    memory length: 582729    epsilon: 0.5220973000103748
steps: 324    lr: 2.5600000000000001e-06    evaluation reward: 7.05
episode: 2420  score: 9.0    memory length: 583180    epsilon: 0.5216508100103845
steps: 451    lr: 2.5600000000000001e-06    evaluation reward: 7.07
episode: 2421  score: 6.0    memory length: 583526    epsilon: 0.5213082700103919
steps: 346    lr: 2.5600000000000001e-06    evaluation reward: 7.04
episode: 2422  score: 7.0    memory length: 583953    epsilon: 0.5208855400104011
steps: 427    lr: 2.5600000000000001e-06    evaluation reward: 7.06
episode: 2423  score: 7.0    memory length: 584300    epsilon: 0.5205420100104086
steps: 347    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2424  score: 11.0   memory length: 584808    epsilon:
0.5200390900104195    steps: 508    lr: 2.5600000000000001e-06    evaluation
reward: 7.14
episode: 2425  score: 8.0    memory length: 585277    epsilon: 0.5195747800104296
steps: 469    lr: 2.5600000000000001e-06    evaluation reward: 7.15
episode: 2426  score: 10.0   memory length: 585740    epsilon:
0.5191164100104395    steps: 463    lr: 2.5600000000000001e-06    evaluation
reward: 7.16
episode: 2427  score: 6.0    memory length: 586114    epsilon: 0.5187461500104475
steps: 374    lr: 2.5600000000000001e-06    evaluation reward: 7.14
episode: 2428  score: 4.0    memory length: 586358    epsilon: 0.5185045900104528
steps: 244    lr: 2.5600000000000001e-06    evaluation reward: 7.14
episode: 2429  score: 9.0    memory length: 586858    epsilon: 0.5180095900104635
steps: 500    lr: 2.5600000000000001e-06    evaluation reward: 7.18
episode: 2430  score: 5.0    memory length: 587184    epsilon: 0.5176868500104705
steps: 326    lr: 2.5600000000000001e-06    evaluation reward: 7.14
episode: 2431  score: 6.0    memory length: 587557    epsilon: 0.5173175800104786
steps: 373    lr: 2.5600000000000001e-06    evaluation reward: 7.16
episode: 2432  score: 9.0    memory length: 588018    epsilon: 0.5168611900104885
steps: 461    lr: 2.5600000000000001e-06    evaluation reward: 7.14
episode: 2433  score: 3.0    memory length: 588265    epsilon: 0.5166166600104938
steps: 247    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2434  score: 4.0    memory length: 588520    epsilon: 0.5163642100104993
steps: 255    lr: 2.5600000000000001e-06    evaluation reward: 7.05
episode: 2435  score: 5.0    memory length: 588811    epsilon: 0.5160761200105055
steps: 291    lr: 2.5600000000000001e-06    evaluation reward: 7.06
episode: 2436  score: 1.0    memory length: 588962    epsilon: 0.5159266300105088
steps: 151    lr: 2.5600000000000001e-06    evaluation reward: 7.0
episode: 2437  score: 7.0    memory length: 589306    epsilon: 0.5155860700105162
steps: 344    lr: 2.5600000000000001e-06    evaluation reward: 7.02
episode: 2438  score: 6.0    memory length: 589681    epsilon: 0.5152148200105242

```



```

steps: 375    lr: 2.5600000000000001e-06    evaluation reward: 6.98
episode: 2439    score: 8.0    memory length: 590085    epsilon: 0.5148148600105329
steps: 404    lr: 2.5600000000000001e-06    evaluation reward: 6.99
episode: 2440    score: 11.0    memory length: 590649    epsilon: 0.514256500010545
steps: 564    lr: 2.5600000000000001e-06    evaluation reward: 7.06
episode: 2441    score: 11.0    memory length: 591209    epsilon: 0.513702100010557
steps: 560    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2442    score: 4.0    memory length: 591451    epsilon: 0.5134625200105623
steps: 242    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2443    score: 4.0    memory length: 591693    epsilon: 0.5132229400105675
steps: 242    lr: 2.5600000000000001e-06    evaluation reward: 7.07
episode: 2444    score: 7.0    memory length: 592064    epsilon: 0.5128556500105754
steps: 371    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2445    score: 8.0    memory length: 592494    epsilon: 0.5124299500105847
steps: 430    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2446    score: 5.0    memory length: 592800    epsilon: 0.5121270100105912
steps: 306    lr: 2.5600000000000001e-06    evaluation reward: 7.07
episode: 2447    score: 4.0    memory length: 593042    epsilon: 0.5118874300105964
steps: 242    lr: 2.5600000000000001e-06    evaluation reward: 7.04
episode: 2448    score: 7.0    memory length: 593426    epsilon: 0.5115072700106047
steps: 384    lr: 2.5600000000000001e-06    evaluation reward: 7.0
episode: 2449    score: 6.0    memory length: 593778    epsilon: 0.5111587900106123
steps: 352    lr: 2.5600000000000001e-06    evaluation reward: 6.98
episode: 2450    score: 15.0    memory length: 594344    epsilon:
0.5105984500106244    steps: 566    lr: 2.5600000000000001e-06    evaluation
reward: 7.08
episode: 2451    score: 5.0    memory length: 594618    epsilon: 0.5103271900106303
steps: 274    lr: 2.5600000000000001e-06    evaluation reward: 7.08
episode: 2452    score: 8.0    memory length: 595082    epsilon: 0.5098678300106403
steps: 464    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2453    score: 13.0    memory length: 595659    epsilon:
0.5092966000106527    steps: 577    lr: 2.5600000000000001e-06    evaluation
reward: 7.14
episode: 2454    score: 4.0    memory length: 595934    epsilon: 0.5090243500106586
steps: 275    lr: 2.5600000000000001e-06    evaluation reward: 7.1
episode: 2455    score: 5.0    memory length: 596246    epsilon: 0.5087154700106653
steps: 312    lr: 2.5600000000000001e-06    evaluation reward: 7.03
episode: 2456    score: 8.0    memory length: 596642    epsilon: 0.5083234300106738
steps: 396    lr: 2.5600000000000001e-06    evaluation reward: 7.05
episode: 2457    score: 8.0    memory length: 597043    epsilon: 0.5079264400106824
steps: 401    lr: 2.5600000000000001e-06    evaluation reward: 7.07
episode: 2458    score: 9.0    memory length: 597540    epsilon: 0.5074344100106931
steps: 497    lr: 2.5600000000000001e-06    evaluation reward: 7.09
episode: 2459    score: 8.0    memory length: 597956    epsilon: 0.5070225700107021
steps: 416    lr: 2.5600000000000001e-06    evaluation reward: 7.12
episode: 2460    score: 6.0    memory length: 598292    epsilon: 0.5066899300107093
steps: 336    lr: 2.5600000000000001e-06    evaluation reward: 7.15
episode: 2461    score: 7.0    memory length: 598685    epsilon: 0.5063008600107177

```

```

steps: 393    lr: 2.5600000000000001e-06    evaluation reward: 7.15
episode: 2462    score: 11.0    memory length: 599231    epsilon:
0.5057603200107295    steps: 546    lr: 2.5600000000000001e-06    evaluation
reward: 7.21
episode: 2463    score: 8.0    memory length: 599662    epsilon: 0.5053336300107387
steps: 431    lr: 2.5600000000000001e-06    evaluation reward: 7.19
episode: 2464    score: 9.0    memory length: 600115    epsilon: 0.5048851600107485
steps: 453    lr: 1.0240000000000005e-06    evaluation reward: 7.25
episode: 2465    score: 6.0    memory length: 600472    epsilon: 0.5045317300107561
steps: 357    lr: 1.0240000000000005e-06    evaluation reward: 7.28
episode: 2466    score: 8.0    memory length: 600949    epsilon: 0.5040595000107664
steps: 477    lr: 1.0240000000000005e-06    evaluation reward: 7.27
episode: 2467    score: 10.0    memory length: 601470    epsilon:
0.5035437100107776    steps: 521    lr: 1.0240000000000005e-06    evaluation
reward: 7.32
episode: 2468    score: 3.0    memory length: 601714    epsilon: 0.5033021500107828
steps: 244    lr: 1.0240000000000005e-06    evaluation reward: 7.32
episode: 2469    score: 10.0    memory length: 602253    epsilon:
0.5027685400107944    steps: 539    lr: 1.0240000000000005e-06    evaluation
reward: 7.37
episode: 2470    score: 9.0    memory length: 602712    epsilon: 0.5023141300108043
steps: 459    lr: 1.0240000000000005e-06    evaluation reward: 7.39
episode: 2471    score: 7.0    memory length: 603137    epsilon: 0.5018933800108134
steps: 425    lr: 1.0240000000000005e-06    evaluation reward: 7.37
episode: 2472    score: 8.0    memory length: 603574    epsilon: 0.5014607500108228
steps: 437    lr: 1.0240000000000005e-06    evaluation reward: 7.36
episode: 2473    score: 18.0    memory length: 604097    epsilon: 0.500942980010834
steps: 523    lr: 1.0240000000000005e-06    evaluation reward: 7.51
episode: 2474    score: 7.0    memory length: 604483    epsilon: 0.5005608400108423
steps: 386    lr: 1.0240000000000005e-06    evaluation reward: 7.5
episode: 2475    score: 6.0    memory length: 604822    epsilon: 0.5002252300108496
steps: 339    lr: 1.0240000000000005e-06    evaluation reward: 7.52
episode: 2476    score: 9.0    memory length: 605296    epsilon: 0.4997559700108598
steps: 474    lr: 1.0240000000000005e-06    evaluation reward: 7.55
episode: 2477    score: 3.0    memory length: 605546    epsilon: 0.4995084700108652
steps: 250    lr: 1.0240000000000005e-06    evaluation reward: 7.52
episode: 2478    score: 8.0    memory length: 605988    epsilon: 0.4990708900108747
steps: 442    lr: 1.0240000000000005e-06    evaluation reward: 7.51
episode: 2479    score: 9.0    memory length: 606461    epsilon:
0.49860262001088484    steps: 473    lr: 1.0240000000000005e-06    evaluation
reward: 7.46
episode: 2480    score: 11.0    memory length: 606969    epsilon:
0.49809970001089576    steps: 508    lr: 1.0240000000000005e-06    evaluation
reward: 7.51
episode: 2481    score: 4.0    memory length: 607210    epsilon:
0.49786111001090094    steps: 241    lr: 1.0240000000000005e-06    evaluation
reward: 7.46
episode: 2482    score: 6.0    memory length: 607569    epsilon:

```

0.49750570001090866 steps: 359 lr: 1.0240000000000005e-06 evaluation
 reward: 7.49
 episode: 2483 score: 10.0 memory length: 607961 epsilon:
 0.4971176200109171 steps: 392 lr: 1.0240000000000005e-06 evaluation
 reward: 7.53
 episode: 2484 score: 8.0 memory length: 608377 epsilon: 0.496705780010926
 steps: 416 lr: 1.0240000000000005e-06 evaluation reward: 7.47
 episode: 2485 score: 11.0 memory length: 608896 epsilon:
 0.4961919700109372 steps: 519 lr: 1.0240000000000005e-06 evaluation
 reward: 7.52
 episode: 2486 score: 4.0 memory length: 609176 epsilon: 0.4959147700109432
 steps: 280 lr: 1.0240000000000005e-06 evaluation reward: 7.46
 episode: 2487 score: 9.0 memory length: 609655 epsilon: 0.4954405600109535
 steps: 479 lr: 1.0240000000000005e-06 evaluation reward: 7.53
 episode: 2488 score: 12.0 memory length: 610249 epsilon:
 0.49485250001096626 steps: 594 lr: 1.0240000000000005e-06 evaluation
 reward: 7.58
 episode: 2489 score: 8.0 memory length: 610699 epsilon: 0.4944070000109759
 steps: 450 lr: 1.0240000000000005e-06 evaluation reward: 7.59
 episode: 2490 score: 7.0 memory length: 611073 epsilon:
 0.49403674001098397 steps: 374 lr: 1.0240000000000005e-06 evaluation
 reward: 7.6
 episode: 2491 score: 8.0 memory length: 611523 epsilon:
 0.49359124001099364 steps: 450 lr: 1.0240000000000005e-06 evaluation
 reward: 7.57
 episode: 2492 score: 7.0 memory length: 611932 epsilon: 0.4931863300110024
 steps: 409 lr: 1.0240000000000005e-06 evaluation reward: 7.5
 episode: 2493 score: 10.0 memory length: 612444 epsilon:
 0.49267945001101343 steps: 512 lr: 1.0240000000000005e-06 evaluation
 reward: 7.54
 episode: 2494 score: 12.0 memory length: 613027 epsilon:
 0.49210228001102596 steps: 583 lr: 1.0240000000000005e-06 evaluation
 reward: 7.59
 episode: 2495 score: 14.0 memory length: 613538 epsilon:
 0.49159639001103694 steps: 511 lr: 1.0240000000000005e-06 evaluation
 reward: 7.7
 episode: 2496 score: 10.0 memory length: 614060 epsilon:
 0.49107961001104816 steps: 522 lr: 1.0240000000000005e-06 evaluation
 reward: 7.74
 episode: 2497 score: 9.0 memory length: 614605 epsilon: 0.4905400600110599
 steps: 545 lr: 1.0240000000000005e-06 evaluation reward: 7.76
 episode: 2498 score: 11.0 memory length: 615182 epsilon:
 0.4899688300110723 steps: 577 lr: 1.0240000000000005e-06 evaluation
 reward: 7.82
 episode: 2499 score: 7.0 memory length: 615643 epsilon: 0.4895124400110822
 steps: 461 lr: 1.0240000000000005e-06 evaluation reward: 7.8
 episode: 2500 score: 20.0 memory length: 616440 epsilon:
 0.4887234100110993 steps: 797 lr: 1.0240000000000005e-06 evaluation

```

reward: 7.89
episode: 2501  score: 8.0  memory length: 616847  epsilon:
0.48832048001110806  steps: 407  lr: 1.0240000000000005e-06  evaluation
reward: 7.91
episode: 2502  score: 8.0  memory length: 617323  epsilon: 0.4878492400111183
steps: 476  lr: 1.0240000000000005e-06  evaluation reward: 7.93
episode: 2503  score: 6.0  memory length: 617663  epsilon: 0.4875126400111256
steps: 340  lr: 1.0240000000000005e-06  evaluation reward: 7.79
episode: 2504  score: 3.0  memory length: 617874  epsilon:
0.48730375001113013  steps: 211  lr: 1.0240000000000005e-06  evaluation
reward: 7.75
episode: 2505  score: 8.0  memory length: 618307  epsilon:
0.48687508001113944  steps: 433  lr: 1.0240000000000005e-06  evaluation
reward: 7.74
episode: 2506  score: 9.0  memory length: 618812  epsilon: 0.4863751300111503
steps: 505  lr: 1.0240000000000005e-06  evaluation reward: 7.79
episode: 2507  score: 8.0  memory length: 619198  epsilon: 0.4859929900111586
steps: 386  lr: 1.0240000000000005e-06  evaluation reward: 7.84
episode: 2508  score: 6.0  memory length: 619533  epsilon: 0.4856613400111658
steps: 335  lr: 1.0240000000000005e-06  evaluation reward: 7.74
episode: 2509  score: 11.0  memory length: 620099  epsilon:
0.48510100001117795  steps: 566  lr: 1.0240000000000005e-06  evaluation
reward: 7.78
episode: 2510  score: 10.0  memory length: 620592  epsilon:
0.48461293001118855  steps: 493  lr: 1.0240000000000005e-06  evaluation
reward: 7.8
episode: 2511  score: 14.0  memory length: 621167  epsilon:
0.4840436800112009  steps: 575  lr: 1.0240000000000005e-06  evaluation
reward: 7.85
episode: 2512  score: 15.0  memory length: 621719  epsilon:
0.48349720001121277  steps: 552  lr: 1.0240000000000005e-06  evaluation
reward: 7.92
episode: 2513  score: 9.0  memory length: 622219  epsilon: 0.4830022000112235
steps: 500  lr: 1.0240000000000005e-06  evaluation reward: 7.91
episode: 2514  score: 11.0  memory length: 622679  epsilon:
0.4825468000112334  steps: 460  lr: 1.0240000000000005e-06  evaluation
reward: 7.94
episode: 2515  score: 6.0  memory length: 623017  epsilon:
0.48221218001124067  steps: 338  lr: 1.0240000000000005e-06  evaluation
reward: 7.92
episode: 2516  score: 4.0  memory length: 623294  epsilon: 0.4819379500112466
steps: 277  lr: 1.0240000000000005e-06  evaluation reward: 7.93
episode: 2517  score: 8.0  memory length: 623679  epsilon: 0.4815568000112549
steps: 385  lr: 1.0240000000000005e-06  evaluation reward: 7.96
episode: 2518  score: 8.0  memory length: 624101  epsilon:
0.48113902001126396  steps: 422  lr: 1.0240000000000005e-06  evaluation
reward: 7.98
episode: 2519  score: 8.0  memory length: 624573  epsilon: 0.4806717400112741

```

steps: 472 lr: 1.0240000000000005e-06 evaluation reward: 8.01
 episode: 2520 score: 9.0 memory length: 625030 epsilon:
 0.48021931001128393 steps: 457 lr: 1.0240000000000005e-06 evaluation
 reward: 8.01
 episode: 2521 score: 8.0 memory length: 625477 epsilon:
 0.47977678001129354 steps: 447 lr: 1.0240000000000005e-06 evaluation
 reward: 8.03
 episode: 2522 score: 7.0 memory length: 625879 epsilon: 0.4793788000113022
 steps: 402 lr: 1.0240000000000005e-06 evaluation reward: 8.03
 episode: 2523 score: 5.0 memory length: 626205 epsilon: 0.4790560600113092
 steps: 326 lr: 1.0240000000000005e-06 evaluation reward: 8.01
 episode: 2524 score: 4.0 memory length: 626463 epsilon: 0.4788006400113147
 steps: 258 lr: 1.0240000000000005e-06 evaluation reward: 7.94
 episode: 2525 score: 5.0 memory length: 626752 epsilon:
 0.47851453001132094 steps: 289 lr: 1.0240000000000005e-06 evaluation
 reward: 7.91
 episode: 2526 score: 10.0 memory length: 627281 epsilon:
 0.4779908200113323 steps: 529 lr: 1.0240000000000005e-06 evaluation
 reward: 7.91
 episode: 2527 score: 11.0 memory length: 627883 epsilon:
 0.47739484001134525 steps: 602 lr: 1.0240000000000005e-06 evaluation
 reward: 7.96
 episode: 2528 score: 8.0 memory length: 628293 epsilon:
 0.47698894001135406 steps: 410 lr: 1.0240000000000005e-06 evaluation
 reward: 8.0
 episode: 2529 score: 9.0 memory length: 628768 epsilon:
 0.47651869001136427 steps: 475 lr: 1.0240000000000005e-06 evaluation
 reward: 8.0
 episode: 2530 score: 11.0 memory length: 629172 epsilon:
 0.47611873001137295 steps: 404 lr: 1.0240000000000005e-06 evaluation
 reward: 8.06
 episode: 2531 score: 16.0 memory length: 629675 epsilon:
 0.47562076001138376 steps: 503 lr: 1.0240000000000005e-06 evaluation
 reward: 8.16
 episode: 2532 score: 6.0 memory length: 630034 epsilon: 0.4752653500113915
 steps: 359 lr: 1.0240000000000005e-06 evaluation reward: 8.13
 episode: 2533 score: 8.0 memory length: 630335 epsilon:
 0.47496736001139794 steps: 301 lr: 1.0240000000000005e-06 evaluation
 reward: 8.18
 episode: 2534 score: 7.0 memory length: 630728 epsilon: 0.4745782900114064
 steps: 393 lr: 1.0240000000000005e-06 evaluation reward: 8.21
 episode: 2535 score: 7.0 memory length: 631088 epsilon: 0.4742218900114141
 steps: 360 lr: 1.0240000000000005e-06 evaluation reward: 8.23
 episode: 2536 score: 12.0 memory length: 631655 epsilon:
 0.4736605600114263 steps: 567 lr: 1.0240000000000005e-06 evaluation
 reward: 8.34
 episode: 2537 score: 15.0 memory length: 632286 epsilon:
 0.4730358700114399 steps: 631 lr: 1.0240000000000005e-06 evaluation

```

reward: 8.42
episode: 2538   score: 13.0   memory length: 632798   epsilon:
0.4725289900114509   steps: 512   lr: 1.0240000000000005e-06   evaluation
reward: 8.49
episode: 2539   score: 7.0   memory length: 633168   epsilon:
0.47216269001145883   steps: 370   lr: 1.0240000000000005e-06   evaluation
reward: 8.48
episode: 2540   score: 4.0   memory length: 633461   epsilon: 0.4718726200114651
steps: 293   lr: 1.0240000000000005e-06   evaluation reward: 8.41
episode: 2541   score: 4.0   memory length: 633720   epsilon: 0.4716162100114707
steps: 259   lr: 1.0240000000000005e-06   evaluation reward: 8.34
episode: 2542   score: 4.0   memory length: 633962   epsilon: 0.4713766300114759
steps: 242   lr: 1.0240000000000005e-06   evaluation reward: 8.34
episode: 2543   score: 7.0   memory length: 634330   epsilon: 0.4710123100114838
steps: 368   lr: 1.0240000000000005e-06   evaluation reward: 8.37
episode: 2544   score: 5.0   memory length: 634641   epsilon: 0.4707044200114905
steps: 311   lr: 1.0240000000000005e-06   evaluation reward: 8.35
episode: 2545   score: 3.0   memory length: 634893   epsilon: 0.4704549400114959
steps: 252   lr: 1.0240000000000005e-06   evaluation reward: 8.3
episode: 2546   score: 5.0   memory length: 635187   epsilon: 0.4701638800115022
steps: 294   lr: 1.0240000000000005e-06   evaluation reward: 8.3
episode: 2547   score: 7.0   memory length: 635542   epsilon:
0.46981243001150985   steps: 355   lr: 1.0240000000000005e-06   evaluation
reward: 8.33
episode: 2548   score: 6.0   memory length: 635862   epsilon:
0.46949563001151673   steps: 320   lr: 1.0240000000000005e-06   evaluation
reward: 8.32
episode: 2549   score: 8.0   memory length: 636327   epsilon: 0.4690352800115267
steps: 465   lr: 1.0240000000000005e-06   evaluation reward: 8.34
episode: 2550   score: 11.0   memory length: 636883   epsilon:
0.4684848400115387   steps: 556   lr: 1.0240000000000005e-06   evaluation
reward: 8.3
episode: 2551   score: 7.0   memory length: 637253   epsilon: 0.4681185400115466
steps: 370   lr: 1.0240000000000005e-06   evaluation reward: 8.32
episode: 2552   score: 8.0   memory length: 637712   epsilon: 0.4676641300115565
steps: 459   lr: 1.0240000000000005e-06   evaluation reward: 8.32
episode: 2553   score: 6.0   memory length: 638107   epsilon: 0.467273080011565
steps: 395   lr: 1.0240000000000005e-06   evaluation reward: 8.25
episode: 2554   score: 8.0   memory length: 638532   epsilon: 0.4668523300115741
steps: 425   lr: 1.0240000000000005e-06   evaluation reward: 8.29
episode: 2555   score: 7.0   memory length: 638915   epsilon:
0.46647316001158234   steps: 383   lr: 1.0240000000000005e-06   evaluation
reward: 8.31
episode: 2556   score: 10.0   memory length: 639394   epsilon:
0.46599895001159264   steps: 479   lr: 1.0240000000000005e-06   evaluation
reward: 8.33
episode: 2557   score: 6.0   memory length: 639733   epsilon: 0.4656633400115999
steps: 339   lr: 1.0240000000000005e-06   evaluation reward: 8.31

```

episode: 2558 score: 4.0 memory length: 639993 epsilon: 0.4654059400116055
 steps: 260 lr: 1.0240000000000005e-06 evaluation reward: 8.26
 episode: 2559 score: 14.0 memory length: 640619 epsilon:
 0.46478620001161897 steps: 626 lr: 1.0240000000000005e-06 evaluation
 reward: 8.32
 episode: 2560 score: 7.0 memory length: 641023 epsilon:
 0.46438624001162765 steps: 404 lr: 1.0240000000000005e-06 evaluation
 reward: 8.33
 episode: 2561 score: 4.0 memory length: 641318 epsilon: 0.464094190011634
 steps: 295 lr: 1.0240000000000005e-06 evaluation reward: 8.3
 episode: 2562 score: 5.0 memory length: 641645 epsilon: 0.463770460011641
 steps: 327 lr: 1.0240000000000005e-06 evaluation reward: 8.24
 episode: 2563 score: 8.0 memory length: 642055 epsilon:
 0.46336456001164983 steps: 410 lr: 1.0240000000000005e-06 evaluation
 reward: 8.24
 episode: 2564 score: 6.0 memory length: 642427 epsilon: 0.4629962800116578
 steps: 372 lr: 1.0240000000000005e-06 evaluation reward: 8.21
 episode: 2565 score: 7.0 memory length: 642793 epsilon: 0.4626339400116657
 steps: 366 lr: 1.0240000000000005e-06 evaluation reward: 8.22
 episode: 2566 score: 10.0 memory length: 643355 epsilon:
 0.46207756001167777 steps: 562 lr: 1.0240000000000005e-06 evaluation
 reward: 8.24
 episode: 2567 score: 10.0 memory length: 643872 epsilon:
 0.4615657300116889 steps: 517 lr: 1.0240000000000005e-06 evaluation
 reward: 8.24
 episode: 2568 score: 6.0 memory length: 644248 epsilon:
 0.46119349001169696 steps: 376 lr: 1.0240000000000005e-06 evaluation
 reward: 8.27
 episode: 2569 score: 8.0 memory length: 644622 epsilon: 0.460823230011705
 steps: 374 lr: 1.0240000000000005e-06 evaluation reward: 8.25
 episode: 2570 score: 11.0 memory length: 645120 epsilon:
 0.4603302100117157 steps: 498 lr: 1.0240000000000005e-06 evaluation
 reward: 8.27
 episode: 2571 score: 6.0 memory length: 645459 epsilon: 0.459994600011723
 steps: 339 lr: 1.0240000000000005e-06 evaluation reward: 8.26
 episode: 2572 score: 12.0 memory length: 645929 epsilon:
 0.4595293000117331 steps: 470 lr: 1.0240000000000005e-06 evaluation
 reward: 8.3
 episode: 2573 score: 8.0 memory length: 646343 epsilon: 0.459119440011742
 steps: 414 lr: 1.0240000000000005e-06 evaluation reward: 8.2
 episode: 2574 score: 4.0 memory length: 646584 epsilon:
 0.45888085001174717 steps: 241 lr: 1.0240000000000005e-06 evaluation
 reward: 8.17
 episode: 2575 score: 11.0 memory length: 647113 epsilon:
 0.45835714001175853 steps: 529 lr: 1.0240000000000005e-06 evaluation
 reward: 8.22
 episode: 2576 score: 7.0 memory length: 647502 epsilon: 0.4579720300117669
 steps: 389 lr: 1.0240000000000005e-06 evaluation reward: 8.2

episode: 2577 score: 15.0 memory length: 648068 epsilon:
 0.45741169001177906 steps: 566 lr: 1.0240000000000005e-06 evaluation
 reward: 8.32
 episode: 2578 score: 8.0 memory length: 648508 epsilon: 0.4569760900117885
 steps: 440 lr: 1.0240000000000005e-06 evaluation reward: 8.32
 episode: 2579 score: 11.0 memory length: 649046 epsilon:
 0.4564434700118001 steps: 538 lr: 1.0240000000000005e-06 evaluation
 reward: 8.34
 episode: 2580 score: 10.0 memory length: 649510 epsilon:
 0.45598411001181005 steps: 464 lr: 1.0240000000000005e-06 evaluation
 reward: 8.33
 episode: 2581 score: 9.0 memory length: 649981 epsilon: 0.4555178200118202
 steps: 471 lr: 1.0240000000000005e-06 evaluation reward: 8.38
 episode: 2582 score: 8.0 memory length: 650439 epsilon: 0.45506440001183
 steps: 458 lr: 1.0240000000000005e-06 evaluation reward: 8.4
 episode: 2583 score: 7.0 memory length: 650825 epsilon: 0.4546822600118383
 steps: 386 lr: 1.0240000000000005e-06 evaluation reward: 8.37
 episode: 2584 score: 8.0 memory length: 651262 epsilon: 0.4542496300118477
 steps: 437 lr: 1.0240000000000005e-06 evaluation reward: 8.37
 episode: 2585 score: 9.0 memory length: 651722 epsilon: 0.4537942300118576
 steps: 460 lr: 1.0240000000000005e-06 evaluation reward: 8.35
 episode: 2586 score: 3.0 memory length: 651931 epsilon: 0.4535873200118621
 steps: 209 lr: 1.0240000000000005e-06 evaluation reward: 8.34
 episode: 2587 score: 8.0 memory length: 652353 epsilon:
 0.45316954001187115 steps: 422 lr: 1.0240000000000005e-06 evaluation
 reward: 8.33
 episode: 2588 score: 10.0 memory length: 652856 epsilon:
 0.45267157001188196 steps: 503 lr: 1.0240000000000005e-06 evaluation
 reward: 8.31
 episode: 2589 score: 11.0 memory length: 653423 epsilon:
 0.45211024001189415 steps: 567 lr: 1.0240000000000005e-06 evaluation
 reward: 8.34
 episode: 2590 score: 11.0 memory length: 653827 epsilon:
 0.45171028001190283 steps: 404 lr: 1.0240000000000005e-06 evaluation
 reward: 8.38
 episode: 2591 score: 8.0 memory length: 654253 epsilon: 0.451288540011912
 steps: 426 lr: 1.0240000000000005e-06 evaluation reward: 8.38
 episode: 2592 score: 14.0 memory length: 654832 epsilon:
 0.45071533001192443 steps: 579 lr: 1.0240000000000005e-06 evaluation
 reward: 8.45
 episode: 2593 score: 8.0 memory length: 655237 epsilon:
 0.45031438001193314 steps: 405 lr: 1.0240000000000005e-06 evaluation
 reward: 8.43
 episode: 2594 score: 13.0 memory length: 655760 epsilon:
 0.4497966100119444 steps: 523 lr: 1.0240000000000005e-06 evaluation
 reward: 8.44
 episode: 2595 score: 6.0 memory length: 656113 epsilon:
 0.44944714001195196 steps: 353 lr: 1.0240000000000005e-06 evaluation


```

reward: 8.36
episode: 2596   score: 6.0   memory length: 656475   epsilon:
0.44908876001195974   steps: 362   lr: 1.0240000000000005e-06   evaluation
reward: 8.32
episode: 2597   score: 12.0  memory length: 656915   epsilon:
0.4486531600119692   steps: 440   lr: 1.0240000000000005e-06   evaluation
reward: 8.35
episode: 2598   score: 13.0  memory length: 657414   epsilon:
0.4481591500119799   steps: 499   lr: 1.0240000000000005e-06   evaluation
reward: 8.37
episode: 2599   score: 5.0   memory length: 657706   epsilon: 0.4478700700119862
steps: 292   lr: 1.0240000000000005e-06   evaluation reward: 8.35
episode: 2600   score: 3.0   memory length: 657935   epsilon: 0.4476433600119911
steps: 229   lr: 1.0240000000000005e-06   evaluation reward: 8.18
episode: 2601   score: 8.0   memory length: 658406   epsilon:
0.44717707001200124   steps: 471   lr: 1.0240000000000005e-06   evaluation
reward: 8.18
episode: 2602   score: 8.0   memory length: 658839   epsilon:
0.44674840001201055   steps: 433   lr: 1.0240000000000005e-06   evaluation
reward: 8.18
episode: 2603   score: 9.0   memory length: 659341   epsilon:
0.44625142001202134   steps: 502   lr: 1.0240000000000005e-06   evaluation
reward: 8.21
episode: 2604   score: 11.0  memory length: 659945   epsilon:
0.4456534600120343   steps: 604   lr: 1.0240000000000005e-06   evaluation
reward: 8.29
episode: 2605   score: 9.0   memory length: 660432   epsilon: 0.4451713300120448
steps: 487   lr: 1.0240000000000005e-06   evaluation reward: 8.3
episode: 2606   score: 7.0   memory length: 660788   epsilon:
0.44481889001205244   steps: 356   lr: 1.0240000000000005e-06   evaluation
reward: 8.28
episode: 2607   score: 15.0  memory length: 661419   epsilon: 0.444194200012066
steps: 631   lr: 1.0240000000000005e-06   evaluation reward: 8.35
episode: 2608   score: 7.0   memory length: 661793   epsilon:
0.44382394001207404   steps: 374   lr: 1.0240000000000005e-06   evaluation
reward: 8.36
episode: 2609   score: 9.0   memory length: 662256   epsilon: 0.443365570012084
steps: 463   lr: 1.0240000000000005e-06   evaluation reward: 8.34
episode: 2610   score: 11.0  memory length: 662806   epsilon:
0.4428210700120958   steps: 550   lr: 1.0240000000000005e-06   evaluation
reward: 8.35
episode: 2611   score: 6.0   memory length: 663139   epsilon:
0.44249140001210296   steps: 333   lr: 1.0240000000000005e-06   evaluation
reward: 8.27
episode: 2612   score: 2.0   memory length: 663321   epsilon: 0.4423112200121069
steps: 182   lr: 1.0240000000000005e-06   evaluation reward: 8.14
episode: 2613   score: 4.0   memory length: 663563   epsilon: 0.4420716400121121
steps: 242   lr: 1.0240000000000005e-06   evaluation reward: 8.09

```

episode: 2614 score: 15.0 memory length: 664107 epsilon:
 0.44153308001212377 steps: 544 lr: 1.0240000000000005e-06 evaluation
 reward: 8.13
 episode: 2615 score: 5.0 memory length: 664394 epsilon:
 0.44124895001212994 steps: 287 lr: 1.0240000000000005e-06 evaluation
 reward: 8.12
 episode: 2616 score: 9.0 memory length: 664881 epsilon: 0.4407668200121404
 steps: 487 lr: 1.0240000000000005e-06 evaluation reward: 8.17
 episode: 2617 score: 7.0 memory length: 665231 epsilon: 0.4404203200121479
 steps: 350 lr: 1.0240000000000005e-06 evaluation reward: 8.16
 episode: 2618 score: 5.0 memory length: 665516 epsilon:
 0.44013817001215405 steps: 285 lr: 1.0240000000000005e-06 evaluation
 reward: 8.13
 episode: 2619 score: 7.0 memory length: 665890 epsilon: 0.4397679100121621
 steps: 374 lr: 1.0240000000000005e-06 evaluation reward: 8.12
 episode: 2620 score: 7.0 memory length: 666310 epsilon: 0.4393521100121711
 steps: 420 lr: 1.0240000000000005e-06 evaluation reward: 8.1
 episode: 2621 score: 4.0 memory length: 666569 epsilon: 0.4390957000121767
 steps: 259 lr: 1.0240000000000005e-06 evaluation reward: 8.06
 episode: 2622 score: 12.0 memory length: 667069 epsilon:
 0.4386007000121874 steps: 500 lr: 1.0240000000000005e-06 evaluation
 reward: 8.11
 episode: 2623 score: 8.0 memory length: 667530 epsilon:
 0.43814431001219734 steps: 461 lr: 1.0240000000000005e-06 evaluation
 reward: 8.14
 episode: 2624 score: 8.0 memory length: 667953 epsilon: 0.4377255400122064
 steps: 423 lr: 1.0240000000000005e-06 evaluation reward: 8.18
 episode: 2625 score: 9.0 memory length: 668452 epsilon:
 0.43723153001221715 steps: 499 lr: 1.0240000000000005e-06 evaluation
 reward: 8.22
 episode: 2626 score: 12.0 memory length: 668894 epsilon:
 0.43679395001222665 steps: 442 lr: 1.0240000000000005e-06 evaluation
 reward: 8.24
 episode: 2627 score: 10.0 memory length: 669382 epsilon:
 0.43631083001223714 steps: 488 lr: 1.0240000000000005e-06 evaluation
 reward: 8.23
 episode: 2628 score: 5.0 memory length: 669688 epsilon: 0.4360078900122437
 steps: 306 lr: 1.0240000000000005e-06 evaluation reward: 8.2
 episode: 2629 score: 9.0 memory length: 670124 epsilon: 0.4355762500122531
 steps: 436 lr: 1.0240000000000005e-06 evaluation reward: 8.2
 episode: 2630 score: 8.0 memory length: 670558 epsilon: 0.4351465900122624
 steps: 434 lr: 1.0240000000000005e-06 evaluation reward: 8.17
 episode: 2631 score: 11.0 memory length: 671110 epsilon:
 0.4346001100122743 steps: 552 lr: 1.0240000000000005e-06 evaluation
 reward: 8.12
 episode: 2632 score: 5.0 memory length: 671416 epsilon:
 0.43429717001228085 steps: 306 lr: 1.0240000000000005e-06 evaluation
 reward: 8.11

episode: 2633 score: 5.0 memory length: 671729 epsilon: 0.4339873000122876
 steps: 313 lr: 1.0240000000000005e-06 evaluation reward: 8.08
 episode: 2634 score: 8.0 memory length: 672132 epsilon:
 0.43358833001229624 steps: 403 lr: 1.0240000000000005e-06 evaluation
 reward: 8.09
 episode: 2635 score: 8.0 memory length: 672598 epsilon:
 0.43312699001230626 steps: 466 lr: 1.0240000000000005e-06 evaluation
 reward: 8.1
 episode: 2636 score: 14.0 memory length: 673162 epsilon:
 0.4325686300123184 steps: 564 lr: 1.0240000000000005e-06 evaluation
 reward: 8.12
 episode: 2637 score: 5.0 memory length: 673485 epsilon: 0.4322488600123253
 steps: 323 lr: 1.0240000000000005e-06 evaluation reward: 8.02
 episode: 2638 score: 9.0 memory length: 674001 epsilon: 0.4317380200123364
 steps: 516 lr: 1.0240000000000005e-06 evaluation reward: 7.98
 episode: 2639 score: 8.0 memory length: 674429 epsilon: 0.4313143000123456
 steps: 428 lr: 1.0240000000000005e-06 evaluation reward: 7.99
 episode: 2640 score: 8.0 memory length: 674907 epsilon: 0.4308410800123559
 steps: 478 lr: 1.0240000000000005e-06 evaluation reward: 8.03
 episode: 2641 score: 4.0 memory length: 675169 epsilon: 0.4305817000123615
 steps: 262 lr: 1.0240000000000005e-06 evaluation reward: 8.03
 episode: 2642 score: 6.0 memory length: 675492 epsilon:
 0.43026193001236845 steps: 323 lr: 1.0240000000000005e-06 evaluation
 reward: 8.05
 episode: 2643 score: 9.0 memory length: 675982 epsilon: 0.429776830012379
 steps: 490 lr: 1.0240000000000005e-06 evaluation reward: 8.07
 episode: 2644 score: 8.0 memory length: 676428 epsilon:
 0.42933529001238857 steps: 446 lr: 1.0240000000000005e-06 evaluation
 reward: 8.1
 episode: 2645 score: 3.0 memory length: 676657 epsilon: 0.4291085800123935
 steps: 229 lr: 1.0240000000000005e-06 evaluation reward: 8.1
 episode: 2646 score: 6.0 memory length: 677016 epsilon: 0.4287531700124012
 steps: 359 lr: 1.0240000000000005e-06 evaluation reward: 8.11
 episode: 2647 score: 3.0 memory length: 677248 epsilon: 0.4285234900124062
 steps: 232 lr: 1.0240000000000005e-06 evaluation reward: 8.07
 episode: 2648 score: 11.0 memory length: 677797 epsilon: 0.427979980012418
 steps: 549 lr: 1.0240000000000005e-06 evaluation reward: 8.12
 episode: 2649 score: 16.0 memory length: 678400 epsilon:
 0.42738301001243095 steps: 603 lr: 1.0240000000000005e-06 evaluation
 reward: 8.2
 episode: 2650 score: 7.0 memory length: 678787 epsilon:
 0.42699988001243927 steps: 387 lr: 1.0240000000000005e-06 evaluation
 reward: 8.16
 episode: 2651 score: 6.0 memory length: 679151 epsilon: 0.4266395200124471
 steps: 364 lr: 1.0240000000000005e-06 evaluation reward: 8.15
 episode: 2652 score: 6.0 memory length: 679465 epsilon:
 0.42632866001245384 steps: 314 lr: 1.0240000000000005e-06 evaluation
 reward: 8.13

episode: 2653 score: 7.0 memory length: 679868 epsilon: 0.4259296900124625
 steps: 403 lr: 1.0240000000000005e-06 evaluation reward: 8.14
 episode: 2654 score: 10.0 memory length: 680366 epsilon:
 0.4254366700124732 steps: 498 lr: 1.0240000000000005e-06 evaluation
 reward: 8.16
 episode: 2655 score: 3.0 memory length: 680614 epsilon:
 0.42519115001247854 steps: 248 lr: 1.0240000000000005e-06 evaluation
 reward: 8.12
 episode: 2656 score: 4.0 memory length: 680877 epsilon: 0.4249307800124842
 steps: 263 lr: 1.0240000000000005e-06 evaluation reward: 8.06
 episode: 2657 score: 9.0 memory length: 681352 epsilon: 0.4244605300124944
 steps: 475 lr: 1.0240000000000005e-06 evaluation reward: 8.09
 episode: 2658 score: 7.0 memory length: 681698 epsilon:
 0.42411799001250183 steps: 346 lr: 1.0240000000000005e-06 evaluation
 reward: 8.12
 episode: 2659 score: 4.0 memory length: 681939 epsilon: 0.423879400012507
 steps: 241 lr: 1.0240000000000005e-06 evaluation reward: 8.02
 episode: 2660 score: 11.0 memory length: 682485 epsilon:
 0.42333886001251875 steps: 546 lr: 1.0240000000000005e-06 evaluation
 reward: 8.06
 episode: 2661 score: 9.0 memory length: 682938 epsilon: 0.4228903900125285
 steps: 453 lr: 1.0240000000000005e-06 evaluation reward: 8.11
 episode: 2662 score: 7.0 memory length: 683309 epsilon:
 0.42252310001253646 steps: 371 lr: 1.0240000000000005e-06 evaluation
 reward: 8.13
 episode: 2663 score: 7.0 memory length: 683717 epsilon: 0.4221191800125452
 steps: 408 lr: 1.0240000000000005e-06 evaluation reward: 8.12
 episode: 2664 score: 7.0 memory length: 684125 epsilon: 0.421715260012554
 steps: 408 lr: 1.0240000000000005e-06 evaluation reward: 8.13
 episode: 2665 score: 8.0 memory length: 684535 epsilon: 0.4213093600125628
 steps: 410 lr: 1.0240000000000005e-06 evaluation reward: 8.14
 episode: 2666 score: 17.0 memory length: 685166 epsilon:
 0.42068467001257637 steps: 631 lr: 1.0240000000000005e-06 evaluation
 reward: 8.21
 episode: 2667 score: 4.0 memory length: 685427 epsilon: 0.420426280012582
 steps: 261 lr: 1.0240000000000005e-06 evaluation reward: 8.15
 episode: 2668 score: 10.0 memory length: 685951 epsilon:
 0.41990752001259324 steps: 524 lr: 1.0240000000000005e-06 evaluation
 reward: 8.19
 episode: 2669 score: 3.0 memory length: 686164 epsilon: 0.4196966500125978
 steps: 213 lr: 1.0240000000000005e-06 evaluation reward: 8.14
 episode: 2670 score: 7.0 memory length: 686551 epsilon:
 0.41931352001260613 steps: 387 lr: 1.0240000000000005e-06 evaluation
 reward: 8.1
 episode: 2671 score: 12.0 memory length: 686996 epsilon:
 0.4188729700126157 steps: 445 lr: 1.0240000000000005e-06 evaluation
 reward: 8.16
 episode: 2672 score: 10.0 memory length: 687503 epsilon:

0.4183710400126266 steps: 507 lr: 1.0240000000000005e-06 evaluation
 reward: 8.14
 episode: 2673 score: 12.0 memory length: 688040 epsilon:
 0.41783941001263813 steps: 537 lr: 1.0240000000000005e-06 evaluation
 reward: 8.18
 episode: 2674 score: 10.0 memory length: 688515 epsilon:
 0.41736916001264834 steps: 475 lr: 1.0240000000000005e-06 evaluation
 reward: 8.24
 episode: 2675 score: 11.0 memory length: 689071 epsilon:
 0.4168187200126603 steps: 556 lr: 1.0240000000000005e-06 evaluation
 reward: 8.24
 episode: 2676 score: 9.0 memory length: 689560 epsilon: 0.4163346100126708
 steps: 489 lr: 1.0240000000000005e-06 evaluation reward: 8.26
 episode: 2677 score: 5.0 memory length: 689849 epsilon: 0.416048500012677
 steps: 289 lr: 1.0240000000000005e-06 evaluation reward: 8.16
 episode: 2678 score: 14.0 memory length: 690355 epsilon:
 0.4155475600126879 steps: 506 lr: 1.0240000000000005e-06 evaluation
 reward: 8.22
 episode: 2679 score: 7.0 memory length: 690777 epsilon:
 0.41512978001269696 steps: 422 lr: 1.0240000000000005e-06 evaluation
 reward: 8.18
 episode: 2680 score: 8.0 memory length: 691203 epsilon: 0.4147080400127061
 steps: 426 lr: 1.0240000000000005e-06 evaluation reward: 8.16
 episode: 2681 score: 5.0 memory length: 691512 epsilon:
 0.41440213001271275 steps: 309 lr: 1.0240000000000005e-06 evaluation
 reward: 8.12
 episode: 2682 score: 4.0 memory length: 691756 epsilon: 0.414160570012718
 steps: 244 lr: 1.0240000000000005e-06 evaluation reward: 8.08
 episode: 2683 score: 13.0 memory length: 692213 epsilon:
 0.4137081400127278 steps: 457 lr: 1.0240000000000005e-06 evaluation
 reward: 8.14
 episode: 2684 score: 5.0 memory length: 692537 epsilon: 0.4133873800127348
 steps: 324 lr: 1.0240000000000005e-06 evaluation reward: 8.11
 episode: 2685 score: 7.0 memory length: 692883 epsilon: 0.4130448400127422
 steps: 346 lr: 1.0240000000000005e-06 evaluation reward: 8.09
 episode: 2686 score: 5.0 memory length: 693173 epsilon:
 0.41275774001274845 steps: 290 lr: 1.0240000000000005e-06 evaluation
 reward: 8.11
 episode: 2687 score: 7.0 memory length: 693621 epsilon: 0.4123142200127581
 steps: 448 lr: 1.0240000000000005e-06 evaluation reward: 8.1
 episode: 2688 score: 8.0 memory length: 694038 epsilon:
 0.41190139001276704 steps: 417 lr: 1.0240000000000005e-06 evaluation
 reward: 8.08
 episode: 2689 score: 8.0 memory length: 694437 epsilon: 0.4115063800127756
 steps: 399 lr: 1.0240000000000005e-06 evaluation reward: 8.05
 episode: 2690 score: 9.0 memory length: 694899 epsilon:
 0.41104900001278555 steps: 462 lr: 1.0240000000000005e-06 evaluation
 reward: 8.03

episode: 2691 score: 12.0 memory length: 695501 epsilon:
 0.4104530200127985 steps: 602 lr: 1.0240000000000005e-06 evaluation
 reward: 8.07
 episode: 2692 score: 6.0 memory length: 695827 epsilon: 0.4101302800128055
 steps: 326 lr: 1.0240000000000005e-06 evaluation reward: 7.99
 episode: 2693 score: 11.0 memory length: 696362 epsilon: 0.409600630012817
 steps: 535 lr: 1.0240000000000005e-06 evaluation reward: 8.02
 episode: 2694 score: 7.0 memory length: 696699 epsilon:
 0.40926700001282423 steps: 337 lr: 1.0240000000000005e-06 evaluation
 reward: 7.96
 episode: 2695 score: 14.0 memory length: 697250 epsilon:
 0.4087215100128361 steps: 551 lr: 1.0240000000000005e-06 evaluation
 reward: 8.04
 episode: 2696 score: 7.0 memory length: 697608 epsilon:
 0.40836709001284377 steps: 358 lr: 1.0240000000000005e-06 evaluation
 reward: 8.05
 episode: 2697 score: 10.0 memory length: 698170 epsilon:
 0.40781071001285585 steps: 562 lr: 1.0240000000000005e-06 evaluation
 reward: 8.03
 episode: 2698 score: 6.0 memory length: 698491 epsilon:
 0.40749292001286275 steps: 321 lr: 1.0240000000000005e-06 evaluation
 reward: 7.96
 episode: 2699 score: 6.0 memory length: 698821 epsilon:
 0.40716622001286984 steps: 330 lr: 1.0240000000000005e-06 evaluation
 reward: 7.97
 episode: 2700 score: 7.0 memory length: 699187 epsilon: 0.4068038800128777
 steps: 366 lr: 1.0240000000000005e-06 evaluation reward: 8.01
 episode: 2701 score: 8.0 memory length: 699604 epsilon:
 0.40639105001288667 steps: 417 lr: 1.0240000000000005e-06 evaluation
 reward: 8.01
 episode: 2702 score: 16.0 memory length: 700090 epsilon:
 0.4059099100128971 steps: 486 lr: 4.09600000000000023e-07 evaluation
 reward: 8.09
 episode: 2703 score: 13.0 memory length: 700555 epsilon:
 0.4054495600129071 steps: 465 lr: 4.09600000000000023e-07 evaluation
 reward: 8.13
 episode: 2704 score: 11.0 memory length: 701106 epsilon:
 0.40490407001291895 steps: 551 lr: 4.09600000000000023e-07 evaluation
 reward: 8.13
 episode: 2705 score: 6.0 memory length: 701485 epsilon: 0.4045288600129271
 steps: 379 lr: 4.09600000000000023e-07 evaluation reward: 8.1
 episode: 2706 score: 11.0 memory length: 702017 epsilon:
 0.4040021800129385 steps: 532 lr: 4.09600000000000023e-07 evaluation
 reward: 8.14
 episode: 2707 score: 10.0 memory length: 702532 epsilon:
 0.4034923300129496 steps: 515 lr: 4.09600000000000023e-07 evaluation
 reward: 8.09
 episode: 2708 score: 8.0 memory length: 702935 epsilon:

0.40309336001295826 steps: 403 lr: 4.09600000000000023e-07 evaluation
 reward: 8.1
 episode: 2709 score: 10.0 memory length: 703407 epsilon:
 0.4026260800129684 steps: 472 lr: 4.09600000000000023e-07 evaluation
 reward: 8.11
 episode: 2710 score: 4.0 memory length: 703668 epsilon: 0.402367690012974
 steps: 261 lr: 4.09600000000000023e-07 evaluation reward: 8.04
 episode: 2711 score: 17.0 memory length: 704226 epsilon: 0.401815270012986
 steps: 558 lr: 4.09600000000000023e-07 evaluation reward: 8.15
 episode: 2712 score: 14.0 memory length: 704780 epsilon:
 0.4012668100129979 steps: 554 lr: 4.09600000000000023e-07 evaluation
 reward: 8.27
 episode: 2713 score: 8.0 memory length: 705202 epsilon: 0.400849030013007
 steps: 422 lr: 4.09600000000000023e-07 evaluation reward: 8.31
 episode: 2714 score: 10.0 memory length: 705693 epsilon:
 0.40036294001301753 steps: 491 lr: 4.09600000000000023e-07 evaluation
 reward: 8.26
 episode: 2715 score: 5.0 memory length: 705964 epsilon:
 0.40009465001302336 steps: 271 lr: 4.09600000000000023e-07 evaluation
 reward: 8.26
 episode: 2716 score: 13.0 memory length: 706520 epsilon:
 0.3995442100130353 steps: 556 lr: 4.09600000000000023e-07 evaluation
 reward: 8.3
 episode: 2717 score: 13.0 memory length: 707146 epsilon:
 0.39892447001304876 steps: 626 lr: 4.09600000000000023e-07 evaluation
 reward: 8.36
 episode: 2718 score: 6.0 memory length: 707504 epsilon:
 0.39857005001305645 steps: 358 lr: 4.09600000000000023e-07 evaluation
 reward: 8.37
 episode: 2719 score: 10.0 memory length: 708011 epsilon:
 0.39806812001306735 steps: 507 lr: 4.09600000000000023e-07 evaluation
 reward: 8.4
 episode: 2720 score: 12.0 memory length: 708613 epsilon:
 0.3974721400130803 steps: 602 lr: 4.09600000000000023e-07 evaluation
 reward: 8.45
 episode: 2721 score: 11.0 memory length: 709172 epsilon:
 0.3969187300130923 steps: 559 lr: 4.09600000000000023e-07 evaluation
 reward: 8.52
 episode: 2722 score: 10.0 memory length: 709667 epsilon:
 0.39642868001310294 steps: 495 lr: 4.09600000000000023e-07 evaluation
 reward: 8.5
 episode: 2723 score: 7.0 memory length: 710035 epsilon:
 0.39606436001311085 steps: 368 lr: 4.09600000000000023e-07 evaluation
 reward: 8.49
 episode: 2724 score: 5.0 memory length: 710344 epsilon: 0.3957584500131175
 steps: 309 lr: 4.09600000000000023e-07 evaluation reward: 8.46
 episode: 2725 score: 6.0 memory length: 710679 epsilon: 0.3954268000131247
 steps: 335 lr: 4.09600000000000023e-07 evaluation reward: 8.43

episode: 2726 score: 10.0 memory length: 711146 epsilon:
 0.3949644700131347 steps: 467 lr: 4.09600000000000023e-07 evaluation
 reward: 8.41
 episode: 2727 score: 5.0 memory length: 711418 epsilon: 0.3946951900131406
 steps: 272 lr: 4.09600000000000023e-07 evaluation reward: 8.36
 episode: 2728 score: 9.0 memory length: 711905 epsilon:
 0.39421306001315104 steps: 487 lr: 4.09600000000000023e-07 evaluation
 reward: 8.4
 episode: 2729 score: 4.0 memory length: 712165 epsilon: 0.3939556600131566
 steps: 260 lr: 4.09600000000000023e-07 evaluation reward: 8.35
 episode: 2730 score: 5.0 memory length: 712491 epsilon:
 0.39363292001316363 steps: 326 lr: 4.09600000000000023e-07 evaluation
 reward: 8.32
 episode: 2731 score: 4.0 memory length: 712753 epsilon:
 0.39337354001316926 steps: 262 lr: 4.09600000000000023e-07 evaluation
 reward: 8.25
 episode: 2732 score: 11.0 memory length: 713282 epsilon:
 0.39284983001318063 steps: 529 lr: 4.09600000000000023e-07 evaluation
 reward: 8.31
 episode: 2733 score: 9.0 memory length: 713763 epsilon:
 0.39237364001319097 steps: 481 lr: 4.09600000000000023e-07 evaluation
 reward: 8.35
 episode: 2734 score: 10.0 memory length: 714271 epsilon:
 0.3918707200132019 steps: 508 lr: 4.09600000000000023e-07 evaluation
 reward: 8.37
 episode: 2735 score: 7.0 memory length: 714665 epsilon:
 0.39148066001321036 steps: 394 lr: 4.09600000000000023e-07 evaluation
 reward: 8.36
 episode: 2736 score: 14.0 memory length: 715299 epsilon: 0.390853000013224
 steps: 634 lr: 4.09600000000000023e-07 evaluation reward: 8.36
 episode: 2737 score: 6.0 memory length: 715623 epsilon:
 0.39053224001323095 steps: 324 lr: 4.09600000000000023e-07 evaluation
 reward: 8.37
 episode: 2738 score: 8.0 memory length: 716029 epsilon: 0.3901303000132397
 steps: 406 lr: 4.09600000000000023e-07 evaluation reward: 8.36
 episode: 2739 score: 10.0 memory length: 716554 epsilon:
 0.38961055001325096 steps: 525 lr: 4.09600000000000023e-07 evaluation
 reward: 8.38
 episode: 2740 score: 12.0 memory length: 717131 epsilon:
 0.38903932001326336 steps: 577 lr: 4.09600000000000023e-07 evaluation
 reward: 8.42
 episode: 2741 score: 13.0 memory length: 717761 epsilon:
 0.3884156200132769 steps: 630 lr: 4.09600000000000023e-07 evaluation
 reward: 8.51
 episode: 2742 score: 5.0 memory length: 718085 epsilon:
 0.38809486001328386 steps: 324 lr: 4.09600000000000023e-07 evaluation
 reward: 8.5
 episode: 2743 score: 8.0 memory length: 718557 epsilon: 0.387627580013294


```

steps: 472    lr: 4.09600000000000023e-07    evaluation reward: 8.49
episode: 2744    score: 7.0    memory length: 718949    epsilon:
0.38723950001330243    steps: 392    lr: 4.09600000000000023e-07    evaluation
reward: 8.48
episode: 2745    score: 12.0    memory length: 719521    epsilon:
0.3866732200133147    steps: 572    lr: 4.09600000000000023e-07    evaluation
reward: 8.57
episode: 2746    score: 5.0    memory length: 719809    epsilon: 0.3863881000133209
steps: 288    lr: 4.09600000000000023e-07    evaluation reward: 8.56
episode: 2747    score: 14.0    memory length: 720339    epsilon:
0.3858634000133323    steps: 530    lr: 4.09600000000000023e-07    evaluation
reward: 8.67
episode: 2748    score: 9.0    memory length: 720792    epsilon:
0.38541493001334204    steps: 453    lr: 4.09600000000000023e-07    evaluation
reward: 8.65
episode: 2749    score: 6.0    memory length: 721100    epsilon:
0.38511001001334866    steps: 308    lr: 4.09600000000000023e-07    evaluation
reward: 8.55
episode: 2750    score: 7.0    memory length: 721507    epsilon: 0.3847070800133574
steps: 407    lr: 4.09600000000000023e-07    evaluation reward: 8.55
episode: 2751    score: 5.0    memory length: 721788    epsilon:
0.38442889001336344    steps: 281    lr: 4.09600000000000023e-07    evaluation
reward: 8.54
episode: 2752    score: 9.0    memory length: 722289    epsilon: 0.3839329000133742
steps: 501    lr: 4.09600000000000023e-07    evaluation reward: 8.57
episode: 2753    score: 5.0    memory length: 722562    epsilon: 0.3836626300133801
steps: 273    lr: 4.09600000000000023e-07    evaluation reward: 8.55
episode: 2754    score: 13.0    memory length: 723022    epsilon:
0.38320723001338997    steps: 460    lr: 4.09600000000000023e-07    evaluation
reward: 8.58
episode: 2755    score: 10.0    memory length: 723522    epsilon:
0.3827122300134007    steps: 500    lr: 4.09600000000000023e-07    evaluation
reward: 8.65
episode: 2756    score: 8.0    memory length: 723928    epsilon:
0.38231029001340944    steps: 406    lr: 4.09600000000000023e-07    evaluation
reward: 8.69
episode: 2757    score: 9.0    memory length: 724378    epsilon: 0.3818647900134191
steps: 450    lr: 4.09600000000000023e-07    evaluation reward: 8.69
episode: 2758    score: 12.0    memory length: 724940    epsilon:
0.3813084100134312    steps: 562    lr: 4.09600000000000023e-07    evaluation
reward: 8.74
episode: 2759    score: 9.0    memory length: 725376    epsilon:
0.38087677001344056    steps: 436    lr: 4.09600000000000023e-07    evaluation
reward: 8.79
episode: 2760    score: 9.0    memory length: 725865    epsilon:
0.38039266001345107    steps: 489    lr: 4.09600000000000023e-07    evaluation
reward: 8.77
episode: 2761    score: 8.0    memory length: 726310    epsilon:

```

0.37995211001346063 steps: 445 lr: 4.09600000000000023e-07 evaluation
 reward: 8.76
 episode: 2762 score: 9.0 memory length: 726746 epsilon: 0.37952047001347
 steps: 436 lr: 4.09600000000000023e-07 evaluation reward: 8.78
 episode: 2763 score: 12.0 memory length: 727297 epsilon:
 0.37897498001348184 steps: 551 lr: 4.09600000000000023e-07 evaluation
 reward: 8.83
 episode: 2764 score: 9.0 memory length: 727775 epsilon: 0.3785017600134921
 steps: 478 lr: 4.09600000000000023e-07 evaluation reward: 8.85
 episode: 2765 score: 10.0 memory length: 728215 epsilon:
 0.3780661600135016 steps: 440 lr: 4.09600000000000023e-07 evaluation
 reward: 8.87
 episode: 2766 score: 12.0 memory length: 728801 epsilon:
 0.37748602001351417 steps: 586 lr: 4.09600000000000023e-07 evaluation
 reward: 8.82
 episode: 2767 score: 3.0 memory length: 729029 epsilon:
 0.37726030001351907 steps: 228 lr: 4.09600000000000023e-07 evaluation
 reward: 8.81
 episode: 2768 score: 8.0 memory length: 729480 epsilon:
 0.37681381001352876 steps: 451 lr: 4.09600000000000023e-07 evaluation
 reward: 8.79
 episode: 2769 score: 9.0 memory length: 729980 epsilon: 0.3763188100135395
 steps: 500 lr: 4.09600000000000023e-07 evaluation reward: 8.85
 episode: 2770 score: 6.0 memory length: 730321 epsilon:
 0.37598122001354684 steps: 341 lr: 4.09600000000000023e-07 evaluation
 reward: 8.84
 episode: 2771 score: 15.0 memory length: 731038 epsilon:
 0.37527139001356224 steps: 717 lr: 4.09600000000000023e-07 evaluation
 reward: 8.87
 episode: 2772 score: 9.0 memory length: 731491 epsilon: 0.374822920013572
 steps: 453 lr: 4.09600000000000023e-07 evaluation reward: 8.86
 episode: 2773 score: 20.0 memory length: 732109 epsilon:
 0.37421110001358526 steps: 618 lr: 4.09600000000000023e-07 evaluation
 reward: 8.94
 episode: 2774 score: 7.0 memory length: 732502 epsilon: 0.3738220300135937
 steps: 393 lr: 4.09600000000000023e-07 evaluation reward: 8.91
 episode: 2775 score: 6.0 memory length: 732857 epsilon:
 0.37347058001360134 steps: 355 lr: 4.09600000000000023e-07 evaluation
 reward: 8.86
 episode: 2776 score: 10.0 memory length: 733369 epsilon:
 0.37296370001361234 steps: 512 lr: 4.09600000000000023e-07 evaluation
 reward: 8.87
 episode: 2777 score: 9.0 memory length: 733864 epsilon: 0.372473650013623
 steps: 495 lr: 4.09600000000000023e-07 evaluation reward: 8.91
 episode: 2778 score: 11.0 memory length: 734408 epsilon:
 0.3719350900136347 steps: 544 lr: 4.09600000000000023e-07 evaluation
 reward: 8.88
 episode: 2779 score: 11.0 memory length: 734976 epsilon:

0.3713727700136469 steps: 568 lr: 4.09600000000000023e-07 evaluation
 reward: 8.92
 episode: 2780 score: 4.0 memory length: 735220 epsilon: 0.3711312100136521
 steps: 244 lr: 4.09600000000000023e-07 evaluation reward: 8.88
 episode: 2781 score: 20.0 memory length: 735927 epsilon:
 0.3704312800136673 steps: 707 lr: 4.09600000000000023e-07 evaluation
 reward: 9.03
 episode: 2782 score: 10.0 memory length: 736388 epsilon:
 0.3699748900136772 steps: 461 lr: 4.09600000000000023e-07 evaluation
 reward: 9.09
 episode: 2783 score: 7.0 memory length: 736753 epsilon:
 0.36961354001368507 steps: 365 lr: 4.09600000000000023e-07 evaluation
 reward: 9.03
 episode: 2784 score: 5.0 memory length: 737098 epsilon: 0.3692719900136925
 steps: 345 lr: 4.09600000000000023e-07 evaluation reward: 9.03
 episode: 2785 score: 5.0 memory length: 737425 epsilon: 0.3689482600136995
 steps: 327 lr: 4.09600000000000023e-07 evaluation reward: 9.01
 episode: 2786 score: 7.0 memory length: 737835 epsilon: 0.3685423600137083
 steps: 410 lr: 4.09600000000000023e-07 evaluation reward: 9.03
 episode: 2787 score: 7.0 memory length: 738263 epsilon: 0.3681186400137175
 steps: 428 lr: 4.09600000000000023e-07 evaluation reward: 9.03
 episode: 2788 score: 15.0 memory length: 738885 epsilon:
 0.3675028600137309 steps: 622 lr: 4.09600000000000023e-07 evaluation
 reward: 9.1
 episode: 2789 score: 12.0 memory length: 739479 epsilon:
 0.36691480001374366 steps: 594 lr: 4.09600000000000023e-07 evaluation
 reward: 9.14
 episode: 2790 score: 10.0 memory length: 739993 epsilon:
 0.3664059400137547 steps: 514 lr: 4.09600000000000023e-07 evaluation
 reward: 9.15
 episode: 2791 score: 8.0 memory length: 740419 epsilon:
 0.36598420001376386 steps: 426 lr: 4.09600000000000023e-07 evaluation
 reward: 9.11
 episode: 2792 score: 8.0 memory length: 740785 epsilon: 0.3656218600137717
 steps: 366 lr: 4.09600000000000023e-07 evaluation reward: 9.13
 episode: 2793 score: 12.0 memory length: 741322 epsilon:
 0.36509023001378327 steps: 537 lr: 4.09600000000000023e-07 evaluation
 reward: 9.14
 episode: 2794 score: 16.0 memory length: 741895 epsilon:
 0.3645229600137956 steps: 573 lr: 4.09600000000000023e-07 evaluation
 reward: 9.23
 episode: 2795 score: 22.0 memory length: 742637 epsilon:
 0.36378838001381153 steps: 742 lr: 4.09600000000000023e-07 evaluation
 reward: 9.31
 episode: 2796 score: 9.0 memory length: 743064 epsilon: 0.3633656500138207
 steps: 427 lr: 4.09600000000000023e-07 evaluation reward: 9.33
 episode: 2797 score: 11.0 memory length: 743665 epsilon:
 0.3627706600138336 steps: 601 lr: 4.09600000000000023e-07 evaluation

```

reward: 9.34
episode: 2798   score: 10.0   memory length: 744139   epsilon:
0.3623014000138438   steps: 474   lr: 4.09600000000000023e-07   evaluation
reward: 9.38
episode: 2799   score: 10.0   memory length: 744659   epsilon: 0.361786600013855
steps: 520   lr: 4.09600000000000023e-07   evaluation reward: 9.42
episode: 2800   score: 7.0   memory length: 745027   epsilon: 0.3614222800138629
steps: 368   lr: 4.09600000000000023e-07   evaluation reward: 9.42
episode: 2801   score: 9.0   memory length: 745547   epsilon:
0.36090748001387407   steps: 520   lr: 4.09600000000000023e-07   evaluation
reward: 9.43
episode: 2802   score: 11.0   memory length: 746074   epsilon:
0.3603857500138854   steps: 527   lr: 4.09600000000000023e-07   evaluation
reward: 9.38
episode: 2803   score: 8.0   memory length: 746498   epsilon: 0.3599659900138945
steps: 424   lr: 4.09600000000000023e-07   evaluation reward: 9.33
episode: 2804   score: 13.0   memory length: 747019   epsilon:
0.3594502000139057   steps: 521   lr: 4.09600000000000023e-07   evaluation
reward: 9.35
episode: 2805   score: 19.0   memory length: 747591   epsilon: 0.358883920013918
steps: 572   lr: 4.09600000000000023e-07   evaluation reward: 9.48
episode: 2806   score: 6.0   memory length: 747937   epsilon:
0.35854138001392544   steps: 346   lr: 4.09600000000000023e-07   evaluation
reward: 9.43
episode: 2807   score: 9.0   memory length: 748432   epsilon: 0.3580513300139361
steps: 495   lr: 4.09600000000000023e-07   evaluation reward: 9.42
episode: 2808   score: 8.0   memory length: 748844   epsilon:
0.35764345001394493   steps: 412   lr: 4.09600000000000023e-07   evaluation
reward: 9.42
episode: 2809   score: 4.0   memory length: 749107   epsilon: 0.3573830800139506
steps: 263   lr: 4.09600000000000023e-07   evaluation reward: 9.36
episode: 2810   score: 13.0   memory length: 749564   epsilon:
0.3569306500139604   steps: 457   lr: 4.09600000000000023e-07   evaluation
reward: 9.45
episode: 2811   score: 9.0   memory length: 749986   epsilon: 0.3565128700139695
steps: 422   lr: 4.09600000000000023e-07   evaluation reward: 9.37
episode: 2812   score: 9.0   memory length: 750449   epsilon: 0.3560545000139794
steps: 463   lr: 4.09600000000000023e-07   evaluation reward: 9.32
episode: 2813   score: 10.0   memory length: 750925   epsilon:
0.35558326001398965   steps: 476   lr: 4.09600000000000023e-07   evaluation
reward: 9.34
episode: 2814   score: 9.0   memory length: 751268   epsilon: 0.355243690013997
steps: 343   lr: 4.09600000000000023e-07   evaluation reward: 9.33
episode: 2815   score: 8.0   memory length: 751688   epsilon:
0.35482789001400605   steps: 420   lr: 4.09600000000000023e-07   evaluation
reward: 9.36
episode: 2816   score: 6.0   memory length: 752042   epsilon:
0.35447743001401366   steps: 354   lr: 4.09600000000000023e-07   evaluation

```

```

reward: 9.29
episode: 2817   score: 12.0   memory length: 752461   epsilon:
0.35406262001402267   steps: 419   lr: 4.09600000000000023e-07   evaluation
reward: 9.28
episode: 2818   score: 6.0   memory length: 752787   epsilon: 0.3537398800140297
steps: 326   lr: 4.09600000000000023e-07   evaluation reward: 9.28
episode: 2819   score: 6.0   memory length: 753141   epsilon: 0.3533894200140373
steps: 354   lr: 4.09600000000000023e-07   evaluation reward: 9.24
episode: 2820   score: 7.0   memory length: 753490   epsilon: 0.3530439100140448
steps: 349   lr: 4.09600000000000023e-07   evaluation reward: 9.19
episode: 2821   score: 11.0   memory length: 754082   epsilon:
0.3524578300140575   steps: 592   lr: 4.09600000000000023e-07   evaluation
reward: 9.19
episode: 2822   score: 13.0   memory length: 754648   epsilon:
0.35189749001406967   steps: 566   lr: 4.09600000000000023e-07   evaluation
reward: 9.22
episode: 2823   score: 7.0   memory length: 755035   epsilon: 0.351514360014078
steps: 387   lr: 4.09600000000000023e-07   evaluation reward: 9.22
episode: 2824   score: 7.0   memory length: 755432   epsilon: 0.3511213300140865
steps: 397   lr: 4.09600000000000023e-07   evaluation reward: 9.24
episode: 2825   score: 9.0   memory length: 755891   epsilon: 0.3506669200140964
steps: 459   lr: 4.09600000000000023e-07   evaluation reward: 9.27
episode: 2826   score: 13.0   memory length: 756465   epsilon:
0.3500986600141087   steps: 574   lr: 4.09600000000000023e-07   evaluation
reward: 9.3
episode: 2827   score: 10.0   memory length: 756903   epsilon:
0.34966504001411813   steps: 438   lr: 4.09600000000000023e-07   evaluation
reward: 9.35
episode: 2828   score: 8.0   memory length: 757315   epsilon: 0.349257160014127
steps: 412   lr: 4.09600000000000023e-07   evaluation reward: 9.34
episode: 2829   score: 9.0   memory length: 757779   epsilon:
0.34879780001413696   steps: 464   lr: 4.09600000000000023e-07   evaluation
reward: 9.39
episode: 2830   score: 10.0   memory length: 758295   epsilon:
0.34828696001414805   steps: 516   lr: 4.09600000000000023e-07   evaluation
reward: 9.44
episode: 2831   score: 12.0   memory length: 758776   epsilon:
0.3478107700141584   steps: 481   lr: 4.09600000000000023e-07   evaluation
reward: 9.52
episode: 2832   score: 16.0   memory length: 759522   epsilon:
0.3470722300141744   steps: 746   lr: 4.09600000000000023e-07   evaluation
reward: 9.57
episode: 2833   score: 7.0   memory length: 759943   epsilon:
0.34665544001418347   steps: 421   lr: 4.09600000000000023e-07   evaluation
reward: 9.55
episode: 2834   score: 12.0   memory length: 760573   epsilon: 0.346031740014197
steps: 630   lr: 4.09600000000000023e-07   evaluation reward: 9.57
episode: 2835   score: 5.0   memory length: 760883   epsilon:

```

```

0.34572484001420367    steps: 310    lr: 4.09600000000000023e-07    evaluation
reward: 9.55
episode: 2836    score: 9.0    memory length: 761206    epsilon: 0.3454050700142106
steps: 323    lr: 4.09600000000000023e-07    evaluation reward: 9.5
episode: 2837    score: 8.0    memory length: 761598    epsilon:
0.34501699001421904    steps: 392    lr: 4.09600000000000023e-07    evaluation
reward: 9.52
episode: 2838    score: 15.0    memory length: 762215    epsilon:
0.3444061600142323    steps: 617    lr: 4.09600000000000023e-07    evaluation
reward: 9.59
episode: 2839    score: 8.0    memory length: 762609    epsilon:
0.34401610001424077    steps: 394    lr: 4.09600000000000023e-07    evaluation
reward: 9.57

```

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.

```

[8]: 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

```

```

[9]: agent.load_policy_net('save_model/breakout_ddqn.pth')
     agent.update_target_net()

```

```

[10]: 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()

```

<IPython.core.display.HTML object>

[10]: <pyvirtualdisplay.display.Display at 0x7f30495b4550>

[]: